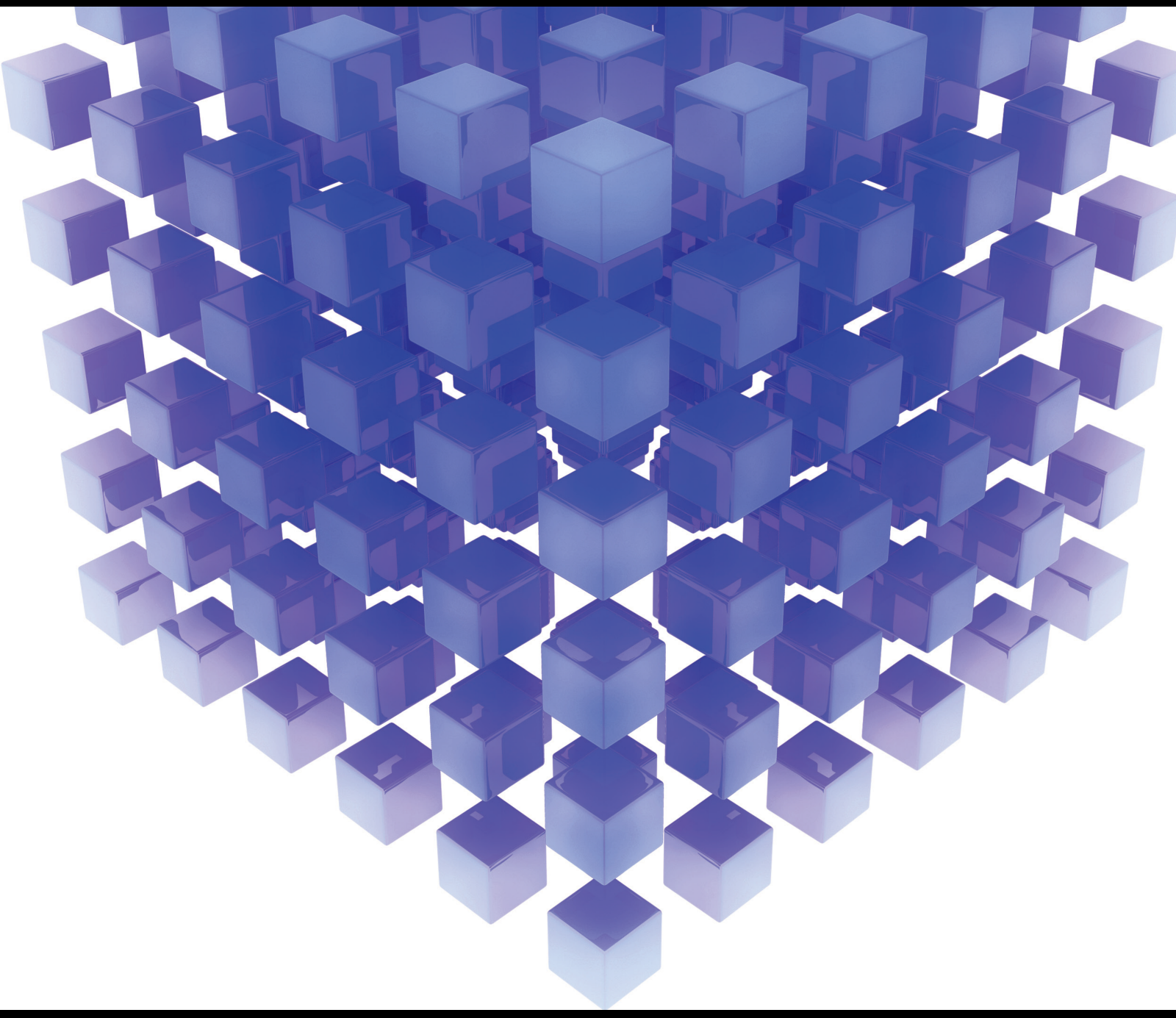
Mathematical Problems in Engineering

Recent Advances in Industrial Mathematics and Applications

Lead Guest Editor: Gengxin Sun
Guest Editors: Mingshun Yang and Hussein Abulkasim



Recent Advances in Industrial Mathematics and Applications

Mathematical Problems in Engineering

Recent Advances in Industrial Mathematics and Applications

Lead Guest Editor: Gengxin Sun
Guest Editors: Mingshun Yang and Hussein Abulkasim

Chief Editor

Academic Editors

Contents

Contents

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 4486311, 8 pages
https://doi.org/10.1155/2022/4486311

Research Article

Step-by-Step Penalized Blind Kriging Methods for Surrogate Modeling

Ziheng Feng,[1] **Min Li**,[1] **Xiao Wang**,[1] **Jianxin Zhao**,[2] **and Xinmin Li**[1]

[1]*School of Mathematics and Statistics, Qingdao University, Qingdao 266 071, China*
[2]*Basic Courses Department, Navy Submarine Academy, Qingdao, China*

Correspondence should be addressed to Xinmin Li; xmli@qdu.edu.cn

Received 17 May 2021; Revised 23 July 2021; Accepted 24 December 2021; Published 21 January 2022

Academic Editor: Suzanne M. Shontz

When building a surrogate model, it is important to identify a proper mean function for the kriging model. The commonly used variable selection method is the penalized blind kriging (PBK) method. But this method could lead to a low time efficiency, which is not suitable for experiments with time-sensitive data. In this paper, we propose three step-by-step approaches for constructing an appropriate mean function to improve the prediction accuracy and time efficiency of the PBK method. Several functions and two engineering examples are used to prove the effectiveness of the proposed methods. From simulation results, we can see that Method 1 (M1) and Method 2 (M2) have been significantly improved in both the prediction accuracy and the time efficiency compared with PBK. Especially, in the Test function, compared with the traditional PBK method, the prediction accuracy of M2 is improved by 69.08% and 26.13%, respectively, under the penalty of Lasso and Elastic Net, and the time efficiency of M1 is improved by 85.15% and 90.33%, respectively, under the penalty of Lasso and Elastic Net. In addition, Method 3 (M3) has been significantly improved in prediction accuracy compared with PBK.

1. Introduction

The surrogate model has been widely used to replace computation-intensive engineering simulation models or black-box systems [1–3]. There are several commonly used surrogate models, such as kriging, support vector regression, and radial basis functions [4]. Their application fields involve the structural dynamics of aeroengine casings, dynamic reliability estimation of turbine blisk, and computationally expensive constrained optimization problems.

The kriging model can be used to fit computer experiments because of its interpolate character. Its error estimation can be extended to many aspects such as failure probability estimation [5], robust design optimization [6], uncertainty analysis [7], and global sequential sampling [8]. As the combination of mean function and stochastic Gaussian process (GP), the kriging model is called ordinary kriging (OK) if using a constant mean to fit the overall trend or universal kriging (UK) if supposing some specified variables in the mean function. However, their prediction accuracy will be reduced if the mean function is specified inappropriately [9]. Computer experiments usually include a large number of input variables, inactive variables which hardly have an impact on the response tending to reduce prediction accuracy; thus, it is a matter to identify active variables into a mean function [10].

In this paper, we consider the problem of variable selection for the kriging model. We want to use the method of variable selection to construct an interpretable kriging model which can select the active variables and remove inactive variables with high prediction accuracy and high time efficiency. The commonly used variable selection method in the kriging model is the penalty likelihood method. For the penalty likelihood method, Hung [11] proposed the penalized blind kriging (PBK) approach to select variables. This method, however, suffers from the limitation of low time efficiency and prediction accuracy. In the field of engineering, because of the time sensitive of data, researchers always want to improve time efficiency greatly while maintaining preferable prediction accuracy. Therefore, efficient variable selection methods are called for.

The idea here is to improve the time efficiency and prediction accuracy of PBK using step-by-step approaches. Three new approaches are proposed for variable selection. The first method is to select the variables using a linear model at the first and then estimate the parameters in the kriging model. The second method is first to estimate the parameters with the OK model and then select the variables via the penalized kriging model and finally refit the kriging model. The third method is an improvement of PBK, which is a refitting based on the variable selection of PBK. These methods are proved to improve the prediction accuracy and time efficiency by several analytic functions and two examples.

The remainder of the paper is organized as follows: in Section 2, three new methods are proposed for variable selection. Section 3 illustrates the effectiveness of proposed methods with four simulated examples and compares the prediction accuracy, identification, and time efficiency. Section 4 implements the proposed methods for two engineering examples. Section 5 provides some discussion and concluding remarks.

2. Step-by-Step Variable Selection Methods

In this section, we first review some of the details; then, three methods are proposed for variable selection based on PBK methods.

2.1. Kriging Preliminaries. The kriging model [12] assumes that the response $y(x)$ is a realization from the stochastic process:

$$y(x) = \sum_{i=1}^{p} \beta_i f_i(x) + z(x) = f(x)^T \beta + z(x), \tag{1}$$

where $f(x) = (f_1(x), \ldots, f_p(x))^T$ is the basis functions, and $\beta = (\beta_1, \ldots, \beta_p)^T$ is a vector of unknown coefficients, respectively. The stochastic process $z(x)$ has mean 0 and covariance $\sigma^2 R$, where $R(\theta, x_i, x_j)$ is a correlation function between x_i and x_j depending on parameter θ. In most engineering applications, the most commonly used correlation function is the stationary Gaussian correlation function $R_{ij}(\theta) = R(\theta, x_i, x_j) = \prod_{k=1}^{d} \exp(-\theta_k (x_{ik} - x_{jk})^2)$, which is chosen in our research.

Given some sampled points $x_i \in R^d$ and the responses $y_i \in R, i = 1, \ldots, n$, the best linear unbiased predictor (BLUP) can be obtained as follows [12]:

$$\widehat{y}(x) = f(x)^T \widehat{\beta} + r(x)^T R^{-1} (y - F\widehat{\beta}), \tag{2}$$

where $r = [R(\theta, x_1, x), \ldots, R(\theta, x_n, x)]^T$ is a correlation between the unsampled point x and the sampled points $x_i, i = 1, \ldots, n$; the matrix $F = [f(x_1), \ldots, f(x_n)]^T$ is constructed by $f(x)$ at the sampled points.

The log-likelihood function can be written as follows:

$$L(\beta, \sigma^2, \theta \mid y) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\sigma^2 R| - \frac{1}{2\sigma^2}(y - F\beta)^T R^{-1}(y - F\beta). \tag{3}$$

Then, the parameters β, σ^2, and θ could be estimated via maximum likelihood estimation (MLE). For given correlation parameters θ, the MLE of β is as follows:

$$\widehat{\beta} = \left(F^T R^{-1} F\right)^{-1} F^T R^{-1} y, \tag{4}$$

and the MLE of σ^2 is as follows:

$$\widehat{\sigma}^2 = \frac{1}{n}(y - F\widehat{\beta})^T R^{-1}(y - F\widehat{\beta}). \tag{5}$$

Since there is no analytical solution for the parameters θ, optimization methods are needed to obtain as follows:

$$\widehat{\theta} = \text{argmin}_\theta \left(n \log \widehat{\sigma}^2 + \log|R|\right). \tag{6}$$

Finally, the estimated parameters $\widehat{\beta}, \widehat{\sigma}^2, \widehat{\theta}$ are substituted into equation (2) to obtain the predictor of responses for the unsampled points.

2.2. Step-by-Step Variable Selection Methods. Although the kriging model has a character of interpolator due to its stochastic process, penalty methods still could be used to select the important variables from the candidates to improve prediction accuracy [13].

Two penalty functions are used for variable selection in the PBK model by Hung [11], i.e., Lasso [14] and adaptive Lasso [15]. It estimates the regression coefficients β by minimizing the negative penalized log-likelihood function.

$$Q\left(\beta, \theta, \sigma^2\right) = \frac{(y - F\beta)^T R^{-1}(y - F\beta)}{2\sigma^2} + \sum_{j=1}^{p} P_\lambda\left(\left|\beta_j\right|\right), \tag{7}$$

where $P_\lambda(|\beta_j|)$ is the regularization term called "penalty function," and λ is the regularization parameter. In order to efficiently obtain the estimation of parameters in the penalty likelihood function, Hung [11] proposed the iteratively reweighted least angle regression (IRLARS) algorithm for Lasso penalty and adaptive Lasso penalty.

Besides above penalty methods, Elastic Net penalty proposed by Zou and Hastie [16] also could be used for variable selection. It has regularization parameters λ and α. The Elastic Net estimation of the parameters β can be solved by the following formulation:

$$\widehat{\beta}_{EN} = \text{argmin}_\beta \left(\frac{(y - F\beta)^T R^{-1}(y - F\beta)}{2\sigma^2} + \lambda\alpha\|\beta\|_1 + \lambda(1 - \alpha)\|\beta\|_2^2 \right), \tag{8}$$

where $\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|$ and $\|\beta\|_2^2 = \sum_{j=1}^{p} \beta_j^2$. Thus, it can be seen that the penalty function of Elastic Net is a convex combination of ridge regression and Lasso. Particularly, the Elastic Net penalty becomes ridge regression when $\alpha = 0$,

and the Elastic Net penalty becomes Lasso penalty when $\alpha = 1$. According to the IRLARS algorithm for Lasso penalty, we describe the IRLARS algorithm for Elastic Net penalty in Algorithm 1.

Regularization parameters are obtained based on 10-fold cross validation, and the estimation of parameters β, θ, σ^2 are obtained based on the IRLARS algorithm in this research.

Although PBK effectively reduces the prediction error, further improvements are needed for low time efficiency. Thus, we propose three step-by-step variable selection methods to improve the prediction accuracy and time efficiency. We describe them in detail as follows:

2.2.1. Method One (M1). We integrate penalty into the linear model $y(x) = f(x)^T\beta + \varepsilon$ firstly, where $\varepsilon \sim N(0, \sigma^2)$; then, we select active variables by solving the penalized likelihood function for β:

$$\widehat{\beta} = \text{argmin}_{\beta}\left(\|y - F\beta\|^2 + \sum_{j=1}^{p} P_{\lambda}\left(\left|\beta_j\right|\right)\right), \tag{9}$$

where λ is obtained based on cross validation. Then, selected basis functions $f_i'(x), i = 1, \ldots, m$ corresponding nonzero regression coefficients β are substituted into the Kriging model (1) as new variables, and correlation parameters θ are updated as θ'. $\beta', \sigma'^2, \theta'$ are estimated by formulas (4)–(6), respectively. Finally, the estimated parameters $\widehat{\beta}', \widehat{\sigma}'^2, \widehat{\theta}'$ are used to predict the responses of unsampled points by substituting into formula (2), and the predicted model can be expressed as follows:

$$\widehat{y}'(x) = f'(x)^T\widehat{\beta}' + \widehat{r}'(x)^T\widehat{R}'^{-1}\left(y - F'\widehat{\beta}'\right). \tag{10}$$

The proposed method M1 is given in Algorithm 2.

2.2.2. Method Two (M2). Firstly, we estimate the parameters $\widehat{\theta}_{OK}$ and $\widehat{\sigma}^2_{OK}$ with the ordinary kriging model $y(x) = \mu + z(x)$ by solving the following formulas:

$$\widehat{\mu}_{OK} = \left(1_n^T R^{-1} 1_n\right)^{-1} 1_n^T R^{-1} y, \tag{11}$$

$$\widehat{\sigma}^2_{OK} = \frac{1}{n}\left(y - 1_n\widehat{\mu}_{OK}\right)^T R^{-1}\left(y - 1_n\widehat{\mu}_{OK}\right), \tag{12}$$

$$\widehat{\theta}_{OK} = \text{argmin}_{\theta}\left(n \log \widehat{\sigma}^2_{OK} + \log|R|\right), \tag{13}$$

where 1_n is an $n \times 1$ vector of ones. Then, we can select active variables by minimizing the negative penalized log-likelihood function $Q(\beta, \widehat{\theta}_{OK}, \widehat{\sigma}^2_{OK})$ and update correlation parameters. Finally, the new kriging model is fitted similarly to that in M1.The proposed method M2 is given in Algorithm 3.

2.2.3. Method Three (M3). Firstly, the PBK method is used to select active variables β based on the kriging model via minimizing the negative penalized log-likelihood function (7). It considers correlation between sample points, which is different from M1. Then, the selected variables based on PBK were substituted into the model (1) as new variables, and correlation parameters are updated. Finally, the estimated parameters $\widehat{\beta}', \widehat{\sigma}'^2, \widehat{\theta}'$ estimated by formulas (4)–(6) are used to predict the responses of unsampled points by substituting into (2).

The variable selection part of M3 is using IRLARS of PBK in Algorithm 1. The fitting steps are the same as Steps 2–4 in Algorithm 2 after selecting variables. We will not go into details here.

Theoretically, M1 and M2 do not need the iterative step of parameter estimation of β such as PBK in variable selection; thus, they can improve time efficiency greatly. In addition, the last step of the proposed three methods is to update θ and refit the kriging model, so the prediction accuracy could be improved.

3. Simulation

The effectiveness of the proposed methods is proved in terms of simulation performance. We validate the simulation performance of the methods compared with back propagation (BP) neural network, generalized regression (GR) neural network, and PBK method according to time efficiency and prediction accuracy. We use the following indicators to measure time and prediction accuracy.

(a) The root mean square prediction error (RMSPE) is defined as

$$RMSPE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y\left(x_i^{\text{test}}\right) - \widehat{y}\left(x_i^{\text{test}}\right)\right)^2}, \tag{18}$$

where N is the number of testing samples. The mean of RMSPE (MRMSPE) of the simulations is calculated to evaluate the prediction performance.

(b) The time of consumption (CPU) shows the uptime of the system.

Especially, we need identification to measure their accuracy of variable selection in linear functions using the following indicators:

(c) The average number of inactive effect identified rate (IEIR).

(d) The average number of active effect identified rate (AEIR).

(e) The average size of identified mean function (MEAN).

Obviously, when AEIR is larger, MRMSPE, IEIR, and CPU are smaller, and if MEAN is closer to the target, the model is better.

Four analytical functions are displayed in the following equations:

Step 1: set the initial values for $\beta^{(0)}, \theta^{(0)}, \sigma^{2(0)}$
Step 2: with $R(\widehat{\theta}^{(k)})^{-1}/\widehat{\sigma}^{(2(k))} = C^T C$, $y' = Cy$, and $F' = CF$, solve penalized likelihood function for λ and α obtained by cross validation,$\widehat{\beta}^{(k+1)} = \text{argmin}_\beta (\|y' - F'\beta\|^2 + \lambda\alpha \sum_{j=1}^{p} |\beta|_j + \lambda(1-\alpha) \sum_{j=1}^{p} |\beta|_j^2)$
Step 3: estimate σ^2, θ via maximizing the log-likelihood function (3),$\widehat{\sigma}^{2(k+1)} = 1/n(y - F\widehat{\beta}^{(k+1)})^T R^{-1}(y - F\widehat{\beta}^{(k+1)})$, $\widehat{\theta}^{(k+1)} = \text{argmin}_\theta (n \log \widehat{\sigma}^{2(k+1)} + \log|R|)$
Step 4: if the convergence is attained, get the estimation $\widehat{\beta}, \widehat{\theta}, \widehat{\sigma}^2$; otherwise, return to step 2

Algorithm 1: (IRLARS algorithm for Elastic Net penalty).

(1) A known function [11]:

$$y(x) = 0.4x_1 + 0.3x_2 + 0.2x_3 + 0.1x_4 + 0.05x_5 + 0.01x_6 + z(x). \quad (14)$$

(2) Test function [17]:

$$y(x) = 9 + 15\sum_{i=1}^{6} x_i + \frac{25}{2}\sum_{i=1}^{5} x_i x_{i+1} + 10\sum_{i=1}^{6} x_i^2. \quad (15)$$

(3) Gramacy and Lee function [13]:

$$y(x) = \exp\left[\sin\left(0.9(x_1 + 0.48)^{10}\right)\right] + x_2 x_3 + x_4. \quad (16)$$

(4) Borehole function [18]:

$$y(x) = 2\pi x_3 (x_4 - x_6) \times \log\left(\frac{x_2}{x_1}\right)\left(1 + 2\frac{x_3 x_4}{\log(x_2/x_1)x_1^2 x_8}\right)^{-1}. \quad (17)$$

A known function (KF) includes twelve-dimensional inputs, where the first six variables $(x_1, x_2, \ldots, x_6)$ decrease the effects on the responses, the remaining variables $(x_7, x_8, \ldots, x_{12})$, are unrelated (i.e., zero coefficients) to the outputs, and $z(x)$ has mean 0 and covariance $\sigma^2 R(\theta)$, where $\sigma^2 = 0.05$ and $\theta = 1_6$. The Test function (TF) includes twelve-dimensional inputs, where the last six variables are irrelevant to the outputs. The Gramacy and Lee function (GL) includes six-dimensional inputs, where the last two variables are irrelevant to the outputs. The Borehole function (BF) includes eight-dimensional inputs. And the nonlinearity gradually gets more obvious from the function KF to function BF. In function KF, we consider 60, 80, and 100 sample points for training in order to compare the effect of different sample sizes. In function TF, GL, and BF, we consider 100 sample points for training in order to compare the effect of nonlinearity. 1,000 sample points are randomly selected as testing data. We evaluate the accuracy of prediction, identification, and time efficiency based on 500 simulations.

In order to compare the effectiveness of the proposed methods, simulation studies are conducted to evaluate the performance of the BP neural network and GR neural network methods as other surrogate models firstly. The number of neurons in the hidden layer of the BP neural network is the optimal number of neurons among the 15 candidate neurons. The transfer functions from the input layer to the hidden layer and the hidden layer to the output layer choose "logsig" and "purelin," respectively. The simulation results for four functions are given in Table 1.

Due to the lack of interpretability of a neural network, neural network methods cannot select variables. In this paper, we will focus on variable selection for the kriging model. We carry out simulation studies using PBK, three methods in Section 2 with Lasso penalty and Elastic Net penalty. The first-order polynomial basis function and the stationary Gaussian correlation function are used. And Latin Hypercube designs [19] are used to select sample points for training data in the design domain because they can be produced with minimal computational cost, and the design space is fulfilled well. The simulation results for four functions are given in Tables 2–6, respectively.

The principal results of these simulations are as follows:

(1) According to Tables 2–6, the prediction accuracy of the proposed three methods is better than that of the PBK, BP, and GR neural networks, where M3 is the better one. In addition, their high prediction accuracy is obvious with escalating nonlinearity. Particularly, we take Table 4 as an example; the RMSPE for the predictors with Lasso penalty and Elastic Net in M3 gets low to (69.30% = (0.3528 − 0.1083) /0.3528)and 58.53% smaller than the PBK predictors, respectively.

(2) According to simulation results of the function KF (i.e., Tables 2 and 3), the method M1 could identify more variables, which lead to its highest AEIR and IEIR. Although AEIR of M2 is not high, its IEIR is the lowest and MEAN is close to the target. As to M3, its identification is basically the same as PBK. It indicates that M2 and M3 have better identification in variable selection.

(3) In the aspect of time, the efficiency of M1 and M2 is obviously better than that of PBK, BP, and GR neural networks, where M1 is the most efficient. But the efficiency of M3 is little worse than that of PBK. Particularly, we take $n = 100$ in Tables 2 and 3 as an example the time of M1 with Lasso penalty and Elastic Net penalty gets low to 79.16% and 86.50% smaller than PBK, respectively. For other cases, the optimization of time efficiency of M1 is almost higher than 80%.

In summary, the BP neural network is mostly worse than the traditional PBK method in the accuracy of prediction, and it is not faster than the PBK method with Lasso penalty in the time efficiency. Although the GR neural network is faster than the PBK method in the time efficiency, it is mostly worse than the traditional PBK method in the accuracy of prediction. We

Step 1: Screen variables using formula (9).
Step 2: Select the new basis function $f_i'(x)$ as new variables. Update the correlation parameter to θ'.
Step 3: Estimate the parameters $\beta', \sigma'^2, \theta'$ via formulas (4)–(6).
Step 4: The estimated parameters $\widehat{\beta}', \widehat{\sigma}'^2, \widehat{\theta}'$ are used to predict the responses by substituting into formula (10).

Algorithm 2: (variable selection algorithm for M1).

Step 1: give $\mu^{(0)}$, $\theta^{(0)}$, and $\sigma^{2(0)}$ as initial variables
Step 2: use formulas (12) and (13) for optimal model parameters $\widehat{\theta}_{OK}, \widehat{\sigma}^2_{OK}$
Step 3: with $R(\widehat{\theta}_{OK})^{-1}/\widehat{\sigma}^2_{OK} = C^T C$, $y^* = Cy$, and $F^* = CF$ solve the penalized likelihood function for λ, $\widehat{\beta} = \operatorname{argmin}_\beta (\|y^* - F^*\beta\|^2 + \sum_{j=1}^{p} P_\lambda(|\beta_j|))$, where λ is obtained based on cross validation
Step 4: the remaining steps of fitting are the same as Steps 2–4 in Algorithm 2 and will not be repeated here

Algorithm 3: (variable selection algorithm for M2).

Table 1: Data simulation results of neural network.

Function	Sample Size	BP		GR	
		MRMSPE	CPU(s)	MRMSPE	CPU(s)
KF	$n = 60$	0.110 6	3008.31	0.114 5	2142.63
	$n = 80$	0.108 5	3509.58	0.110 5	2236.98
	$n = 100$	0.107 7	3794.20	0.108 9	2246.46
TF	$n = 100$	3.972 7	5112.79	14.469 4	2429.14
GL	$n = 100$	0.668 5	4354.56	0.672 7	2306.72
BF	$n = 100$	4.341 1	3597.09	22.362 2	2405.54

Table 2: Data simulation results of the function KF with Lasso.

Sample Size	Method	MRMSPE	IEIR(%)	AEIR(%)	MEAN	CPU(s)
$n = 60$	PBK	0.067 8	37.60	92.47	7.80	2815.65
	M1	0.067 5	55.93	96.53	9.15	589.02
	M2	0.067 4	4.37	79.20	5.01	687.74
	M3	0.066 9	37.67	92.43	7.81	2879.55
$n = 80$	PBK	0.066 7	20.53	91.70	6.73	3101.15
	M1	0.066 5	58.17	97.00	9.31	618.83
	M2	0.066 3	5.67	81.20	5.21	953.33
	M3	0.065 7	20.77	91.60	6.74	3384.24
$n = 100$	PBK	0.065 8	12.53	93.17	6.34	3525.52
	M1	0.065 7	58.90	97.50	9.38	734.71
	M2	0.065 0	7.77	82.40	5.41	1128.58
	M3	0.064 9	14.20	93.17	6.44	3881.17

Table 3: Data simulation results of the function KF with Elastic Net.

Sample Size	Method	MRMSPE	IEIR(%)	AEIR(%)	MEAN	CPU(s)
$n = 60$	PBK	0.067 8	44.20	93.40	8.26	10 315.95
	M1	0.067 5	63.00	97.00	9.60	1126.76
	M2	0.067 0	6.33	81.57	5.27	1315.17
	M3	0.067 2	43.40	93.63	8.22	10 980.71
$n = 80$	PBK	0.066 7	25.17	92.27	7.05	10 787.02
	M1	0.066 5	63.10	97.27	9.62	1277.98
	M2	0.066 1	8.33	83.27	5.50	1685.08
	M3	0.065 8	25.37	92.53	7.07	11 161.00
$n = 100$	PBK	0.065 7	17.37	93.63	6.66	11 135.97
	M1	0.065 6	63.57	97.53	9.67	1503.19
	M2	0.065 1	10.73	83.40	5.65	1914.47
	M3	0.065 0	17.90	93.90	6.71	11 761.19

Table 4: Data simulation results of the function TF.

Method	Lasso		Elastic Net	
	MRMSPE	CPU(s)	MRMSPE	CPU(s)
PBK	0.352 8	3896.31	0.355 9	13 416.90
M1	0.205 1	578.45	0.268 5	1297.63
M2	0.109 1	888.38	0.262 9	2286.61
M3	0.108 3	3924.61	0.147 6	14 977.26

Table 5: Data simulation results of the function GL.

Method	Lasso		Elastic Net	
	MRMSPE	CPU(s)	MRMSPE	CPU(s)
PBK	0.755 5	2015.58	0.755 8	7034.73
M1	0.641 6	242.03	0.640 2	766.81
M2	0.633 8	600.29	0.632 6	1237.27
M3	0.634 2	2181.55	0.634 6	8150.91

Table 6: Data simulation results of the function BF.

Method	Lasso		Elastic Net	
	MRMSPE	CPU(s)	MRMSPE	CPU(s)
PBK	0.877 9	2198.33	0.878 7	10 522.03
M1	0.710 4	367.71	0.871 3	1234.14
M2	0.687 0	696.51	0.827 8	1459.26
M3	0.616 4	2775.03	0.796 4	12 315.72

take $n = 100$ in function KF as an example; the RMSPE for BP neural network, the GR neural network and PBK is 0.107 7, 0.108 9, and 0.065 8, respectively, and their time is 3794.20 s, 2246.46 s, and 3525.52 s. In addition, due to the lack of interpretability of BP and GR neural networks, these methods do not have identification. Thus, the neural network is not the preferred candidate method in variable selection methods compared with PBK. Although M1 greatly improves time efficiency by reducing the iterative steps of parameter estimation of β, it identified more inactive variables than PBK, which lead to its high AEIR and IEIR, as the correlation between sample points is not considered in the variable selection. As to M2, it not only maintains the advantage of M1 in time efficiency but also makes up for the shortcomings of M1 by considering the correlation between sample points. Therefore, M2 is superior to M1 in both identification and prediction accuracy and can be used to replace PBK. Although M3 is refitted on the basis of the variable selection of PBK, which leads to its low time efficiency, it has a significant improvement in the prediction accuracy. Therefore, we can use M3 as an improved method of PBK.

4. Empirical Application

4.1. Circuit Simulation. The proposed methods are applied to the circuit-simulation code [12]. Six input variables $x_1, \ldots, x_6$ and one response y are in this experiment. The simulation experiments were conducted 32 times. This data are also used to analyze the relationship between output and input [20]. It can be useful for illustrating the prediction accuracy and time efficiency. The candidate variable set includes all linear effects. To compare the prediction accuracy and time efficiency of different methods, the dataset was separated into two sections: half of the data are used for training and the other half for testing. The experimental results with Lasso and Elastic Net penalty are listed in Tables 7 and 8, respectively.

From the second and third columns of Tables 7 and 8, we could find the prediction accuracy of M1 (0.220 7 and 0.2207), M2 (0.2207 and 0.2207), and M3 (0.1520 and 0.1591) which can achieve better prediction than PBK (0.286 9 and 0.274 0), and the prediction of M1 and M2 are all improved by 23.07% and 19.45%, those of M3 are improved by 47.02% and 41.93% relative to PBK.

As displayed in the last two columns of Tables 7 and 8, we could find the operation time of M1 (0.37 s and 0.81 s) and M2 (0.43 s and 1.08 s) which are less than those of PBK (2.26 s and 8.65 s) and M3 (2.29 s and 9.01 s), and the approximation efficiency of M1 is improved by 83.63% and 90.64%, those of M2 are improved by 80.97% and 87.51% relative to PBK.

Therefore, M1 and M2 can be used as variable selection methods to replacing the traditional PBK method according to prediction accuracy and time efficiency for fitting small samples. And M3 can be used as an improved method of PBK due to better prediction accuracy.

4.2. Piston Clap Noise. Piston secondary motion would cause an engine noise which is unwanted. It is the departure of a piston from the nominal motion prescribed by the mechanism. Six variables were used to minimize the piston clap noise. The variables are the cylinder liner x_1, location of peak pressure x_2, skirt length x_3, skirt profile x_4, skirt ovality x_5, and pin offset x_6, respectively.

Table 7: Experimental results for circuit simulation with Lasso.

Method	RMSPE	Improved (%)	CPU(s)	Improved (%)
PBK	0.286 9	—	2.26	—
M1	0.220 7	23.07	0.37	83.63
M2	0.220 7	23.07	0.43	80.97
M3	0.152 0	47.02	2.29	-1.33

Table 8: Experimental results for circuit simulation with Elastic Net.

Method	RMSPE	Improved (%)	CPU(s)	Improved (%)
PBK	0.274 0	—	8.65	—
M1	0.220 7	19.45	0.81	90.64
M2	0.220 7	19.45	1.08	87.51
M3	0.159 1	41.93	9.01	-4.16

Table 9: Experimental results for piston clap noise with Lasso.

Method	RMSPE	Improved (%)	CPU(s)	Improved (%)
PBK	1.0069	—	18.50	—
M1	0.8316	17.41	2.11	88.59
M2	0.5466	45.71	5.20	71.89
M3	0.6357	36.87	21.66	-17.08

Table 10: Experimental results for piston clap noise with Elastic Net.

Method	RMSPE	Improved (%)	CPU(s)	Improved (%)
PBK	0.9942	—	189.45	—
M1	0.7108	28.51	5.75	96.96
M2	0.5466	45.02	16.89	91.08
M3	0.5466	45.02	207.76	-9.66

The relevant dataset of the example is derived from Fang et al. [2]. The data include 100 observations, 6 input variables, and the candidate variable set includes all linear effects. In this paper, we divide the 100 data into 5 sets, select one of the sets as the test set for calculating RMSPE, and others as the train sets; this process is repeated 5 times and train sets are different every time. PBK, M1, M2, and M3are compared, and the running time is recorded. The simulation results with Lasso and Elastic Net penalty are shown in Tables 9 and 10, respectively.

From the second and third columns of Tables 9 and 10, we could find the prediction accuracy of M2 (0.5466 and 0.5466) which can achieve better prediction than PBK (1.0069 and 0.9942), M1 (0.8316 and 0.7108), and M3 (0.6357 and 0.5466), and the prediction of M1 is improved by 17.41% and 28.51%, those of M2 are improved by 45.71% and 45.02%, and those of M3 are improved by 36.87% and 45.02% relative to PBK.

As displayed in the last two columns of Tables 9 and 10, we can find that the operation time of the M1 (2.11 s and 5.75 s) and M2 (5.20 s and 16.89 s) which are less than those of PBK (18.50 s and 189.45 s) and M3 (21.66 s and 207.76 s), and the approximation efficiency of M1 is improved by 88.59% and 96.96%, and those of M2 are improved by 71.89% and 91.08% relative to PBK.

Therefore, M1 and M2 can achieve the goal which replace the PBK method both in terms of prediction accuracy and time efficiency. And M3 can be viewed as an improvement of PBK.

5. Conclusion and Discussion

In the field of engineering, we always use a simple surrogate model to fit the complex true model for time saving and interpretability [21]. In order to improve the accuracy of prediction and save the cost of time for the PBK method, three new methods are proposed to obtain the trend function to improve the prediction accuracy and time efficiency. M1 is to select the variables using the linear model at the first and then to estimate the parameters in the kriging model. M2 is first to estimate the parameters with the OK model and then to select the variables via the penalized kriging model and finally refit kriging model. M3 is an improvement of PBK, which is a refitting based on the variable selection of PBK.

Several analysis functions with escalating nonlinearity and two engineering examples with different sample sizes are used to test our proposed step-by-step variable selection methods. The experimental results show that M2 which is superior to PBK, BP, and GR neural networks and other

methods can be used as a favorable method of variable selection in both prediction accuracy and time efficiency. If we want to pursue higher time efficiency while maintaining certain prediction accuracy, M1 might be a good choice. As well, M3 could be used as an improved method of PBK due to good prediction accuracy.

Data Availability

The data used to support the findings of this study are included in this paper.

Disclosure

Ziheng Feng and Min Li are co-first authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 11871294 and 12101346) and Shandong Provincial Natural Foundation, China (Nos. ZR2021QA053 and ZR2021QA044).

References

[1] T. J. Santner, B. J. Williams, and W. I. Notz, *The Design and Analysis of Computer Experiments*, Springer, New York, 2003.

[2] K. T. Fang, R. Li, and A. Sudjianto, *Design and Modeling for Computer Experiments*, CRC Press, Florida, 2005.

[3] H. Dong, P. Wang, C. Fu, and B. Song, "Kriging-assisted teaching-learning-based optimization (KTLBO) to solve computationally expensive constrained problems," *Information Sciences*, vol. 556, pp. 404–435, 2021.

[4] H. Dong, P. Wang, B. Song, Y. Zhang, and X. An, "Kriging-assisted Discrete Global Optimization (KDGO) for black-box problems with costly objective and constraints," *Applied Soft Computing*, vol. 94, p. 106429, 2020.

[5] B. Echard, N. Gayton, and M. Lemaire, "Ak-mcs: an active learning reliability method combining Kriging and Monte Carlo simulation," *Structural Safety*, vol. 33, no. 2, pp. 145–154, 2011.

[6] S. Zhang, P. Zhu, W. Chen, and P. Arendt, "Concurrent treatment of parametric uncertainty and metamodeling uncertainty in robust design," *Structural and Multidisciplinary Optimization*, vol. 47, no. 1, pp. 63–76, 2013.

[7] P. Kersaudy, B. Sudret, N. Varsier, O. Picon, and J. Wiart, "A new surrogate modeling technique combining Kriging and polynomial chaos expansions - application to uncertainty analysis in computational dosimetry," *Journal of Computational Physics*, vol. 286, pp. 103–117, 2015.

[8] H. Liu, Y.-S. Ong, and J. Cai, "A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design," *Structural and Multidisciplinary Optimization*, vol. 57, no. 1, pp. 393–416, 2018.

[9] J. D. Martin and T. W. Simpson, "Use of kriging models to approximate deterministic computer models," *AIAA Journal*, vol. 43, no. 4, pp. 853–863, 2005.

[10] V. R. Joseph, Y. Hung, and A. Sudjianto, "Blind kriging: a new method for developing metamodels," *Journal of Mechanical Design*, vol. 130, no. 3, pp. 350–353, 2008.

[11] Y. Hung, "Penalized blind kriging in computer experiments," *Statistica Sinica*, vol. 21, no. 3, pp. 1171–1190, 2011.

[12] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp. 409–423, 1989.

[13] Y. Zhang, W. Yao, S. Ye, and X. Chen, "A regularization method for constructing trend function in Kriging model," *Structural and Multidisciplinary Optimization*, vol. 59, no. 4, pp. 1221–1239, 2019.

[14] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[15] H. Zou, "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.

[16] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B*, vol. 67, no. 2, pp. 301–320, 2005.

[17] Y. Zhang, W. Yao, X. Chen, and S. Ye, "A penalized blind likelihood Kriging method for surrogate modeling," *Structural and Multidisciplinary Optimization*, vol. 61, no. 2, pp. 457–474, 2020.

[18] M. D. Morris, T. J. Mitchell, and D. Ylvisaker, "Bayesian design and analysis of computer experiments: use of derivatives in surface prediction," *Technometrics*, vol. 35, no. 3, pp. 243–255, 1993.

[19] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.

[20] W. J. Welch, R. J. Buck, J. Sacks, H. P. Wynn, T. J. Mitchell, and M. D. Morris, "Screening, predicting, and computer experiments," *Technometrics*, vol. 34, no. 1, pp. 15–25, 1992.

[21] M. Plumlee and V. R. Joseph, "Orthogonal Gaussian process models," *Statistica Sinica*, vol. 28, no. 2, pp. 601–619, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 4553872, 6 pages
https://doi.org/10.1155/2021/4553872

Research Article

A Novel Method for Integration of Online Educational Resources via Teaching Information Remote Scheduling

Kang Zhao

College of Literature and Media, Xi'an Fanyi University, Xi'an 710105, China

Correspondence should be addressed to Kang Zhao; zhaokang@xafy.edu.cn

Received 11 June 2021; Revised 21 July 2021; Accepted 25 July 2021; Published 4 October 2021

Academic Editor: Hussein Abulkasim

The traditional online education resource integration technology ignores the calculation of the characteristic variance contribution rate of online educational resources, and the evaluation of educational resources is biased, which leads to the low ability of resource integration. Therefore, the online education resource integration technology based on remote scheduling of teaching information is proposed. Based on the theory of remote scheduling of teaching information, the quantitative value of online educational resources integration is obtained. By calculating the integration probability of online educational resources, the influencing factors of online educational resources in the process of integration are determined. The principle of teaching information remote scheduling is integrated into the feature vector extraction of online educational resources, and the contribution rate of feature variance of online educational resources is calculated. By extracting the number of main factors of the online educational resources feature vector, the feature vector of online educational resources is decomposed. Taking the wavelet entropy of online educational resources as the fusion weight of online educational resources, the integration result of online educational resources is obtained through wavelet transform. The experimental results show that the online education resource integration technology based on remote scheduling of teaching information has ideal application performance in integration probability, the goodness of fit, and robustness.

1. Introduction

The development and progress of education informatization have contributed to the accumulation of more and more information resources in the education information system [1, 2]. However, due to factors such as the level of technology and management system, islands of information exist in online education information resources. To solve this problem, online educational resources need to be developed and integrated in a rational way, so that these resources serve for online education [3].

In order to improve the capability in retrieving educational resources, the literature [4] analyzed the core structure model of multimedia resource distribution in embedded distance education networks, sorted the nonlinear time of educational resource information flow, divided the time window of multimedia resources into a set of time slots uniformly distributed using the smoothing filter method, and then explored the accessibility of multimedia resources in distance education, through which the balance between resource allocation and resource utilization were effectively improved. The literature [5] points out that, in the traditional method, there was no statistics on the number of specific departments in computing resource integration, which led to the low accuracy of integration; it proposes the method of integrating educational resources for college students based on the factor analysis theory, so as to obtain many empirical values, quantify these values, and get the main influencing elements of resource integration, while based on the factor analysis theory and the influence on the integration of educational resources of college students, the coupling matrix is obtained for the integration of the main educational resources of college students, and the number of basic departments is calculated through the integration of these resources, so as to analyze and study the current situation of the integration of educational resources of

college students, and it has been proved by experiment that this method has good robustness and can better integrate the educational resources for students.

Based on the above research background, this paper applies the remote scheduling of teaching information to the integration of online educational resources, which improves the ability of online educational resources integration.

2. Online Education Resource Integration Technology

2.1. Influencing Factors of Online Educational Resources in the Process of Integration. In the process of integration of online educational resources, based on the theory of remote scheduling of teaching information [6], different quantitative values existed in the integration of online educational resources were obtained to extract the factors influencing the integration of online educational resources as follows.

Suppose that s denotes the factor of multifaceted conflicts that influence the integration of online educational resources, x denotes the abnormal measurement value in the process of integration of online educational resources, n denotes the time frame in the process of integration of online educational resources, and j denotes the quantitative value of the integration performance of online educational resources. Combined with the theory of remote scheduling of teaching information [7], the quantitative value of the integration of online educational resources generated in the process of integration of online educational resources due to the expectation discrepancy that educational resources have is obtained and calculated as follows:

$$O_h = \sum_{j=1}^{n} \frac{s}{x_j} + \sum_{j=1}^{n} \frac{c_{kh} \times \nu_s}{x_j}, \tag{1}$$

where c_{kh} denotes the regular performance of online educational resources in the integration process and ν_s denotes the adaptive capacity shown by online educational resources in the integration process. Then the integration probability of online educational resources can be calculated using the following equation:

$$\vartheta\left(V_y\right) = \frac{\mu_b \times \kappa_r}{x_j} c_{kh} O_h, \tag{2}$$

where κ_r denotes the current status of integration of online educational resources and μ_b denotes the characteristics of changes in the process of integration of online educational resources.

Suppose that h_r denotes the characteristics of online educational resources to determine the influencing factors of online educational resources in the process of integration as follows:

$$K_v = \frac{h_r \times z_o}{\tau_m} \times \frac{F^D}{\iota(i) \cdot a_d}, \tag{3}$$

where $\iota(i)$ denotes the attributes of online educational resources, i denotes the attribute vector of factors influencing the integration of online educational resources, F^D denotes the features selected, z_o denotes the time required in the process of integration of online educational resources, τ_m denotes the probability of failure in the integration of online educational resources, and a_d denotes the integration cycle of online educational resources.

Based on the theory of remote scheduling of teaching information, the quantitative value of online educational resources integration is obtained. By calculating the integration probability of online educational resources, the influencing factors of online educational resources in the process of integration are determined.

2.2. Extraction of Online Educational Resources Feature Vector. In the process of integration of online educational resources, the principle of teaching information remote scheduling is integrated into the feature vector extraction of online educational resources [8], and the contribution rate of feature variance of online educational resources is calculated. The observable random vector of online educational resource features is given. By extracting the number of main factors of online educational resources feature vector as follows, the feature vector of online educational resources is decomposed [9].

Suppose that n denotes the number of feature variables of the original online educational resources, X denotes the condition of $X = x_1, x_2, \ldots, x_n$ that the n feature variables in the original online educational resources samples need to satisfy to carry out orthogonal transformation of online educational resources [10] and to integrate u' online educational resources feature variables (y_1, y_2, y_3), and R denotes the correlation coefficient matrix of online educational resources samples. The theory of remote scheduling of teaching information [11] is used to establish the feature equation of online educational resources as follows:

$$\lambda^n(i) = \frac{\{R \otimes X\}^n}{\{(y_1, y_2, y_3)\}} \otimes \frac{(x_1, x_2, \ldots, x_n)}{u'}. \tag{4}$$

Suppose that λ_i denotes the number of nonnegative feature values of the correlation coefficient matrix of online educational resources samples. Under the condition that satisfies $\lambda_1 \geq \lambda_2 \geq \lambda_n \geq 0$, λ_i sorting is conducted, and then the first m online educational resource features can be extracted as follows:

$$\Phi(p) = \frac{m \otimes (\lambda_i)}{\xi(e)} \otimes \eta(r), \tag{5}$$

where $\xi(e)$ denotes the noise interference of online educational resources feature and $\eta(r)$ denotes the uncertainty of online educational resources feature vector.

Suppose that α denotes the variance of the first m online educational resources features. The calculating formula is as follows:

$$\beta(p) = \frac{m \otimes \alpha}{\mu(R)} \times \nu(e)\left(\sigma^* \otimes \kappa\right), \tag{6}$$

where $\mu(R)$ denotes the weights of online educational resources feature samples, $\nu(e)$ denotes the information

entropy of different online educational resource features, σ^* denotes the optimal threshold value of online educational resource feature variables, and κ denotes the observed variables of student features.

$X = x_1, x_2, \ldots, x_n$ is defined as the random vector of online educational resource features, a_{ij} denotes the factor loading of online educational resources feature vector, and the observable random vector of online educational resource features is calculated as follows:

$$\partial(X) = \frac{X \otimes F}{\left(a_{ij}\right)_{n \mp m}} \otimes c_i \otimes \varepsilon_i \times X_i, \tag{7}$$

where F denotes the unobservable vector of online educational resources, c_i denotes the factor loadings of particular online educational resources, and ε_i denotes the unique factors affecting factor loadings c_i.

The degree of uncertainty of the online educational resources information source X^* is measured by the equation as follows:

$$H\left(X^*\right) = \frac{p_i \log\, p_i}{H\left(p_1, p_2, \ldots, p_n\right)} I\left(a_i\right), \tag{8}$$

where $I(a_i)$ denotes the space of online educational resources information source X^*, p_i denotes the probability of occurrence of online educational resources integration strategy, and $H(p_1, p_2, \ldots, p_n)$ denotes the probability of occurrence of discrete random variables.

Suppose that $I(\lambda_i)$ denotes the information function of online educational resources integration. The probability of online educational resources integration behavior can be defined as

$$\xi(w) = \frac{I\left(\lambda_i\right) \otimes H\left(X^*\right)}{\zeta(k)} \otimes \vartheta(R), \tag{9}$$

where $\zeta(k)$ denotes the information volume of online educational resources and $\vartheta(R)$ denotes the cumulative information contribution rate of online educational resources integration.

According to the abovementioned, the principle of remote scheduling of teaching information is integrated into the feature vector extraction of online educational resources, the contribution rate of feature variance of online educational resources is calculated, and by extracting the number of main factors of the feature vector of online educational resources, the feature vector of online educational resources is decomposed, which lays the foundation for the integration of online educational resources.

2.3. Integration of Online Educational Resources. According to the feature vector extraction results of online educational resources, a complete online educational resources dataset is established. Through data fusion algorithm [12], all the features of online educational resources in the online educational resources dataset are reconstructed [13], and then, the reconstructed online educational resources dataset is decomposed. The wavelet variance of online educational resources is calculated according to the decomposition result [14], which is used as the weight fusion wavelet coefficient to obtain the integration result of online educational resources.

Suppose that q explorers manage multiple online educational resources X simultaneously. The value of the online educational resources reserve in each explorer can be expressed as

$$Z_i = X + V_i, \quad i = 1, 2, \ldots, m, \tag{10}$$

where V_i denotes the parameters of online educational resources in the explorers.

Suppose that δ_i denotes the weight in each explorer, σ_i denotes the variance of parameters of online educational resources, the integration value of online educational resources finally obtained is $Z = \sum_{i=1}^{q} \delta_i Z_i$, and $\sum_{i=1}^{q} \delta_i = 1$, $0 < \delta_i < 1$. The total mean square error of online educational resources integration is as follows:

$$\sigma^2 = E\left[(X - Z)^2\right] = \sum_{i=1}^{q} \delta_i^2 \sigma_i^2. \tag{11}$$

According to the theory for the extremum of multivariable functions [15], the minimum value of equation (11) can be calculated as $1/\sum_{i=1}^{q} 1/\sigma_i^2$, and the weighting factor corresponding to the minimum value is

$$\delta_i = \frac{1/\sigma_i^2}{\sum_{i=1}^{q} 1/\sigma_i^2}. \tag{12}$$

The weighting factor obtained from equation (12) is used to reconstruct the features of online educational resources [16], and the equation is

$$Z_J = \sum_{k} h_n(J,k)\varphi_{J,k} + \sum_{J} \sum_{k} g_n(J,k)\psi_{J,k}, \tag{13}$$

where Z_J denotes the reconstructed online educational resources, $\varphi_{J,k}$ denotes the scaling function of online educational resources at layer J, $h_n(J,k)$ denotes the k^{th} approximate coefficient of online educational resources, $\psi_{J,k}$ denotes the wavelet resource function of online educational resources at layer J, and $g_n(J,k)$ denotes the k^{th} detail resource function of online educational resources.

Suppose that $W_{h,i}$ denotes the weighting factors of the approximate parameters of online educational resources and $W_{g,i}$ denotes the weighting factors of the detail parameters of online educational resources. Then it can be obtained as follows:

$$\sigma^2 = \sigma_h^2 + \sigma_g^2 = E\left[Z_V^2 - \sum_{i=1}^{q} W_{h,i}^2 \frac{1}{2}\eta_i \varphi_{J,k}\right]. \tag{14}$$

According to equation (14), it can be known that the mean variance of the weighted results of online educational resources integration is less than or equal to that of the optimal weighting of online educational resources in the time domain, and the remote scheduling of online teaching resources through teaching information [17] leads to an increasingly smaller mean variance of online educational resources with the decomposition of the scale, and the final

conclusion provides a theoretical basis for the integration of online educational resources [18].

Suppose that X denotes the parameter sequence of online educational resources managed. Based on the teaching information remote scheduling theory, it can be known that wavelet transform has the property of energy conservation under orthogonal wavelet basis, and based on the time sequence [19], the energy can be divided on the scale domain, and using the wavelet basis function, the energy on the explorers of online educational resources can be decomposed as

$$\|X\|^2 = \sum_{i=1}^{J} \|W_J\|^2 + \|V_J\|^2. \tag{15}$$

The measured variance of online educational resources is given by

$$\sigma_X^2 = \frac{1}{N}\left\| \sum_{N=1}^{N} (X - \overline{X})^2 \right\| = \left(\|W_{j,n}\|^2 + \|V_{j,n}\|^2 \right). \tag{16}$$

As V_j approximates $\overline{X}$, the wavelet variance and the average wavelet energy on scale j in equation (16) are normalized as follows:

$$p_j(E) = \frac{E_j}{E} = \frac{1/N\|W_j\|^2}{E}, \tag{17}$$

where E denotes the total energy of online educational resources, $E = \sum_{j=1}^{J} E_j$, and the energy sequence $p_j(E)$ of online educational resources normalized denotes the ratio of the energy of online educational resources to the total energy. The wavelet entropy of online educational resources is obtained based on the distribution of the energy sequence of online educational resources [20–22]. The specific procedure is as follows:

$$H_{we} = H(P) = H\left(p_1(E), p_2(E), \ldots, p_J(E)\right). \tag{18}$$

To sum up, taking the wavelet entropy of online educational resources as the fusion weight of online educational resources, the integration result of online educational resources is obtained through wavelet transform, realizing the integration of online educational resources.

3. Experimental Comparative Analysis

An experiment was conducted to verify the effectiveness of online educational resources integration technology based on remote scheduling of teaching information. Six thousand three hundred and twenty-five types of online educational resources in varied majors of a university were selected as objects of the experiment, from which 700 types were chosen for integration. The experiment results were subject to check analysis.

Tests were conducted with the online educational resources integration technology based on teaching information remote scheduling, and the integration technologies of online educational resources in the literature [4, 5]. The results obtained from equation (2) are used to compare the integration probability and the actual value of the three integration technologies, which are shown in Figure 1.

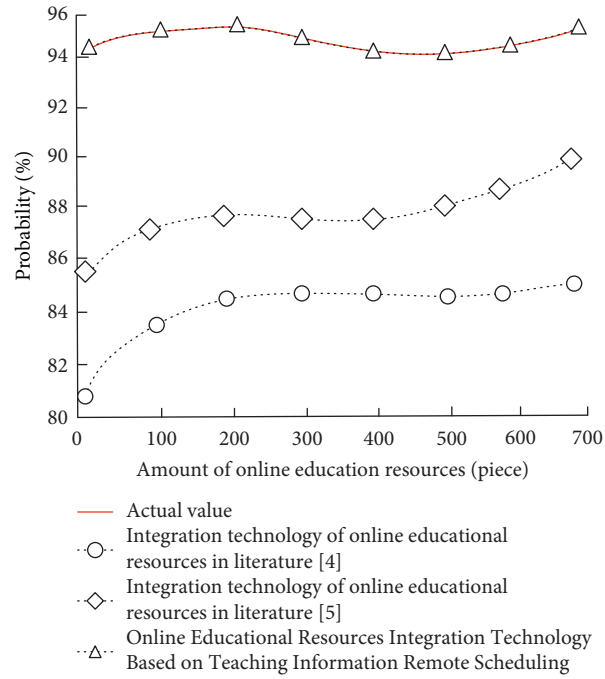


Figure 1: Comparison results of online educational resources integration probability.

It can be seen from the results in Figure 1 that the probability of having influencing factors on online educational resources integration technology based on teaching information remote scheduling coincides with the actual value. The reason is that the technology allows beforehand calculation of the probability of factors having influence on online educational resources integration to extract features of online educational resources, based on which the online educational resources integration probability matches the actual value.

To further prove the rationality of online educational resources integration technology based on teaching information remote scheduling, the goodness of fit of the technology is verified through relevant checking residuals by using the following equation:

$$R^{\text{new}} = 1 - \frac{Q}{2\sum \widehat{y}_2} \times 100\%, \tag{19}$$

where new denotes the coefficient of determination, $\widehat{y}_2$ denotes the fitting degree, and Q denotes the residual sum of squares.

The test results of goodness of fit by online educational resources integration analysis based on three different technologies are shown in Figure 2.

It is shown in Figure 2 that the online educational resources integration technology based on teaching information remote scheduling is superior to the other two technologies in terms of goodness of fit because the quantitative value of online education resources integration features is obtained based on the technology, presenting the

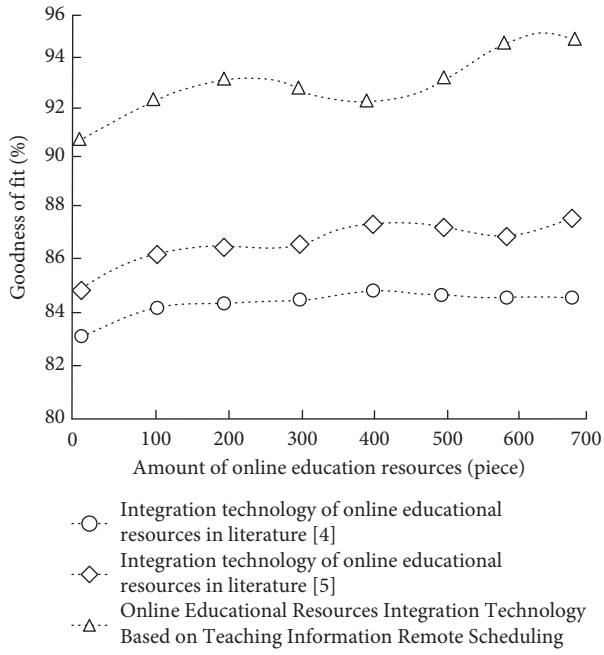


Figure 2: Test results of goodness of fit by online educational resources integration analysis.

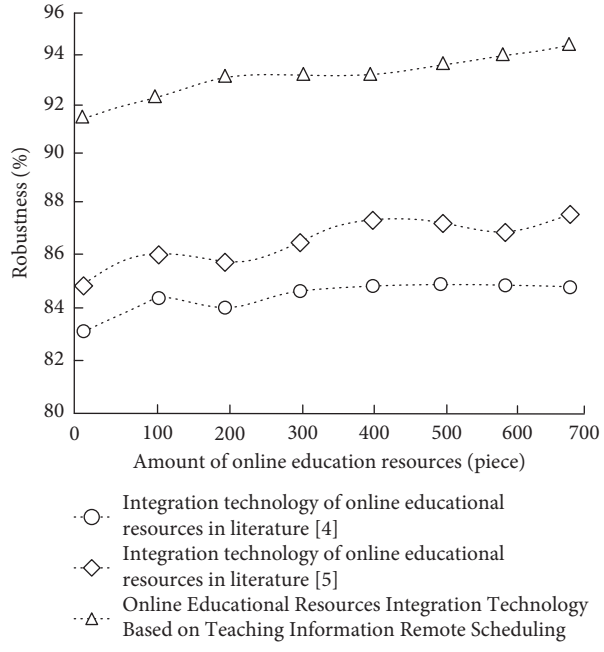


Figure 3: Robustness of online educational resources integration.

correlation between influencing factors on online educational resources integration and online educational resources integration. Thus, the technology shows a better goodness of fit.

Based on the results above, the robustness of the three online educational resources integration technologies was tested, and the results are as shown in Figure 3.

It can be seen from Figure 3 that the online educational resources integration technology based on teaching information remote scheduling shows a better robustness than the other two technologies. The reason is that this technology analyzes the relation between the influencing factors on online educational resources and online educational resources, determining the integration degree of influencing factors on online educational resources. Thus, such technology has a better robustness.

4. Conclusion

A research on online educational resources integration technology based on teaching information remote scheduling is put forth in this paper. By determining influencing factors on online educational resources integration, the feature vector of online educational resources integration is extracted, and online educational resources are integrated pursuant to the online educational resources integration process. The results show that this technology can effectively improve the integration capability of online educational resources.

Data Availability

All the data can be requested from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] J. Wolfe and G. Hatsidimitris, "Physclips: multimedia, multilevel learning, and teaching resources," *Journal of the Acoustical Society of America*, vol. 132, no. 3, p. 2011, 2012.

[2] T. Okemakinde, S. O. Adedeji, and J. Ssempebwa, "Utilization of teaching resources and academic performance in technical colleges in Oyo state, Nigeria," *Monthly Notices of the Royal Astronomical Society*, vol. 380, no. 2, pp. 741–748, 2012.

[3] C. Klupiec, S. Pope, R. Taylor, D. Carroll, M. Ward, and P. Celi, "Development and evaluation of online video teaching resources to enhance student knowledge of livestock handling," *Australian Veterinary Journal*, vol. 92, no. 7, pp. 235–239, 2014.

[4] M. Nong, "Research on multimedia resource scheduling technology in distance education network," *Modern Electronics Technique*, vol. 40, no. 24, pp. 68–70, 2017.

[5] X.-Y. Qiao, "Optimization of college students' educational resources integration," *Computer Simulation*, vol. 34, no. 8, pp. 239–242, 2017.

[6] T. Wilkinson, "Primal pictures anatomy teaching resources: 3D anatomy software and 3D real-time viewer," *Journal of Anatomy*, vol. 220, no. 1, pp. 118-119, 2011.

[7] M. L. Aikens and E. L. Dolan, "Teaching quantitative biology: goals, assessments, and resources," *Molecular Biology of the Cell*, vol. 25, no. 22, pp. 3478–3481, 2014.

[8] S. B. Issenberg, S. Pringle, R. M. Harden, S. Khogali, and M. S. Gordon, "Adoption and integration of simulation-based learning technologies into the curriculum of a UK undergraduate education programme," *Medical Education*, vol. 37, no. s1, pp. 42–49, 2003.

[9] C. C. Maponga and S. A. Feresu, "A community-based clinical pharmacy teaching, research, and service program in Zimbabwe," *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, vol. 15, no. 5, pp. 614–620, 2012.
[10] I. A. Furtado, L. L. Neto, and A. J. Gonçalves-Ferreira, "Paths for anatomy education in the digital era: options of the institute of anatomy of the faculty of medicine of lisbon," *Annals of Medicine*, vol. 51, no. s1, p. 54, 2019.
[11] S. Pullen and K. Brinkert, "SolEn for a sustainable future: developing and teaching a multidisciplinary course on solar energy to further sustainable education in chemistry," *Journal of Chemical Education*, vol. 91, no. 10, pp. 1569–1573, 2014.
[12] M. Gupta and P. Lata, "Development of one-dimensional KLM model based on ABCD matrices," *Diagnostic Ultrasound Imaging: Inside Out*, vol. 8, no. 13, pp. 787–790, 2014.
[13] M. Stead, R. Stradling, M. Macneil et al., "Bridging the gap between evidence and practice: a multi-perspective examination of real-world drug education," *Drugs: Education, Prevention & Policy*, vol. 17, no. 1, pp. 1–20, 2010.
[14] R. S. Wigton, "Use of linear models to analyze physicians' decisions," *Medical Decision Making*, vol. 8, no. 4, pp. 241–252, 1988.
[15] A/RES, "Towards full integration of persons with disabilities into society: a continuing world programme of action: resolution/adopted by the general assembly," *Anaesthesia*, vol. 65, no. 8, pp. 867–871, 2012.
[16] M. A. M. Arthur, R. Battat, and T. F. Brewer, "Teaching the basics: core competencies in global health," *Infectious Disease Clinics of North America*, vol. 25, no. 2, pp. 347–358, 2011.
[17] T. S. Klitzner, L. A. Rabbitt, and R.-K. R. Chang, "Benefits of care coordination for children with complex disease: a pilot medical home project in a resident teaching clinic," *The Journal of Pediatrics*, vol. 156, no. 6, pp. 1006–1010, 2010.
[18] S. Alsaggaf, S. S. Ali, N. N. Ayuob, B. S. Eldeek, and A. Elhaggagy, "A model of horizontal and vertical integration of teaching on the cadaveric heart," *Annals of Anatomy—Anatomischer Anzeiger*, vol. 192, no. 6, pp. 373–377, 2010.
[19] M. I. Kleib, "Integration of health informatics in baccalaureate nursing education: effectiveness of face-to-face vs. online teaching methods," *Mathematische Zeitschrift*, vol. 263, no. 4, pp. 1279–1296, 2012.
[20] I. L. Johnson, F. E. Scott, N. P. Byrne, K. A. MacRury, and J. Rosenfield, "Integration of community health teaching in the undergraduate medicine curriculum at the university of toronto," *American Journal of Preventive Medicine*, vol. 41, no. 4-supp-S3, pp. S176–S180, 2011.
[21] S. Mlambo, P. Rambe, and L. Schlebusch, "Effects of Gauteng province's educators' ICT self-efficacy on their pedagogical use of ICTS in classrooms," *Heliyon*, vol. 6, no. 4, Article ID e03730, 2020.
[22] A. Pakdaman-Savoji, J. C. Nesbit, and N. Gajdamaschko, "The conceptualisation of cognitive tools in learning and technology: a review," *Australasian Journal of Educational Technology*, vol. 35, no. 2, pp. 1–24, 2019.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 3462846, 8 pages
https://doi.org/10.1155/2021/3462846

Research Article

Generation of Digital Art Composition Using a Multilabel Learning Algorithm

Wei Li and Xin Gong

Department of Plastics Arts, Apparel Art Design College, Xi'an polytechnic university, Xi'an 710048, China

Correspondence should be addressed to Wei Li; weili_2021@21cn.com

Received 24 May 2021; Accepted 8 August 2021; Published 9 September 2021

Academic Editor: Hussein Abulkasim

The traditional methods for generating digital art composition have the disadvantage of capturing incomplete geometric information, which leads to obvious defects in the generation results. Therefore, a digital art composition generation method based on the multilabel learning algorithm is proposed in this research. Firstly, a preset series of grids are prepared to generate sampling and fractal pixels on the drawing base. Then, the preset grid construction is constructed by the interactive program of the preset grid library. After the stroke is drawn by the user, the actual motion trajectory of the pen is sampled by the digital panel, and the stroke information in the motion trajectory is obtained by the multilabel learning algorithm. Next, the steps of generating art composition are designed, including generating the skeleton of art composition, generating the geometric network structure of the skeleton, generating the sampling pixel and connecting the fractal pixel, and initializing other attributes of the mesh. Experimental results show that the proposed method has higher sampling rate and geometric information capture rate and has better application performance and prospect.

1. Introduction

As a new form of painting, digital painting is produced and perfected with the continuous development of digital technology [1]. Digital painting inherits the development characteristics of contemporary and modern painting art, including diversification and practicality. Specifically, diversification refers to the diversification of styles and expressions of painting. Before the 20th century, painting styles in each historical period often only had one mainstream form, and one style could represent a specific historical era, such as the Impressionism, Renaissance, and Christianity in the west, as well as scholars, academies, and Taoist schools in China. However, after the 20th century, there were many painting genres at the same time, such as Pop Art, Abstract Expressionism, Surrealism, Expressionism, Abstractionism, Dadaism, Cubism, Fauvism, and other painting genres with different forms of artistic thinking, which emerged one after another and then developed together [2]. In digital painting, these genres are prosperous. Practicality refers to the continuous enhancement of the practicality of the painting and the continuous expansion of the application field. Before the 20th century, whether it is in the east or the west, the expression of the painting was a work of art, and the rich and powerful were keen to regard it as a luxury for collection. In addition to preserving this form, modern painting has also penetrated into various fields. Digital painting has become a basic technical means in many fields, including film, animation, costume design, public art design, architectural landscape design, illustration, and advertising [3]. In the field of digital painting, the problem of digital art composition generation is the current research hotspot [4].

Great attention is paid to the research on the generation of digital art composition. Among them, due to the in-depth research on the digital painting, the international research on the generation of digital art composition has been carried out very early. Some scholars have proposed a digital art composition generation method based on the agent model. The Chinese research on the generation of digital art

composition has only lasted for decades, but certain research results have also been achieved. Some scholars have proposed a digital art composition generation method based on interactive technology.

Because the above methods have problems in sampling and capturing geometric information, a multilabel learning algorithm is applied in the process of research on the digital art composition generation method, and a digital art composition generation method based on multilabel learning algorithm is proposed.

2. Related Works

In this section, related works are reviewed and analyzed.

The objective of Ramanto et al. [5] was to make a framework that could haphazardly produce music that fits the mood from the manual client input. Markov chain is a stochastic model utilized in displaying the segments of the music structure. For the procedurally created to satisfy the mindset of the client, distinctive boundary esteems for every synthesis segment are apportioned for every mindset. These segments incorporate tempo, pitch range, note esteems, harmony type strength, and tune notes. The execution of the procedural music generation framework is then assessed by overviews and analyses. The assessment yielded results that guarantee the capacity of the music generation framework to fit the mood input.

In [6], a modeling technique for advancement is created utilizing an arithmetical model of plan details, prerequisites, and binary codes to apply formal strategies for check, model-based testing, just as techniques for mathematical coordinating. The details of the arithmetical equipment model have conducted polynomial math decided on a bunch of activities and practices.

In [7], a smart music composition neural network is proposed to naturally create a particular style of music. In this construction, the music grouping is obtained through an entertainer's long transient memory and then, at that point, fixed the likelihood of arrangement by a prize-based system that fills in as criticism to improve the exhibition of music synthesis. The music hypothetical principle is acquainted with compelling the style of produced music. A subjective approval in the trial is additionally proposed to ensure the predominance of the model contrasted and cutting edge works.

In [8], first, through the elaboration of the neural network techniques dependent on deep learning, the recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU) networks were presented, and the deep learning-based GRU-RNN programmed piece model was built. Second, the blockchain innovation was broken down and communicated, and the issues in the conventional copyright security and the board of computerized music were dissected. The three perspectives, i.e., possession, right of utilization, and right insurance, were completely thought of, and the blockchain innovation was incorporated into the copyright security and the board of computerized music. At last, the manual investigation assessment and delay examination were chosen as the pointers to dissect and portray the music structure nature of the GRU-RNN model, just as breaking down the improvement of the computerized music market incorporated with blockchain innovation. The outcomes positively affect advancing the turn of events and the utilization of deep learning strategies and blockchain innovation in computerized music.

It is generally acknowledged that computational advances shape the relationship of performers, instrument manufacturers, and writers with music, influencing different socio-social authenticities in music. In [9], it is reported that the possible ways of music production arise as a social build even because of the common collaboration with human performers and AI-controlled self-ruling instruments. It is contended that structure, making, and performing with a computerized instrument have gone through a slow sociomechanical change that has influenced workmanship, science, innovation, culture, and networks in general.

In [10], a novel technique for creating theoretical artworks is presented. The public artwork dataset WikiArt is utilized, and a K-means calculation is planned that consequently tracks down the ideal K worth to perform shading division on these pictures and partitions the images into various shading blocks. Tests showed that the subsequent unique artistic creation has an extraordinary visual effect, and some of them have been introduced as enhancements openly and private spaces, just as craftsmanship foundations. Likewise, a few specialists and architects have utilized the outcomes in their work.

An algorithm is introduced in [11] to produce digital painting lighting impacts from a solitary picture. The proposed calculation is content-mindful, with created lighting impacts normally adjusting to picture structures, and can be utilized as an intuitive instrument to improve on current labor-intensive work processes for producing lighting impacts for advanced and matte canvases. Likewise, this procedure can deliver usable lighting impacts for photos or three-dimensional delivered pictures. Results show that this methodology can altogether decrease the required collaboration time.

Computer-based intelligence innovation is generally utilized in different fields, including painting. Nonetheless, the mechanical idea of AI painting's own expressive procedures restricts its differentiated determination of painting materials and the setting of brush boundaries, and its showstoppers cannot accomplish the distinctive and acculturated expressive style of the conventional canvas. As a result of science and innovation, with the persistent advancement of science and innovation and the progression of time, the creative substance of works will increment. Advanced innovation ought to be viewed as an apparatus and means for imaginative creation. Just the ideal mix of computerized innovation and conventional composition structures can be helpful for the inventive advancement of workmanship and the feasible improvement of science and innovation and humanities. In light of this, the examination directed in [12] has concentrated on how to rationalistically take a gander at the incorporation of AI innovation and conventional artwork artistic expressions.

According to the discussions that took place, great attention is paid to the research on the generation of digital art composition. Among them, due to the in-depth research on digital painting, international research on the generation of digital art composition has been carried out very early. Generally, some scholars have proposed a digital art composition generation method based on the agent model. In addition, some others have proposed a digital art composition generation method based on interactive technology. However, the current methods have problems in sampling and capturing geometric information. Therefore, in this paper, a new digital art composition generation method based on a multilabel learning algorithm is proposed.

3. Design of the Digital Art Composition Generation Method Based on Multilabel Learning Algorithm

3.1. Preset Grid Construction. In the generation of digital art composition, a series of preset grids must be prepared first to generate sampling primitives and fractal primitives on the painting base [13]. To construct the preset grid through the preset grid library interactive program, the specific operations are as follows:

(1) Initialize the rectangular grid: the user inputs the grid resolution and the rectangular range, and the preset grid library interactive program can generate a uniformly distributed quadrilateral grid in the preset rectangular range and display it to the user [14].

(2) Read a grid in the preset grid library: the user selects a grid from the stored grid library, and the preset grid library interactive program displays the grid to the user [15].

(3) Select the 2D space transformation and apply it to the selected grid: select a transformation in the transformation function library, and enter the default parameters or the parameters that require transformation based on the interface prompts. When inputting the parameters, the user can preview the transformed grid. After the user confirms, the interactive program implements the transformation to the grid [16].

(4) Input the complex variable function to implement the complex transformation of the grid: input the function expression in the form of a string, and the interactive program implements the function to the grid [17]. When applying a complex variable function, the ordinate and abscissa of the grid are calculated as the imaginary and real parts of the complex number.

(5) Save the current grid in the preset grid library.

The specific process of creating the preset grid is shown in Figure 1 [18].

3.2. Get Stroke Information. The tilt, pressure, and position of the digital pen can be sensed by a digital panel. After the user draws the strokes, the digital panel is used to sample the actual motion trajectory of the digital pen, and the stroke information is obtained from the motion trajectory through the multilabel learning algorithm [19].

Five sampling sequences with the same length can be obtained during sampling, as shown in Table 1 [20].

Among the five acquired sampling sequences of the same length, the first four are data acquired by direct sensing digital pens, which need to be preprocessed as follows.

Sequence T is processed into a normalized time within a stroke; that is, the following calculation is performed for the absolute time value [21].

Step 1: conversion to relative time, which is achieved by

$$\begin{cases} t_n = t_n - t_1, \\ t_{n-1} = t_{n-1} - t_1, \\ \dots \\ t_2 = t_2 - t_1, \\ t_1 = 0. \end{cases} \tag{1}$$

Step 2: normalization, which is achieved by

$$t_i = \frac{t_i}{\max\left(t_1, t_2, \dots, t_n\right)}. \tag{2}$$

After the above transformation, $t_1 = 0$, $t_n = 1$, and other parameters are between t_1 and t_n [14].

All the coordinate values of the sequence C are directly transformed from the coordinate values of the drawing board to the coordinate values of the bottom of the drawing [15].

Sequence F is normalized by

$$f_i = \frac{f_i}{\max\left(f_1, f_2, \dots, f_n\right)}. \tag{3}$$

The value in the sequence Θ is directly converted into the corresponding radian value of the included angle of the x-axis [16].

Sequence V is output through these four sequences, and the specific steps are as follows.

Step 1: calculate the absolute speed, which is achieved by

$$\begin{cases} v_{xi} = \dfrac{x_i - x_{i-1}}{t_i - t_{i-1}}, \\ v_{yi} = \dfrac{y_i - y_{i-1}}{t_i - t_{i-1}}, \\ i > 1,\ v_{x1} = v_{x2},\ v_{y1} = v_{y2}. \end{cases} \tag{4}$$

Step 2: normalize the speed, which is achieved by

$$\begin{cases} v_{xi} \longleftarrow \dfrac{v_{xi}}{v_{\max}}, \\ \\ v_{yi} \longleftarrow \dfrac{v_{yi}}{v_{\max}}, \end{cases} \tag{5}$$

where $v_{\max}$ represents the maximum running speed of the digital pen, which is obtained by

$$\begin{cases} v_{\max} & = \max\left(\sqrt{v_{xi}^2 + v_{yi}^2}\right), \\ i & = 1, 2, \ldots, n. \end{cases} \tag{6}$$

The above sequences are all discrete stroke information, which needs to be transformed into a continuous parameter equation [22]. For five sequences, the sequence T is regarded as the parameters of other sequences, and the other sequences are transformed into continuous parameter equations through linear interpolation as expressed by

$$\begin{cases} C(t) = \tau C_i + (1-\tau)C_{i+1}, \\ F(t) = \tau F_i + (1-\tau)F_{i+1}, \\ \Theta(t) = \tau \Theta_i + (1-\tau)\Theta_{i+1}, \\ V(t) = \tau V_i + (1-\tau)V_{i+1}, \end{cases} \quad t \in [t_i, t_{i+1}], \tag{7}$$

where τ represents the continuous parameter threshold, and it can be calculated by

$$\begin{cases} \tau = \dfrac{t - t_i}{t_{i+1} - t_i}, \\ \\ i = 1, 2, \ldots, n. \end{cases} \tag{8}$$

According to the above parameter equations, a multilabel learning algorithm is used to classify and derive the corresponding geometric information of the stroke trajectory [18].

Derived $C(t)$ is expressed in the component form, as given by

$$C(t) = \begin{cases} x(t), \\ y(t), \end{cases} \tag{9}$$

where $x(t)$ and $y(t)$ represent the component coordinates of $C(t)$ [19].

According to the above formula, the corresponding equation of the corresponding tangential parameter $C'(t)$ of the stroke trajectory is given by

$$C'(t) = \begin{cases} x'(t), \\ y'(t), \end{cases} \tag{10}$$

where $x'(t)$ and $y'(t)$ represent the component coordinates of the tangential parameter $C'(t)$ [20].

The equation corresponding to the unit tangential parameter is as follows:

$$C''(t) = \begin{cases} C_x(t) \\ C_y(t) \end{cases} = \begin{cases} \dfrac{x'(t),}{C'(t),} \\ \\ \dfrac{y'(t),}{C'(t),} \end{cases} \tag{11}$$

where $C''(t)$ represents the unit tangential parameter and $C_x(t)$ and $C_y(t)$ represent the component coordinates of the unit tangential parameter.

The normal parameter equation corresponding to the stroke trajectory is as follows:

$$N(t) = \begin{cases} N_x(t) = x'(t)\cos\left(-\dfrac{\pi}{2}\right) - y'(t)\sin\left(-\dfrac{\pi}{2}\right), \\ \\ N_y(t) = x'(t)\sin\left(-\dfrac{\pi}{2}\right) - y'(t)\cos\left(-\dfrac{\pi}{2}\right), \end{cases} \tag{12}$$

where $N(t)$ represents the normal parameter corresponding to the stroke trajectory and $N_x(t)$ and $N_y(t)$ represent the component coordinates of the normal parameter [21].

In this way, one stroke can be used to derive the multiparameter equations of geometric information through the following:

$$\begin{cases} S(t) = \langle C(t), C'(t), C'(t), N(t), N(t), F(t), \Theta(t), V(t) \rangle, \\ t \in [0, 1], \end{cases} \tag{13}$$

where $S(t)$ represents the multiparameter equation group of the derived geometric information, i.e., the stroke information.

3.3. Art Composition Generation. Based on the obtained stroke information, the art composition is generated in four steps. The first step is to generate the skeleton of the art composition, the second step is to generate the geometric network structure of the skeleton, and the third step is to generate sampling primitives and connect fractals. For primitives, the fourth step is to initialize other properties of the grid [23].

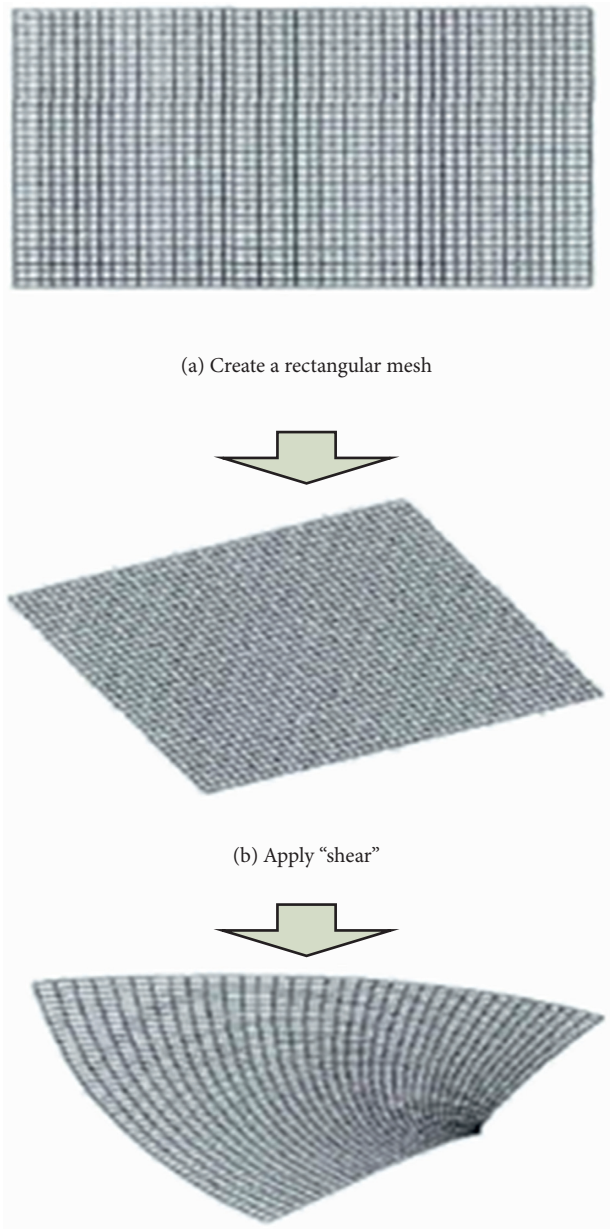
The specific steps are as follows:

Step 1: the skeleton of the art composition is the central axis, which is a series of coordinate values obtained by determining the coordinate values of each vertex.

Step 2: extend each node of the skeleton to both sides to generate branch lines with nodes, and connect all branch lines to form a grid. The extension mechanism is controlled by adjusting the preset parameters to control the overall structure of the grid.

Step 3: generate a quadrilateral grid with the same resolution as the fractal primitive; that is, sample the primitive, and connect it with the fractal primitive.

Step 4: the first three steps have determined the specific position of the art composition, and then initialize other parameters based on the preset parameters, including attenuation parameters, image processing functions, activity periods, color channels, and input terminals.



(a) Create a rectangular mesh

(b) Apply "shear"

(c) Imposed transformation "f (z) = Z^2"

Figure 1: The specific process of creating a preset grid.

4. Experiment of Art Composition Generation

4.1. Experimental Design. In order to verify the performance of the digital art composition generation method based on the multilabel learning algorithm, an experiment was carried out. The tool used in the experiment was a digital painting tool, and its framework is shown in Figure 2.

The digital painting tool can implement a total of five painting modes, including collage, prepainting, selection, brush, and color mode. The specific implementation methods are shown in Table 2.

This digital painting tool was used to test the performance of the digital art composition generation method based on multilabel learning algorithm. The sampling rate and geometric information capture rate of this method were obtained and used as experimental data. In order to avoid the lack of convincing results of this experiment, the two existing digital art composition generation methods were used for comparison in the experiment. These two existing digital art composition generation methods included agent-based and interactive technology-based digital art composition generation methods. The performance test of these

Table 1: Five sampling sequences of the same length obtained during sampling.

Sequence marker	Sequence element tags	Meaning
T	$t_1, t_2, \ldots, t_m$	Represents the normalized relative time of each sampling point
C	$C_1 = (x_1, y_1), C_2 = (x_2, y_2), \ldots, C_n = (x_n, y_n)$	Trajectory coordinates of the digital pen on the bottom
F	$f_1, f_2, \ldots, f_n$	Pressure applied to the drawing board by the digital pen
Θ	$\theta_1, \theta_2, \ldots, \theta_n$	Tilt direction of the digital pen
V	$(vx_1, vy_1), (vx_2, vy_2), \ldots, (vx_n, vy_n)$	Indicates the speed of the digital pen during operation

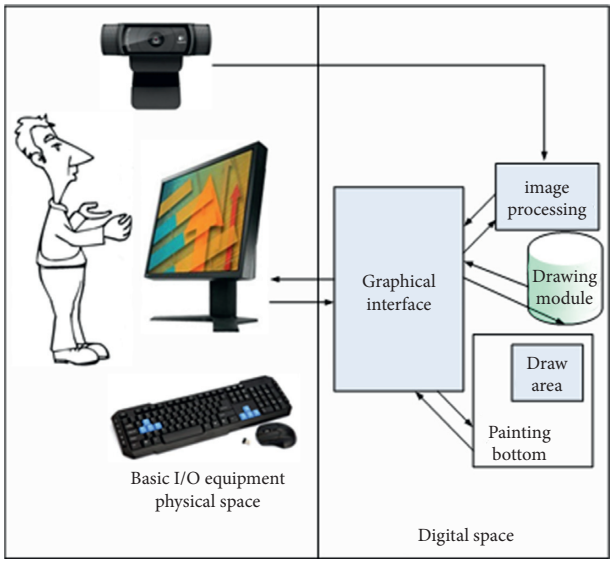


Figure 2: Framework of the digital painting tool.

Table 2: Realization of painting modes.

Number	Painting mode	Implementation mode
1	Color matching mode	Add some color processing methods in image processing, and let users interact through graphics.
2	Brush	The interface specifies the processing parameters and controls the color of the output image.
3	Constituency	An image processing mechanism, which can filter out a special target in the video image.
4	Preprocessing the background	The user can draw with the object in the way of "stroke" in front of the camera.
5	Collage	Let the user specify a specific area on the bottom of the painting. When painting, the pigment can only affect the selected area.

two methods was also performed, and the sampling rate and geometric information capture rate of these two methods were obtained and used as comparative experimental data. The sampling rate and geometric information capture rate of the three digital art composition generation methods are analyzed and compared.

4.2. Experimental Results

4.2.1. Sampling Rate. In the test of digital art composition generation, the sampling rate data of the digital art composition generation method based on the multilabel learning algorithm and based on the agent model and interactive technology are shown in Table 3. In addition, the results are also displayed in Figure 3 for a better comparison.

According to Table 3 as well as Figure 3, in the test of digital art composition generation, the sampling rate of the digital art composition generation method based on the multilabel learning algorithm was higher than that of the method based on the agent model and interactive technology.

4.2.2. Geometric Information Capture Rate. Table 4 presents the geometric information capture rate of the digital art composition generation method based on the multilabel learning algorithm, agent model, and interactive technology. In addition, the results are also displayed in Figure 4 for a better comparison.

According to Table 4 as well as Figure 4, in the test of digital art composition generation, the geometric information capture rate of the digital art composition generation method based on the multilabel learning algorithm was higher than that of the method based on the agent model and interactive technology.

Based on the above experimental results, the sampling rate and geometric information capture rate of the digital art

Table 3: Experimental data of sampling rate.

Grid size	Sampling rate (frame/s)		
	Method based on multilabel learning algorithm	Method based on the agent model	Method based on interactive technology
1 × 1	32.5	25.3	22.3
2 × 2	30.8	23.6	20.4
3 × 3	28.6	22.9	17.3
4 × 4	26.7	21.3	16.8
5 × 5	25.1	21.0	15.2
6 × 6	23.8	18.0	13.7
7 × 7	21.7	15.9	11.7
8 × 8	18.2	15.2	10.2
9 × 9	17.9	13.2	9.3
10 × 10	17.2	11.8	9.0

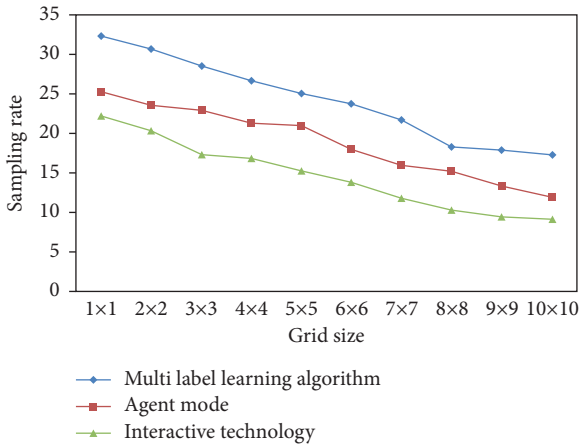


Figure 3: Results of the sampling rate.

Table 4: Experimental data of geometric information capture rate.

Grid size	Geometric information capture rate (%)		
	Method based on multilabel learning algorithm	Method based on the agent model	Method based on interactive technology
1 × 1	98.34	94.23	89.62
2 × 2	98.32	94.22	89.60
3 × 3	98.30	94.19	89.59
4 × 4	98.27	94.17	89.61
5 × 5	98.25	94.15	89.58
6 × 6	98.21	94.13	89.50
7 × 7	98.20	94.11	89.54
8 × 8	98.18	94.18	89.53
9 × 9	98.17	94.08	89.50
10 × 10	98.13	94.10	89.55

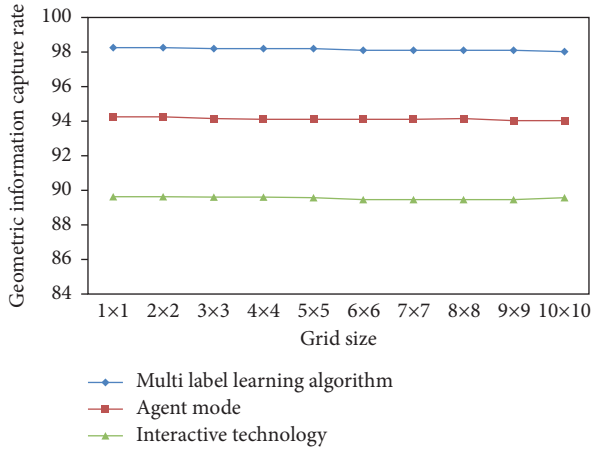


Figure 4: Results of the geometric information capture rate.

composition generation method based on the multilabel learning algorithm were higher than those of the two comparative digital art composition generation methods, indicating its better performance.

5. Conclusion

By designing a digital art composition generation method based on multilabel learning algorithm, the sampling rate and geometric information capture rate are improved. However, the stroke information data obtained in the design of the method contain redundant data, and the processing of the redundant data will be investigated in the future.

Data Availability

All the data used to support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] B. Coşkun Onan, "Implementing contemporary art theory to instruction of painting in undergraduate level: a phenomenology study," *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, vol. 30, no. 2, pp. 725–755, 2018.

[2] C. Schmidt, M. Santos, L. Bohn et al., "Dynamic balance and mobility explain quality of life in HFpEF, outperforming all the other physical fitness components," *Arquivos Brasileiros de Cardiologia*, vol. 114, no. 4, pp. 701–707, 2020.

[3] D. Aras, C. Ç. Bildircin, Ö. Güler, M. Gülü, and F. Akça, "Spor tirmaniş yetenek Seçimi test bataryasi Örneği," *Ankara Üniversitesi Beden Eğitimi ve Spor Yüksekokulu SPORMETRE Beden Eğitimi ve Spor Bilimleri Dergisi*, vol. 17, no. 2, pp. 41–52, 2019.

[4] F. Schwalbe, "The importance of student leagues on medical training in neurosurgery and residency choice," *Arquivos Brasileiros De Neurocirurgia Brazilian Neurosurgery*, vol. 37, no. 01, pp. 13–18, 2018.

[5] R. Ramanto, A. Sigit, and N. U. Maulidevi in *Proceedings of the ADADA 2016 14TH international Conference for Asia Digital Art and Design Association-Mood-based Procedural Music Generation Using Markov Chains*, Bali, Indonesia, November 2016.

[6] O. O. Letychevskyi, V. S. Peschanenko, V. S. Kharchenko, V. A. Volkov, and O. M. Odarushchenko, "Modeling method for development of digital system Algorithms based on programmable logic devices," *Cybernetics and Systems Analysis*, vol. 56, no. 5, pp. 710–717, 2020.

[7] C. Jin, Y. Tie, Y. Bai, X. Lv, and S. Liu, "A style-specific music composition neural network," *Neural Processing Letters*, vol. 52, no. 3, pp. 1893–1912, 2020.

[8] H. Li, "Piano automatic computer composition by deep learning and blockchain technology," *IEEE Access*, vol. 8, pp. 188951–188958, 2020.

[9] K. Tahiroğlu, "Ever-shifting roles in building, composing and performing with digital musical instruments," *Journal of New Music Research*, vol. 50, no. 2, pp. 155–164, 2021.

[10] M. Li, J. Lv, J. Wang, and Y. Sang, "An abstract painting generation method based on deep generative model," *Neural Processing Letters*, vol. 52, pp. 1–12, 2019.

[11] L. Zhang, E. Simo-Serra, Ji Yi, and C. Liu, "Generating digital painting lighting effects via RGB-space geometry," *ACM Transactions on Graphics*, vol. 39, no. 2, pp. 1–13, 2020.

[12] X. Liu, "Artistic reflection on artificial intelligence digital painting," in *Journal of Physics: Conference Series*vol. 1648, no. 3, IOP Publishing, Article ID 032125, 2020.

[13] H. Liu, Z. Mi, L. Lin et al., "Shifting plant species composition in response to climate change stabilizes grassland primary production," *Proceedings of the National Academy of Sciences*, vol. 115, no. 16, pp. 4051–4056, 2018.

[14] G. Incerti, E. Cecconi, F. Capozzi et al., "Infraspecific variability in baseline element composition of the epiphytic lichen Pseudevernia furfuracea in remote areas: implications for biomonitoring of air pollution," *Environmental Science and Pollution Research*, vol. 24, no. 9, pp. 1–13, 2020.

[15] X. Gan, G. Xu, G. Zhao, M. Zhou, and Z. Cai, "Composition optimization of Nb-Ti microalloyed high strength steel," *Journal of Wuhan University of Technology-Mater. Sci. Ed.*, vol. 33, no. 5, pp. 1193–1197, 2018.

[16] J. Liu, S. Du, F. Li, H. Zhang, and S. Zhang, "Preparation of ZrB2-SiC powders via carbothermal reduction of zircon and prediction of product composition by back-propagation artificial neural network," *Journal of Wuhan University of Technology-Mater. Science Education*, vol. 33, no. 5, pp. 1062–1069, 2018.

[17] Z. Dongdong, G. Hao, L. Jingfeng et al., "Effect of Cu-Ti-C reaction composition on reinforcing particles size of TiC x/Cu composites," *Journal of Wuhan University of Technology-Mater. Science Education*, vol. 33, no. 1, pp. 43–48, 2018.

[18] E. S. M. Younis, A. S. Al-Quffail, N. A. Al-Asgah et al., "Effect of dietary fish meal replacement by red algae, Gracilaria arcuata, on growth performance and body composition of Nile tilapia *Oreochromis niloticus*," *Saudi Journal of Biological Sciences*, vol. 25, no. 2, pp. 198–203, 2017.

[19] B. O. Otegbayo, D. J. Oguniyan, B. A. Olunlade, O. O. Oroniran, and O. E. Atobatele, "Characterizing genotypic variation in biochemical composition, anti-nutritional and mineral bioavailability of some Nigerian yam (Dioscorea spp.) land races," *Journal of Food Science and Technology*, vol. 55, no. 1, pp. 205–216, 2018.

[20] B. Basu, O. Chowdhury, and S. Saha, "Possible link between stress-related factors and altered body composition in women with polycystic ovarian syndrome," *Journal of Human Reproductive Sciences*, vol. 11, no. 1, pp. 10–18, 2018.

[21] J. Leskovec, A. Levart, S. A. Nemec et al., "Effects of supplementation with α-tocopherol, ascorbic acid, selenium, or their combination in linseed oil-enriched diets on the oxidative status in broilers," *Poultry Science*, vol. 97, no. 1, pp. 86–93, 2018.

[22] M. Alessandra, D. Z. Antonio, E. Graziana et al., "Age-related changes to human tear composition," *Investigative Ophthalmology & Visual Science*, vol. 59, no. 5, pp. 2024–2031, 2018.

[23] N. Royaei, T. Shahrabi, and Y. Yaghoubinezhad, "The investigation of the electrocatalytic and corrosion behavior of a TiO2-RuO2 anode modified by graphene oxide and reduced graphene oxide nanosheets via a sol-gel method," *Catalysis Science & Technology*, vol. 8, no. 19, pp. 4957–4974, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 9928020, 5 pages
https://doi.org/10.1155/2021/9928020

Research Article

Path Recognition of the Regional Education Expansion Based on Improved Dragonfly Algorithm

Fang Liu

Scientific Research Department, Xijing University, Xi'an 710123, China

Correspondence should be addressed to Fang Liu; 20020041@xijing.edu.cn

Received 3 June 2021; Revised 7 July 2021; Accepted 28 July 2021; Published 7 September 2021

Academic Editor: Hussein Abulkasim

To solve the problems of low recognition rate, high misrecognition rate, and long recognition time, the path recognition method of the regional education scale expansion based on the improved dragonfly algorithm is proposed. Through a variety of different behaviors utilized in the optimization process, the dragonfly algorithm model has been constructed. The step size and the position vector are introduced to update the dragonfly's location. The dragonfly's foraging behaviors are accurately simulated. Afterward, the dragonfly algorithm is combined with the flower authorization algorithm. The conversion probability is added, and the dragonfly's global development ability is adjusted in real-time. Then, the dragonfly algorithm is improved. The improved dragonfly algorithm is employed to extract the features of the expansion path of the regional education scale. The improved support vector machine is utilized as a classifier to realize the recognition of the regional education scale expansion path. The experimental results denote that the proposed method has a high recognition rate of the regional education scale expansion path and can effectively reduce the misrecognition rate and shorten the recognition time.

1. Introduction

China's higher education has been currently in a brand-new period from scale expansion to connotation enhancement. Because of the obvious differences occurring in the development of higher education in various regions, it is very significant to recognize the path of the regional education scale expansion [1]. Research has been conducted in this field. For example, Dong and Zhang [2] analyzed the diffusion process of the subject and integrated network topology analysis and subject semantic mining for path recognition, which could effectively improve the accuracy of the path recognition research, but it had a longer recognition time. Cai et al. [3] introduced particle filter technology into the field of path deception recognition. They constructed corresponding state equations and observation equations, utilized dynamic data to extract key parameters, and then achieved path recognition. Their method could effectively improve the recognition efficiency, but the feature selection was not considered, which led to a high misrecognition rate of this method. Dealing with the above problems, a path recognition method for the expansion of the regional education scale based on the improved dragonfly algorithm is proposed in this manuscript.

According to the many different behaviors utilized in the optimization process of the dragonfly algorithm, the dragonfly algorithm model is constructed. Then, the step length and position vector are introduced. The position of the dragonfly is updated, and the foraging behavior of the dragonfly is simulated. Combining with the flower authorization algorithm, the conversion probability is added to conduct global search work. Thus, the dragonfly algorithm is improved. The improved dragonfly algorithm is employed to extract the features of the expansion path of the regional education scale, and the improved support vector machine is utilized as a classifier to recognize the expansion path of the regional education scale. The simulation experiment results show that the proposed method can effectively reduce the misrecognition rate and the recognition time while improving the recognition rate.

The rest of the manuscript is organized as follows. Section 2 gives detailed information on both the dragonfly algorithm and flower authorization algorithm. Then, the improved algorithm has been proposed. Section 3 deals with the implementation of the proposed method. Section 4 is

allocated to present the simulation experiments using different datasets. The conclusion is presented in Section 5.

2. Improved Dragonfly Algorithm

2.1. Dragonfly Algorithm. Compared with most intelligent optimization algorithms, the dragonfly algorithm has the global search capability as the foraging behavior of the dragonfly; the local exploration capability is the migration behavior of the dragonfly [4, 5]. The dragonfly optimization algorithm contains many different behaviors in the optimization process.

The first behavior is called collision avoidance behavior. When a dragonfly flies, safety issues need to be prioritized as it needs to keep a safe distance from other dragonflies. The second behavior is called team behavior. Dragonflies are also animals living in groups. Therefore, dragonflies will move in groups in the entire population. To avoid collisions, it is necessary to maintain an appropriate flight distance between each dragonfly [6]. The third behavior is called gathering behavior. As a group, dragonflies have a leader. The dragonflies around the leader need to be close together, and the spacing between the dragonflies needs to be consistent [7, 8]. The fourth behavior is called foraging behavior. To get better and more food, dragonflies need to pass messages between them and keep moving closer to get more and better food. The last behavior is called enemy avoidance behavior. In the process of searching for food, dragonflies may be caught by other creatures [9]. Therefore, when a dragonfly sees its natural enemy, it will avoid it, but the probability of encountering a natural enemy is low.

The abovementioned behaviors are abstracted, and the corresponding algorithm model is formed at the same time. Among them, the displacement of a single dragonfly in the collision behavior can be calculated as follows:

$$S_i = -\sum_{j=1}^{N} X_i - X_j, \tag{1}$$

where S_i represents the displacement formed by the ith dragonfly individual in the collision behavior; X_i represents the current specific coordinate position of the ith dragonfly; X_j represents the iterative position of the ith dragonfly adjacent to the jth dragonfly in the dragonfly population; and N represents the population of dragonflies [10].

Equation (2) is used to give the displacement of a single dragonfly individual in the pairing behavior, namely,

$$A_i = \frac{\sum_{j=1}^{N} X_j}{N}, \tag{2}$$

where A_i represents the displacement formed by the ith dragonfly in the group.

Dragonflies also form displacements in their aggregation behavior [11, 12], and the specific expression is defined by

$$C_i = \frac{\sum_{j=1}^{N} X_j}{N} - X_i, \tag{3}$$

where C_i represents the displacement formed by the ith dragonfly in the group gathering behavior.

The displacement formed by the individual dragonfly during foraging is as follows:

$$F_i = X_{\text{food}} - X_i, \tag{4}$$

where F_i represents the displacement formed by the ith dragonfly in the foraging process and X_{food} represents the exact coordinate position of the food that the dragonfly needs to obtain under the set number of iterations [13–15].

When a dragonfly individual is avoiding the enemy, it will produce a certain displacement. The specific expression is as follows:

$$E_i = X_e + X_i, \tag{5}$$

where E_i represents the displacement formed by the dragonfly individual in the process of avoiding enemies and X_e represents the position of the natural enemy discovered by the dragonfly under the set number of iterations.

In the entire dragonfly algorithm, it is assumed that the entire foraging process of the dragonflies is a combination of the five behaviors mentioned above. To improve the recognition accuracy of the dragonfly algorithm, it is necessary to continuously update the specific position of the dragonflies in the iterative process and, at the same time, accurately simulate the behavior of the dragonflies in foraging. Thus, two vectors are introduced.

(1) Step vector ΔX

(2) Position vector ΔY

The above two vectors are mainly used to update the position of the dragonfly in the next iteration, where the step vector ΔX represents the flying position of the dragonfly individual in the current iteration; the step vector of the ith dragonfly in the $t+1^{\text{th}}$ iteration is

$$\Delta x_{\text{id}}^{t+1} = \Delta Y\left(sS_{\text{id}}^{t} + aA_{\text{id}}^{t} + cC_{\text{id}}^{t} + fF_{\text{id}}^{t} + eE_{\text{id}}^{t}\right) + \Delta X, \tag{6}$$

where $\Delta x_{\text{id}}^{t+1}$ represents the d-dimensional step vector corresponding to the ith dragonfly during the $t+1^{\text{th}}$ iteration process [16]; s represents the impact weight of the collision behavior; S_{id}^{t} represents the deviation of the corresponding d-dimensional due to the collision of the ith dragonfly individual during the ith iteration process; a represents the influence weight of team behavior; A_{id}^{t} represents the deviation of the ith dragonfly due to team behavior during the ith iteration; c represents the inertia weight; C_{id}^{t} represents the deviation of the corresponding d-dimensional of the ith dragonfly individual during the ith iteration due to inertia; f represents the influence weight of dragonfly position; F_{id}^{t} represents the deviation of the corresponding d-dimensional of the ith dragonfly individual during the ith iteration due to the position of the dragonfly; e represents the influence weight of the dragonfly individual; and E_{id}^{t} represents the corresponding d-dimensional of the ith dragonfly individual during the ith iteration due to the formation of the dragonfly individual's deviation.

Analyzing the influence weights of five different behaviors, we can know that the algorithm can adjust dragonfly's global development capability and local detective

capability in real-time during the process of obtaining the optimal solution [17, 18]. Neighboring individuals have a great influence on dragonflies. Therefore, in the entire algorithm, it is assumed that each single dragonfly individual has a certain radius neighbor.

When the step vector calculation is completed, the position of the dragonfly is updated by

$$X_{\mathrm{id}}^{t+1} = X_{\mathrm{id}}^{t} + \Delta X_{\mathrm{id}}^{t+1}, \tag{7}$$

where the position update formula of the ith dragonfly is

$$X_{\mathrm{id}}^{t+1} = X_{\mathrm{id}}^{t} + \mathrm{Levy}(d) \times X_{\mathrm{id}}^{t}, \tag{8}$$

where the Levi flight calculation formula is

$$\mathrm{Levy}(x) = 0.01 \times \frac{r_1 \times \sigma}{\left|r^2\right|^{1/p}}, \tag{9}$$

where $\mathrm{Levy}(x)$ represents the Levi random path, σ represents the step size scaling factor, and p and r represent a random number subject to a uniform distribution.

2.2. Flower Authorization Algorithm. The behavior of flowers using pollination to reproduce offspring is very common in nature. It is necessary to reproduce offspring through all the flowers, and the process of pollination must be combined with energy. Due to the certain differences between different flowers, the pollination methods of flowers also differ greatly. In nature, there exist all kinds of plants, and the flowers that occupy a larger proportion are blooming flowers. In addition to various extreme environments, there exist blooming flowers in all regions. Since most of the blooming flowers cannot exist, the entire pollination process of flowers requires the help of external forces such as bees and insects [19].

In the flower authorization algorithm, the number of pollens is set as N and the dimension of the solution space is D. When pollen is authorized for cross pollen, pollen mainly conducts global search and updates. When pollen is pollinated, local exploration work is mainly performed, and location updates are also required.

In the process of flower authorization, to effectively ensure the simple structure of the whole algorithm and strong optimization ability, the conversion probability p is added to the pollen authorization algorithm to balance and adjust strategy transformation. In the early stage of pollination, r, a random value, will be formed, and the value of r must conform to the law of uniform distribution.

2.3. Improvement of the Dragonfly Algorithm. Combining the dragonfly algorithm with the flower authorization algorithm, we set the value range of each dragonfly during initialization to $[S_{\min}, S_{\max}]$through the difference of the problem-solution space. To effectively avoid the situation where the dragonfly algorithm gathers together in the initial stage, the uniform random distribution is used in the initialization process to perform random initialization processing on each dimension of each dragonfly, namely,

$$x_{\mathrm{id}}^{0} = S_{\min} + \left(S_{\max} - S_{\min}\right). \tag{10}$$

After preprocessing, the two algorithms are further fused to guide the dragonfly to fly to a better position, thereby filling the shortcomings of the dragonfly algorithm.

3. Path Recognition of the Regional Education Expansion Based on Improved Dragonfly Algorithm

To better recognize the regional education scale expansion path, it is necessary to prioritize the feature selection of the regional education scale expansion path. If the value of any feature is 0, it means that the feature does not need to be selected; if the value corresponding to any feature is 1, then the feature is selected. The solution space of the problem can be transformed into two trends $\{0, 1\}$. To make the initial features have a certain difference, each dragonfly needs to obey the distribution as much as possible. In the improved dragonfly algorithm, one-dimensional initialization processing is performed on each dragonfly in the dragonfly population through

$$x_{\mathrm{id}}^{0} = \begin{cases} 0, & r < 0.5, \\ 1, & r \geq 0.5, \end{cases} \tag{11}$$

where x_{id}^{0} represents the specific coordinate position of the ith dragonfly in the d-dimensional ith during the initial iteration, and the unit of r must conform to the random value of the uniform distribution rule.

After the above calculations, the position of the dragonfly can be updated in combination with related formulas, and a global search can be performed. At the same time, the support vector machine is set as the classifier. For the dataset $\{x_i, y_i\}$, d represents the total number of features in the feature set and y_i represents the total output of the dataset. Under normal circumstances, the support vector machine algorithm needs to set a linear decision function, and its specific expression is

$$f(x) = \langle w, \phi_i(x) \rangle + b, \tag{12}$$

where w represents the weight in the support vector machine; b represents a random constant [20]; and ϕ represents the nonlinear function that maps the input feature to a higher feature space. Equation (12) can be transformed into

$$R(C) = \frac{C}{N \sum_{i=1}^{n} L_s\left(f(x_i), y_i\right) + 1/2\|w\|^2}, \tag{13}$$

where $L_s(f(x_i), y_i)$ represents the loss function of ε . When $|f(x) - y| \geq \varepsilon$, $L_s(f(x_i), y_i)$ can be solved by equation (15), and the specific expression is

$$L_s\left(f(x_i), y_i\right) = |f(x) - y| - \varepsilon. \tag{14}$$

In the process of increasing the slack variables ξi and ξi^*, it is necessary to satisfy the three constraints of equation (15) at the same time and minimize the problem, as shown in equation (16)

Table 1: Comparison results of the recognition rate of the regional education scale expansion path by the different methods.

Dataset	The recognition rate of the expansion path of the regional education scale (%)		
	The proposed method	The method in [2]	The method in [3]
01	98.63	96.47	94.25
02	97.14	94.88	92.02
03	96.22	93.96	91.00
04	98.96	92.74	88.36
05	97.14	91.22	87.74
06	96.20	90.52	86.74

Table 2: Comparison results of the recognition time of the regional education scale expansion path by the different methods.

Data volume (MB)	Recognition time of the regional education scale expansion path (s)		
	The proposed method	The method in [2]	The method in [3]
20	1.25	3.41	2.36
40	2.34	5.47	4.42
60	3.37	7.52	5.48
80	4.23	9.41	7.37
100	6.10	10.32	9.23
120	8.20	16.32	12.27

$$\begin{cases} y_i - \langle w, x_i \rangle - b \le \varepsilon + \xi_i, \\ \langle w, x_i \rangle + b - y_i \le \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \ge 0, \end{cases} \tag{15}$$

$$R\left(w, \xi_i^*\right) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right). \tag{16}$$

Based on the above analysis, the regional education scale expansion path classification is carried out through the classifier, and the recognition is finally realized:

$$K\left(x_i, x_j\right) = \phi\left(x_i\right) \cdot \phi\left(x_j\right). \tag{17}$$

4. Simulation Experiment

To verify the comprehensive validation of the proposed method, a simulation experiment was performed on the Windows 10 64-bit system, with Python language as a programming environment.

4.1. The Recognition Rate of the Expansion Path of Regional Education Scale (%). The experiment selected the recognition rate of the regional education scale expansion path as the test index and selected the methods in [2] and in [3] for comparison. The specific experimental comparison results are shown in Table 1.

Analyzing the experimental data in Table 1, it can be seen that, due to the different datasets, the recognition rate of the regional education scale expansion path of the different methods had changed to different degrees. However, when compared with the other two methods, the proposed method remained a relatively more stable and larger recognition rate of the regional education scale expansion path, while the other two methods showed a linear decline. The main reason was that the proposed method used an improved dragonfly algorithm to optimize and recognize it so that it could accurately recognize each path and obtain a higher recognition rate.

4.2. Recognition Time of the Expansion Path of Regional Education Scale (s). The following experiment mainly compared the recognition time of the regional education scale expansion path of the different methods. The specific experimental comparison results are shown in Table 2.

Analyzing the experimental data presented in Table 2, as the amount of test data increased, the recognition time of the regional education scale expansion path increased for all three methods. When the volume of the test data reached 120 MB, the recognition time of the regional education scale expansion path based on the method in [2] was 16.32 s, that of method in [3] was 12.2 7 s, and that of the proposed method was only 8.20 s. It could be seen that the proposed method had a shorter recognition time for the expansion of the regional education scale. This was because the proposed method combined the dragonfly algorithm with the flower authorization algorithm and added the conversion probability, thereby achieving the real-time adjustment of the dragonfly's global development capabilities and reducing the recognition time.

4.3. Misrecognition Rate (%). To fully verify the superiority of the proposed method, the following experimental tests compared with the misrecognition rates of the different methods. The specific experimental comparison results are shown in Figure 1.

The results in Figure 1 show that, as the amount of test data increased, the misrecognition rate of the different methods increased. When the test data volume reached 120 MB, the misrecognition rate of the method in [2] was 15%, that of the method in [3] was 18.2%, and that of the proposed method was 8.3%. It could be seen that the miss-recognition rate of the proposed method was significantly

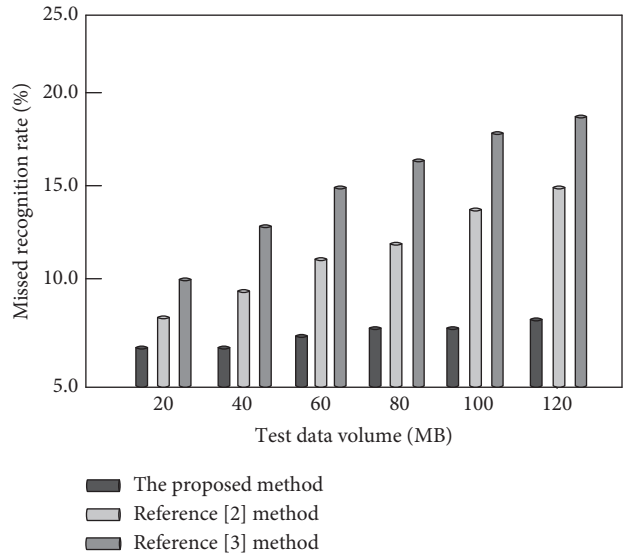


Figure 1: Comparison results of misrecognition rates of different methods.

lower than that of the other two methods, and it also verified the effectiveness and feasibility of adding the improved dragonfly algorithm to the path recognition of the regional education scale expansion.

5. Conclusion

To improve the recognition rate of the regional education scale expansion path, reduce the misrecognition rate, and shorten the recognition time, a method based on the improved dragonfly algorithm is proposed for the recognition of the regional education scale expansion path. Dragonfly algorithm combined with flower authorization algorithm leads to the improved dragonfly algorithm utilizing both step size and position vector and the conversion probability.

Therefore, the improved dragonfly algorithm could be used to extract the features of the expansion path of the regional education scale. Afterward, the improved support vector machine is utilized to recognize the expansion path of the regional education scale. The simulation experiment results show that the proposed method can effectively improve the recognition rate of the regional education expansion path and reduce the recognition time and the misrecognition rate.

Data Availability

All data are available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] B. Liu, C. Yang, and L. Xiong, "Path recognition method of picking robot in orchard natural environment," *Jiangsu Journal of Agriculture*, vol. 35, no. 5, pp. 1222–1231, 2019.

[2] K. Dong and B. Zhang, "Research progress on path recognition of subject knowledge diffusion network," *Information Studies: Theory & Application*, vol. 40, no. 8, pp. 139–144, 2017.

[3] Z. Cai, R. Ju, X. Xu, and S. Wang, "Deceptive path recognition based on particle filter technology," *Systems Engineering and Electronics*, vol. 42, no. 2, pp. 374–380, 2020.

[4] Z. Lin, "Path identification of internet entrepreneurship opportunities based on long tail theory," *China Adult Education*, vol. 10, no. 5, pp. 66–68, 2017.

[5] H. Wang, N. Wang, and J. Lian, "Vibration transmission path identification of hydropower plant based on transfer entropy," *Journal of Hydraulic Engineering*, vol. 49, no. 6, pp. 732–740, 2018.

[6] C. Ding and H. Xiao, "The characteristics and path identification of China's animal husbandry production fluctuations," *Research of Agricultural Modernization*, vol. 39, no. 2, pp. 256–264, 2018.

[7] J. Li, Y. Li, and J. Peng, "Research on path recognition method of agricultural robot visual navigation," *Computer Engineering*, vol. 44, no. 9, pp. 38–44, 2018.

[8] P. Deng, R. Yang, and S. Deng, "A new LPMD-based scattering path recognition algorithm," *Journal of Southwest Jiaotong University*, vol. 53, no. 1, pp. 182–188, 2018.

[9] S. Yan, G. Duan, and S. Zhang, "A device-independent indoor positioning algorithm based on line-of-sight path recognition," *Computer Engineering and Science*, vol. 40, no. 8, pp. 80–87, 2018.

[10] L. Zhao, C. Ye, Y. Zhang, X. Xu, and J. Chen, "Path recognition method of robot visual navigation in an unstructured environment," *Acta Optics*, vol. 38, no. 8, Article ID 0815028, 2018.

[11] H. Wang, W. Linjin, and Q. Zhang, "Path identification and change of the Kuroshio to the east of Taiwan," *Ocean & Limnology*, vol. 49, no. 2, pp. 271–279, 2018.

[12] Z. Mao and Q. Chen, "AGV multi-branch path recognition and tracking based on PCA-LDA and SVM," *Progress in Laser and Optoelectronics*, vol. 55, no. 9, pp. 148–155, 2018.

[13] J. Ding and X. Zhang, "Research on subway passenger path recognition based on mobile phone signaling data," *Journal of China Academy of Electronics*, vol. 14, no. 11, pp. 1194–1201, 2019.

[14] F. Wu, X. Chen, and L. Huang, "Research on technology multi-topic evolution path recognition method based on semantic similarity," *Journal of Information*, vol. 37, no. 5, pp. 91–96, 2018.

[15] S. Fan, X. An, and L. Shan, "Research on subject evolution type and evolution path identification method based on medical reference," *Information Theory and Practice*, vol. 42, no. 3, pp. 114–119, 2019.

[16] Y. Bai and L. Shi, "Spatial spillover of technological innovation in the Yangtze river economic belt: effect measurement, path recognition and cooperative amplification," *Journal of Wuhan University (Natural Science Edition)*, vol. 72, no. 1, pp. 122–135, 2019.

[17] X. Yu and J. Xu, "A bayesian recognition method of highway ambiguity path based on directed graph," *Highway and Transportation Science and Technology*, vol. 35, no. 11, pp. 110–115, 2018.

[18] H. Zou, R. Jia, and C. Zhang, "Research on freescale smart car track recognition and control strategy," *Machine Tool & Hydraulics*, vol. 46, no. 15, pp. 101–105, 2018.

[19] Y. Hou, X. Lv, Y. Chen, J. Zhao, Q. Li, and H. Chen, "Application of deep neural network in visual recognition of forest trails," *Journal of Computer Science and Exploration*, vol. 13, no. 2, pp. 263–274, 2019.

[20] Z. Qian, P. Li, and G. Zhou, "Uncertain and negative scope recognition based on two-way LSTM network," *Journal of Software*, vol. 29, no. 8, pp. 2427–2447, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 8520094, 8 pages
https://doi.org/10.1155/2021/8520094

Research Article

Intelligent Prediction Mathematical Model of Industrial Financial Fraud Based on Data Mining

Xiuqin Geng and Dawei Yang

Shandong Polytechnic, Jinan 250104, China

Correspondence should be addressed to Dawei Yang; 1103954838@qq.com

Received 13 June 2021; Revised 25 June 2021; Accepted 10 July 2021; Published 4 August 2021

Academic Editor: Gengxin Sun

The essence of enterprise financial modeling is to use mathematical models to classify and sort out all kinds of enterprise information according to the main line of value creation and on this basis to complete the analysis, prediction, and value evaluation of enterprise financial situation. A reasonable financial model is also an effective means to reduce financial fraud. In this paper, a financial fraud identification model is constructed based on empirical data. In the process of model construction, the primary feature set is selected according to the financial fraud motivation theory, and then, the original feature set is obtained by Mann–Whitney test on the primary feature set, and the final fraud identification feature set is selected from the original feature set by using Relief and Boruta algorithms. Finally, based on the final fraud identification feature set, the data algorithms such as decision tree, logistic regression, support vector machine, and random forest are used to identify financial fraud. The experimental results show that the combination of financial fraud identification features constructed by the Relief algorithm and random forest model has the best recognition effect. The evaluation indexes of the G mean value and the F value were 75.86% and 78.33%, respectively.

1. Introduction

In the process of enterprise production and operation, enterprise financial management has the function of making enterprise managers understand the operation status of enterprises in time, providing a decision-making basis for enterprise managers, and standardizing the rationality and legitimacy in the process of enterprise operation. With the expansion of the scale of enterprise financial management, we need to change the backward traditional manual way of financial management, through the establishment of financial data-related models and optimization methods to better grasp the business dynamics of enterprises.

Financial fraud [1, 2] is a kind of behavior where the management of a company deliberately manipulates financial information to conceal the true assets and liabilities, operating results, and cash flow of the company in order to achieve some improper purposes and then induces the users of financial statements to make wrong economic decisions based on false financial information. PricewaterhouseCoopers' research report shows that, in recent years, the number of enterprise financial fraud cases has increased year by year. In the past two years, about half of American organizations suffered from financial fraud. In the past few years, the losses caused by financial statement fraud and asset misappropriation in various regions of the world have increased year by year, with a total of about US $3.7 trillion worldwide, and the fraud will cause the company's revenue loss of nearly 5% in the current year.

As an independent third party, auditors are responsible for the reasonable assurance of whether there are material misstatements due to fraud or errors in the financial statements of enterprises. Therefore, improving the ability of auditors to identify financial fraud is of great significance to curb financial fraud and reduce the losses caused by financial fraud. In the digital information environment, data audit has appeared; the new audit mode takes the original data in the audited database as the audit object, establishes the audit

intermediate table through the collection, sorting, and analysis of the original data, and then constructs the model for data analysis by using data mining technologies such as classification, clustering, association analysis, and outlier detection.

In the era of big data, in the face of the explosive growth of data, data mining is more and more widely used by virtue of the ability to find favorable patterns and trends from data sets. Data mining is a process of discovering useful patterns and trends from large data sets. With the deepening of the research on data mining theory, data mining technology has become increasingly diverse. According to the purpose, data mining can be divided into three categories: classification, clustering, and association rule analysis. In audit research, the data mining technology which is often used by the majority of scholars is mainly a classification algorithm.

As an important tool of data processing and information mining, mathematical modeling and data mining are paid more and more attention by audit theory researchers. In the process of modeling, a lot of mathematical theories are involved, such as optimization theory, probability theory, and quantitative statistics. There are many very important mathematical models in accounting and financial management, such as the capital asset pricing model [3], portfolio model [4], securities valuation model [5], and Black-Scholes option pricing model [6]. Dai et al. [7] studied the companies that restate financial statements due to violation of generally accepted accounting principles and found that the sensitivity of option portfolio held by senior managers to stock price has a significant positive correlation with financial statements. Liu et al. [8] studied the correlation between ordinary employees' stock compensation and corporate financial statement fraud. Glancy et al. [1] used the data model to study a large number of financial fraud enterprises and found that the financial fraud enterprises have different degrees of external financing needs during the period of false reporting. By using the data mining method, Albrecht et al. [9] found that, in the period of corporate financial fraud, the number of insiders selling stocks and exercising stock appreciation rights is larger than in other periods. By constructing a mathematical model, Burnes et al. [10] found that the bonus plan is usually directly related to the management's income and has a lower limit; it makes the management easily affected by the game psychology. Khachatryan et al. [11] used an association rule algorithm to study the relationship between financial leverage and corporate fraud and found that financial fraud companies have higher financial leverage than nonfinancial fraud companies.

A large number of studies have found that there is a systematic relationship between financial characteristics and financial fraud. In some cases, financial characteristics are considered to reflect the occurrence mechanism of financial fraud. Louzada et al. [12] proposed a mathematical model to predict the false income. The model found that the percentage change of total assets is positively correlated with the false income, and the percentage change of the number of employees is negatively correlated with the false income. Based on the mathematical model, Tarjo et al. [2] found that when the growth of earnings per share slows down, it shows that they are facing the negative impact of financial performance, and there is a great possibility of financial fraud. Bose et al. [13] used the nonparametric Mann–Whitney test to test 23 financial indicators. The test results showed that 17 financial indicators had significant differences in the samples of judging the possibility of financial report fraud and nonfinancial report fraud. At the same time, the results of Mann–Whitney nonparametric test are further verified by the classification model of logistic regression and artificial neural network algorithm. Sun et al. [14] conducted a t-test on the identification characteristics of fraud risk factors and constructed a fraud identification model by using a support vector machine algorithm and logistic regression analysis.

It can be seen from the above analysis that the mathematical model and data mining technology can extract a large number of information from the financial information and nonfinancial information provided by customers' business activities, which cannot be obtained by the existing audit evidence collection methods. It is of great benefit to improve the audit efficiency and audit effect.

2. Feature Selection of Financial Fraud Identification

The company's fictitious profits, misappropriation of assets, and improper accounting treatment will directly affect the financial statements, which will lead to the abnormality of the company's statement items and various financial indicators calculated according to the statement items. Therefore, some statement items and financial indicators can become the identification attributes of enterprise financial fraud to a certain extent. For those identification attributes that are useful for financial fraud identification, we call them relevant features. Feature selection [15] is the process of selecting relevant feature subsets from the constructed fraud recognition feature set.

Feature selection is an important branch of machine learning. That is, a candidate subset is generated in the initial feature set, and the correlation is evaluated by using the evaluation function. Based on the evaluation results, the next candidate subset is generated, and then the evaluation function is used to evaluate it; the process is repeated until a better feature subset cannot be found.

As shown in Figure 1, feature selection usually takes three processes: firstly, the candidate feature subset is generated by subset search, then the subset goodness-of-fit is evaluated by the selected evaluation function, and finally, a threshold for the evaluation function is set. When the value of the evaluation function reaches the threshold, it can stop searching and output the optimal feature subset.

Subset search is the first key step of feature selection and the process of generating the best candidate subset. According to different subset search patterns, the search algorithm can be divided into a complete search, heuristic search, and random search [16].

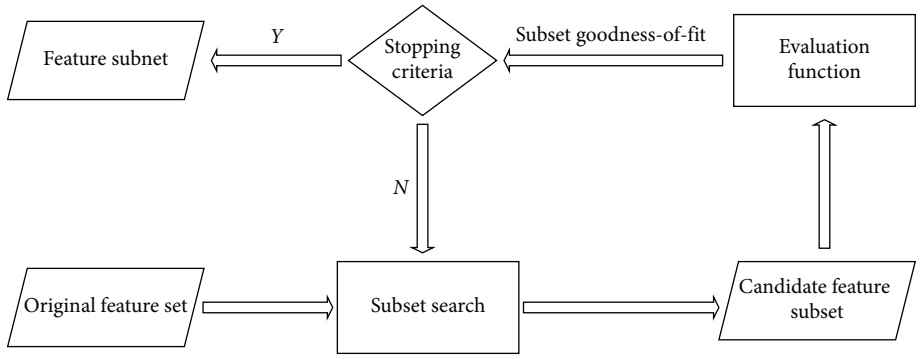


Figure 1: The basic process of feature selection.

The idea of a complete search is to traverse all possible feature subsets in the initial feature set to find feature subsets containing all important information. This method is suitable for the case that the number of features in the original feature set is small.

Heuristic search is divided into the forward search, backward search, bidirectional search, and case-based search. The case-based algorithm mainly includes the Relief algorithm [17]. The algorithm uses distance measure as an evaluation function. Firstly, the user sets a parameter and initializes the weight of each feature to 0, then, an instance is randomly selected from the training set, its near-hit neighbors and near-miss neighbors are calculated according to Euclidean distance, and the weight of each feature is updated by using the near-hit and near-miss neighbors; after all feature weights are calculated, all features whose weights are greater than a certain threshold are selected.

In heuristic search, the forward search algorithm and backward search algorithm often fall into the trap of local optimal value when searching the optimal candidate subset. The random algorithm generates a candidate subset randomly and then performs a forward search algorithm or backward search algorithm on this candidate subset. It makes up for the defect that the forward search algorithm and backward search algorithm cannot jump out of the local optimal value. The random algorithms commonly used in subset search are randomly generated sequence selection algorithm [18], simulated annealing algorithm [19], and genetic algorithm [20].

Subset evaluation is the second key step in feature selection. In this step, the candidate subset goodness-of-fit is evaluated according to the set evaluation function, and whether to stop the subset search process is decided according to the candidate subset goodness-of-fit. The essence of subset evaluation is to evaluate the difference between the current candidate subset's partition of the training data set and the real partition of the training data set. The smaller the difference is, the better the current candidate subset will be. When the difference between the partition of the new candidate subset and the real partition of the training data set reaches the minimum, the candidate subset is stopped. The current candidate subset is the feature subset that contains all the important information.

We introduce the whole idea of the feature selection algorithm from two aspects of subset search and subset evaluation. However, due to the differences in search methods and evaluation methods used in subset search and subset evaluation, feature selection methods can be roughly divided into filtering, packaging, and embedded.

A Relief algorithm is a filtering feature selection algorithm. Firstly, an example x_j is randomly selected from the training data set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$; the nearest neighbor $x_{j,nh}$ is found from the sample set similar to x_j; it is called a near-hit neighbor. Then, the nearest neighbor $x_{j,nm}$ is found from the sample set different from x_j; it is called near-miss neighbor. Finally, the correlation statistic component of the corresponding attribute u can be calculated as follows:

$$\vartheta^u = \sum_j -\operatorname{diff}\left(x_j^u, x_{j,nh}\right)^2 + \operatorname{diff}\left(x_j^u, x_{j,nm}\right)^2, \quad (1)$$

where x_j^u represents the value of sample x_j on feature attribute u.

The above calculation process is repeated for different samples, and then, the estimation results of different samples are averaged to get the relevant statistical components of different attributes u; the larger the component value is, the stronger the classification ability is. The smaller the component value is, the weaker the classification ability is. When the component value is greater than the set threshold, the feature attribute is selected to classify the target.

Boruta algorithm [21] is a package based on a random forest classification algorithm. Its idea is consistent with the basic idea of random forest classifier; that is, by increasing the randomness of the system and collecting the results from the random sample set, the misleading influence of random fluctuation and correlation can be reduced. Boruta algorithm mainly includes the following steps:

Step 1: copy all real features in the data set to create shadow features

Step 2: randomize the newly added shadow features to eliminate the correlation between them

Step 3: a random forest classifier is trained on the extended data set, and the calculated Z-score is collected

Step 4: find out the maximum Z-score value of the shadow feature, and then record a hit for each real feature whose Z-score is higher than the maximum Z-score value, which means that the real feature is more important than the shadow feature

Step 5: the same bilateral tests as maximum Z-score value are performed for real features whose importance cannot be determined

Step 6: attributes that are less important than shadow features are considered unimportant and are permanently removed from the original set

Step 7: the significant higher importance than shadow feature is regarded as important, and delete all shadow attributes

Step 8: repeat this process until the importance of all attributes is assigned, or the algorithm has reached the previously set limit of random forest operation

According to the existing relevant researches [22, 23], this paper selects the appropriate financial indicators from nine categories of profitability, operating capacity, development capacity, per share index, ratio structure, solvency, risk level, disclosure of financial indicators, and cash flow analysis as the features of financial fraud identification.

The feature selection of financial fraud identification is divided into three steps. First of all, the independent sample T-test and Mann–Whitney test are used to test nine categories of financial indicators. Through these two methods, we can screen out the indicators that can significantly distinguish financial fraud companies from nonfinancial fraud companies. However, the independent sample T-test method requires that the indicators obey the normal distribution. Kolmogorov–Smirnov normality test is carried out for the nine categories of indicators, assuming that all the financial indicators follow the normal distribution, and the test results show that the significant values are less than 0.05, which indicates that these indicators do not follow the normal distribution. Therefore, all indicators are tested by Mann–Whitney test, assuming that there is no significant difference between these nine categories of financial indicators and whether the companies are financial fraud. The test results show that the significant value of 75 indicators is less than 0.05, which shows that these 75 financial indicators have a significant role in identifying financial fraud samples. This paper selects these 75 financial indicators as the original feature set of financial fraud identification. Then, Boruta and Relief feature selection algorithms are used to further screen the primary features to reduce the feature dimension and improve the model adaptability of fraud identification features.

Boruta algorithm divides the primary features into three index sets: Confirmed, Rejected, and Tentative. All 18 indicators in Confirmed are chosen as the identification indicators of the fraud classification model. The result of feature selection is shown in Table 1.

Table 1: The result of feature selection based on the Boruta algorithm.

Feature	Feature name
F1	Interest coverage ratio
F2	Cash flow rate
F3	Asset-liability ratio
F4	Total assets net profit margin
F5	Return on invested capital
F6	Weighted average return on net assets after deducting loss
F7	Ratio of receivables to income
F8	Accounts receivable turnover
F9	Total assets turnover
F10	Growth rate of total profit
F11	Sustainable growth rate
F12	Growth rate of owner's equity
F13	Financial leverage
F14	Comprehensive leverage
F15	Basic earnings per share after deducting nonrecurring profit and loss
F16	Retained earnings per share
F17	Net cash flow from operating activities per share
F18	Comprehensive tax rate

These indicators mainly reflect the company's profitability, operation ability, development ability, debt-paying ability, and cash flow status. From the perspective of solvency, the feature set selected by the Boruta algorithm pays more attention to the solvency of business activities to corporate debt and the solvency of business achievements to interest. From the perspective of operational capacity, the feature set selected by the Boruta algorithm pays more attention to the turnover efficiency of enterprise accounts receivable. From the perspective of development capability, the feature set selected by the Boruta algorithm pays more attention to the internal growth power of enterprises in the future. From the perspective of risk level, the feature set screened by the Boruta algorithm considers that the financial risk and operational risk of an enterprise have an effect on the identification of financial fraud.

In the process of feature selection using the Relief algorithm, the selection threshold is set to 0, the features with weight greater than 0 in the original feature set are retained, and the features with weight less than 0 in the original feature set are discarded. Finally, 17 categories of fraud identification feature indicators are obtained. The result of feature selection is shown in Table 2.

These 17 indicators evaluate the financial situation of enterprises from seven dimensions: solvency, profitability, operation ability, development ability, risk level, per share index, and tax burden. Compared with the feature set screened by the Boruta algorithm, the feature set screened by the Relief algorithm thinks that the interest-paying ability of enterprises is lack of recognition degree to identify financial fraud. According to the Relief algorithm, the management of cost and expense, the loss of asset impairment, and the growth of sales expenses have a better recognition degree for the identification of financial fraud.

Table 2: The result of feature selection based on the Relief algorithm.

Feature	Feature name
F1	Cash flow rate
F2	Equity multiplier
F3	Return on equity
F4	Cost profit margin
F5	Asset impairment loss income ratio
F6	Ratio of accounts receivable to income
F7	Accounts receivable turnover
F8	Business cycle
F9	Total assets turnover
F10	Growth rate of return on equity
F11	Growth rate of total profit
F12	Growth rate of sales expenses
F13	Sustainable growth rate
F14	Comprehensive leverage
F15	Undistributed profit per share
F16	Net cash flow from operating activities per share
F17	Comprehensive tax rate

3. Construction of Industrial Financial Fraud Identification Model

The financial fraud identification model is a two-classification model based on a classification algorithm. The common evaluation indexes for the performance of the model are error rate and accuracy. Error rate refers to the proportion of samples with the wrong classification in the total number of samples, while accuracy refers to the proportion of samples with correct classification in the total number of samples. Although the error rate and accuracy are very common, their practicability is not high. In order to better judge the accuracy of the financial fraud identification model, this paper selects the confusion matrix to evaluate the performance of the model. The confusion matrix [24] is an important tool to evaluate the performance of the classification model. It can reflect the number of correct classification and wrong classification of each category in the sample. For the two-classification task of enterprise financial fraud identification, the combination of the real categories of sample enterprises and the prediction categories of fraud identification model can be divided into true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The confusion matrix of classification results is shown in Table 3.

In Table 3, 1 represents the fraud sample, and 0 represents the nonfraud sample.

According to the confusion matrix, some other indicators, shown in Table 4, are designed to evaluate the classification effect, including accuracy, sensitivity, and specificity.

In addition to the above indicators, F value and G mean value are often used. These two indexes give the comprehensive performance evaluation of the fraud identification model.

F value [25] is a comprehensive consideration of sensitivity and accuracy, and its calculation formula is defined as follows:

Table 3: Structure of confusion matrix.

		Prediction category	
		1	0
Actual category	1	TP	FN
	0	FP	TN

Table 4: Indicators of evaluation metrics based on confusion matrix.

Evaluation metrics	Formula
Accuracy	(TP + TN)/(TP + FP + TN + FN)
Sensitivity	TP/(TP + FN)
Precision	TP/(TP + FP)
Specificity	TN/(TN + FP)

$$F = \frac{\left(\rho^2 + 1\right) \times P \times S}{\rho^2 \times P + S}, \tag{2}$$

where S represents the sensitivity of the model, P represents the precision of the model, and ρ is the parameters for adjusting accuracy and sensitivity weights.

If accuracy and sensitivity are considered equally important, then $\rho = 1$. When evaluating the performance of the fraud model, the larger the F value is, the better the performance of the model is.

G mean value is a comprehensive measure of sensitivity and specificity, and it is also a comprehensive index used to evaluate the performance of the model. Its calculation formula is defined as follows:

$$G = \sqrt{S \times M}, \tag{3}$$

where M represents the specificity of the model.

When evaluating the performance of the fraud model, the larger the G value is, the better the performance of the model is.

Support vector machine, decision tree, logistic regression, and random forest were used to build financial fraud recognition models, and the recognition effects of different models were evaluated.

Based on the CSMAR database, this paper obtains 257 listed companies' consolidated financial statements with fictitious profits or assets from 2010 to 2019 as the fraud samples. At the same time, according to the selection principle of control samples, the corresponding number of control samples is selected according to the ratio of 1 : 1. According to the feature selection of fraud identification samples, this paper preprocesses the original samples and the feature samples filtered by the Boruta algorithm and Relief algorithm. 70% of the preprocessed data set is used as a training set and 30% as a test set.

In the fraud identification model experiment, 5-cross validation is used. Due to the instability of classification model, 10 running results are selected for each classification model. The mean value represents the running result of each model, shown in Table 5.

From the results of the above four fraud recognition models, the original feature set samples have good

Table 5: The mean running result of each model.

Classification model	Specificity (%)	Sensitivity (%)	Accuracy (%)	G (%)	F (%)
Decision tree	69.02	65.01	66.79	66.06	66.01
Logistic regression	68.97	69.98	70.05	68.96	69.12
Random forest	76.30	75.17	76.08	75.31	75.27
Support vector machine	72.18	78.92	76.37	74.96	76.16

recognition results in the support vector machine model. The values of G mean and F reach 74.96% and 76.16%, respectively. However, there are 75 fraud features in the original feature set, which is the main reason for the high efficiency of model recognition. Therefore, it is necessary to reduce the dimension of the original feature set to find the feature set with fewer fraud identification features and better model results.

Because of the large number of dimensions in the original feature set, it is not easy to extract and apply the fraud features. Therefore, the initial feature set selected by the Boruta algorithm is used as the feature set of fraud identification. Each classification model selects 10 running results, and the mean value represents the running result of each model, shown in Table 6.

From the results of the above four fraud recognition models, the feature set samples screened by the Boruta algorithm have good recognition results in the random forest model, and the values of G mean and F reach 74.26% and 74.31%, respectively. The number of fraud identification features in the feature set screened by the Boruta algorithm is reduced from 75 to 18, which reduces the dimension of the original fraud identification feature set. However, the identification effect of the feature set samples screened by the Boruta algorithm is not as good as that of the original fraud identification feature set.

The feature set screened by the Boruta algorithm cannot reduce the dimension of the fraud recognition feature and keep the good recognition efficiency of the fraud recognition model. Therefore, the feature set screened by the Relief algorithm is further used as the feature of fraud identification. Each classification model selects 10 running results; the mean value represents the running result of each model, shown in Table 7.

From the results of the above four fraud identification models, it can be seen that the feature set samples screened by the Relief algorithm have good identification results in the random forest model, and the G mean and F reach 75.86% and 78.33%, respectively. Compared with the original fraud identification feature set, the number of fraud identification features in the feature set screened by the Relief algorithm is reduced from 75 to 17, which reduces the dimension of the original fraud identification feature set; the overall recognition effect of the feature set samples screened by the Relief algorithm is better than that of the original fraud recognition feature set samples.

In conclusion, the random forest model has the best performance among the four fraud identification models. From the perspective of financial fraud identification features, the comprehensive identification performance of the fraud identification features selected by the Relief algorithm in the random forest model reaches 78.33%, which can best reflect the differences between fraudulent enterprises and nonfraudulent enterprises. The contribution of this paper is mainly reflected in two aspects: one is to combine the prior knowledge of fraud identification with feature selection algorithm to select the feature set of financial fraud identification based on the Relief algorithm and Boruta algorithm; the other is to verify the above two kinds of fraud identification features by building a financial fraud identification model. It is found that the set of financial fraud recognition features selected based on the Relief algorithm has the best recognition performance.

4. Discussion

It has rich theoretical and practical significance to study the application of data mining and mathematical model in financial fraud identification. On the one hand, it enriches the theoretical system of financial fraud audit; on the other hand, it provides new ideas and methods for financial fraud audit practice. In this paper, the prior knowledge and feature selection algorithm of financial fraud identification are used to study the characteristics of financial fraud identification, and the financial fraud identification model is established based on data mining technology.

Compared with the existing related research results, the research results of this paper are mainly reflected in two aspects. One is to combine the prior knowledge of fraud identification with a feature selection algorithm to select the feature set suitable for industrial financial fraud identification. The other is to verify two kinds of fraud identification features by constructing the financial fraud identification model.

Through this study, it is found that fraudulent enterprises have weak solvency, high debt risk, and strong willingness to finance, and the cash flow generated by operating activities is lower than other normal enterprises in the same industry. Financial indicators such as cash flow ratio, equity multiplier, and net cash flow per share from operating activities are good features for fraud identification. The assets of fraudulent enterprises are in poor condition and slow turnover, and their profitability and growth ability are lower than other normal enterprises in the same industry. The inventory turnover rate, accounts receivable turnover rate, return on net assets, growth rate of return on net assets, sustainable growth rate, and other financial indicators are good fraud identification characteristics. The growth rate of costs and expenses of fraudulent enterprises is higher than that of normal enterprises in the same industry, and the comprehensive tax burden of enterprises is also lower than

Table 6: The mean running result of each model based on the feature set selected by the Boruta algorithm.

Classification model	Specificity (%)	Sensitivity (%)	Accuracy (%)	G (%)	F (%)
Decision tree	66.32	65.72	66.58	66.25	66.17
Logistic regression	66.02	75.93	71.16	70.76	71.02
Random forest	71.27	76.31	74.38	74.26	74.31
Support vector machine	70.09	73.58	71.96	71.72	71.35

Table 7: The mean running result of each model based on the feature set selected by the Relief algorithm.

Classification model	Specificity (%)	Sensitivity (%)	Accuracy (%)	G (%)	F (%)
Decision tree	62.51	68.02	65.28	64.21	66.78
Logistic regression	80.06	60.99	70.76	69.88	68.92
Random forest	73.72	78.29	75.97	75.86	78.33
Support vector machine	85.07	58.62	63.26	56.37	67.06

that of other normal enterprises. The cost-profit rate, asset impairment loss rate, sales expense growth rate, comprehensive tax rate, and other financial indicators are good fraud identification characteristics.

5. Conclusions

This paper mainly studies the application of mathematical models and data mining technology in financial fraud identification. Based on the prior knowledge and feature selection algorithm of financial fraud identification, the financial fraud identification features are studied, and the financial fraud identification model is established based on logistic regression, decision tree, support vector machine, random forest, and other data mining technologies. This paper attempts to provide an effective analysis and prediction method for auditors to improve their ability to identify fraud risks. In future research, we will focus on the parameter adjustment of the model to further improve the recognition performance of the model.

Data Availability

The basic data used in this paper are downloaded from the online public data set: China Stock Market & Accounting Research Database https://www.gtarsc.com/

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by a grant from Shandong Scientific Research Project of China (no. J18RB132).

References

[1] F. H. Glancy and S. B. Yadav, "A computational model for financial reporting fraud detection," *Decision Support Systems*, vol. 50, no. 3, pp. 595–601, 2011.

[2] A. Tarjo and N. Herawati, "Application of beneish M-score models and data mining to detect financial fraud," *Procedia — Social and Behavioral Sciences*, vol. 211, pp. 924–930, 2015.

[3] V. Vovk and G. Shafer, "The game-theoretic capital asset pricing model," *International Journal of Approximate Reasoning*, vol. 49, no. 1, pp. 175–197, 2014.

[4] Z. Qin, S. Kar, and H. Zheng, "Uncertain portfolio adjusting model using semiabsolute deviation," *Soft Computing*, vol. 20, no. 2, pp. 1–9, 2016.

[5] L. A. Zhe and B. Aba, "Normalized nonconformity measures for automated valuation model," *Expert Systems with Applications*, vol. 180, no. 1, Article ID 115165, 2017.

[6] P. Roul and V. M. K. P. Goura, "A compact finite difference scheme for fractional black-scholes option pricing model," *Applied Numerical Mathematics*, vol. 166, pp. 40–60, 2021.

[7] L. Dai, Z. Fu, and Z. Huang, "Option pricing formulas for uncertain financial market based on the exponential ornstein-uhlenbeck model," *Journal of Intelligent Manufacturing*, vol. 28, no. 3, pp. 597–604, 2017.

[8] Y. Liu and D. S. Wang, "Symmetry analysis of the option pricing model with dividend yield from financial markets," *Applied Mathematics Letters*, vol. 24, no. 4, pp. 481–486, 2011.

[9] C. Albrecht, D. Holland, R. Malagueño, S. Dolan, and S. Tzafrir, "The role of power in financial statement fraud schemes," *Journal of Business Ethics*, vol. 131, no. 4, pp. 803–813, 2015.

[10] D. Burnes, C. R. Henderson, C. Sheppard, R. Zhao, K. Pillemer, and M. S. Lachs, "Prevalence of financial fraud and scams among older adults in the United States: a systematic review and meta-analysis," *American Journal of Public Health*, vol. 107, no. 8, p. 1295, 2017.

[11] D. Khachatryan and B. Muehlmann, "Determinants of successful patent applications to combat financial fraud," *Scientometrics*, vol. 111, no. 3, pp. 1–31, 2017.

[12] F. Louzada and A. Ara, "Bagging k-dependence probabilistic networks: an alternative powerful fraud detection tool," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11583–11592, 2012.

[13] I. Bose, S. Piramuthu, and M. J. Shaw, "Quantitative methods for detection of financial fraud," *Decision Support Systems*, vol. 50, no. 3, pp. 557-558, 2011.

[14] N. Sanaz and S. Mehdi, "Cost-sensitive payment card fraud detection based on dynamic random forest and k-nearest neighbors," *Expert Systems with Applications*, vol. 110, no. 11, pp. 381–392, 2018.

[15] G. Sun and S. Bin, "Router-level internet topology evolution model based on multi-subnet composited complex network model," *Journal of Internet Technology*, vol. 18, no. 6, pp. 1275–1283, 2017.

[16] S. Bin and G. Sun, "Optimal energy resources allocation method of wireless sensor networks for intelligent railway systems," *Sensors*, vol. 20, no. 2, p. 482, 2020.

[17] N. Amjady, A. Daraeepour, and F. Keynia, "Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network," *IET Generation, Transmission & Distribution*, vol. 4, no. 3, pp. 432–444, 2010.

[18] S. Bin, G. Sun, N. Cao et al., "Collaborative filtering recommendation algorithm based on multi-relationship social network," *Computers, Materials & Continua*, vol. 60, no. 2, pp. 659–674, 2019.

[19] C. Takeang and A. Aurasopon, "Multiple of hybrid lambda iteration and simulated annealing algorithm to solve economic dispatch problem with ramp rate limit and prohibited operating zones," *Journal of Electrical Engineering & Technology*, vol. 14, no. 1, pp. 111–120, 2019.

[20] G. Sun, C. C. Chen, and S. Bin, "Study of cascading failure in multisubnet composite complex networks," *Symmetry*, vol. 13, no. 3, p. 523, 2021.

[21] M. B. Kursa and W. R. Rudnicki, "Feature selection with boruta package," *Journal of Statistical Software*, vol. 36, no. 11, pp. 1–13, 2010.

[22] T. J. Mock and J. L. Turner, "Auditor identification of fraud risk factors and their impact on audit programs," *International Journal of Auditing*, vol. 9, no. 1, pp. 59–77, 2005.

[23] J. Tang, K. E. Karim, and B. Cooper, "Financial fraud detection and big data analytics — implications on auditors' use of fraud brainstorming session," *Managerial Auditing Journal*, vol. 34, no. 3, pp. 324–337, 2019.

[24] D. Simon and D. L. Simon, "Analytic confusion matrix bounds for fault detection and isolation using a sum-of-squared-residuals approach," *IEEE Transactions on Reliability*, vol. 59, no. 2, pp. 287–296, 2010.

[25] G. Tian, S. Zhou, G. Sun, and C. C. Chen, "A novel intelligent recommendation algorithm based on mass diffusion," *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 4568171, 9 pages, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 3138301, 11 pages
https://doi.org/10.1155/2021/3138301

Research Article

The Numerical Investigation of the Heat Transport in the Nanofluids under the Impacts of Magnetic Field: Applications in Industrial Zone

Adnan[1], **Umar Khan**[2], **Naveed Ahmed**[3], **Syed Tauseef Mohyud-Din**[4], **Ilyas Khan**[5] and **Md. Fayz-Al-Asad**[6]

[1]*Department of Mathematics, Mohi-ud-Din Islamic University, Nerian Sharif AJ&K 12080, Trarkhel, Pakistan*
[2]*Department of Mathematics and Statistics, Hazara University, Mansehra 21120, Pakistan*
[3]*Department of Mathematics, Faculty of Sciences, HITEC University, Taxila Cantt 47070, Taxila, Pakistan*
[4]*University of Multan, Multan 66000, Pakistan*
[5]*Department of Mathematics, College of Science Al-Zulfi, Majmaah University, Al-Majmaah 11952, Saudi Arabia*
[6]*Department of Mathematics, Bangladesh University of Engineering and Technology (BUET), Dhaka 1000, Bangladesh*

Correspondence should be addressed to Md. Fayz-Al-Asad; fayzmath.buet@gmail.com

Received 25 May 2021; Accepted 18 July 2021; Published 28 July 2021

Academic Editor: Gengxin Sun

The dynamics of the nanofluid flow between two plates that are placed parallel to each other is of huge interest due to its numerous applications in different industries. Keeping in view the significance of such flow, investigation of the heat transfer in the Cu-H_2O nanofluid is conducted between parallel rotating plates. For more significant results of the study, the squeezing effects are incorporated over the plates that are electrically conducting. The nondimensional flow model is then treated analytically (VPM), and the results are sketched against the preeminent flow parameters. The remarkable heat transfer in the nanofluid is noticed against the Eckert and Prandtl numbers, whereas the Lorentz forces oppose the fluid temperature. Furthermore, the shear stresses at the walls drop and the local heat transfer rate rises due to increasing flow parameters. Finally, to validate the study, a comparison is made with existing available science literature and noted that the presented results are aligned with them.

1. Introduction

The heat transport investigation in the squeezed flow is substantial for engineering and industrial view point. These are in cooling, fog formation, lubrication system, food processing, and hydrodynamical machines etc. In view of significance of the squeezed flow, researchers and scientists focused to analyze the flow behavior and thermal performance under multiple flow conditions. Firstly, the squeezed flow of lubricants was introduced in [1].

Mustafa et al. [2] pointed out the flow behavior in fluids squeezed between two plates. The influence of multiple preeminent flow parameters on the fluid temperature and mass transport is decorated via graphs and explained comprehensively. Moreover, they found the local thermal performance against ingrained flow parameters. The analysis of magnetized flow regimes by considering the suction/blowing characteristics is imperative. It strengthens roots in civil engineering and industries as well. The analytical investigation of the fluid squeezed between parallel disks with suction/blowing characteristics was perceived in [3]. Further, they analyzed the flow characteristics under varying multiple parameters. The flow of H_2O composed by the Cu nanomaterial was presented by Khan et al. [4]. To enhance the thermal performance in the fluid, they adopted thermal conductance correlation based on multiple geometries of the nanomaterial. They pointed out that thermal performance is prompt for the platelet nanomaterial-based nanofluid.

The thermal transport analysis in the colloidal mixture between the opening/narrowing channel was detected in [5].

The walls of the channel are flexible and with source/sink situated at the culmination of the wall. For novelty of the colloidal model, they ingrained the impact of Lorentz forces in the energy equation. The enhanced thermal conductance of the fluid by dispersing nanomaterial, nanotubes, and various thermal conductance correlations was perceived in [6–8]. In 2014, Goktepe et al. [9] performed comparative inspection between single- and two-phase nanofluid models with convection characteristics. The thermal transport inspection in the unsteady colloidal mixture between parallel plates was perceived in [10]. They conducted the analysis between magnetized plates and decorated the results for the flow regimes. The most relevant investigation in the colloidal mixtures composed by various types of nanomaterials suspended in different host liquids was presented in [11–14].

Dogonchi et al. [15] inspected the temperature behavior in the magnetized nanofluid with radiative heat flux effects. They found that the temperature of the nanofluid enhanced against stronger thermal radiative effects. Further, the enhanced local thermal performance rate was reported in the work for the higher thermal radiation parameter. An analytical treatment of the nanofluid bounded by parallel plates was carried out in [16]. They used the Brownian motion effects, and to enhance the thermal performance in the used colloidal mixture, KKL thermal conductance correlation was ingrained in the energy equation. They pointed out that the local heat transport rate rises against higher volume fraction of the tiny material. Sheikholeslami et al. [17] perceived Brownian motion and thermophoretic effects on the flow characteristics. They revealed that the shear stresses at the plate surface enhanced for higher Hartmann and viscosity parameters. Also, they validated the analysis via comparative analysis.

The thermal and mass transport analysis in the colloidal mixture between rotating plates was perceived in [18]. Further, they revealed that the heat transport is in direct proportion for thermophoresis and Brownian motion parameters, while reverse behavior is pointed out for the concentration field. Singh et al. [19] pointed out the heat and mass transportation in the squeezed flow by imposing the slip effects on the plates. Also, they incorporated the influence of Lorentz forces and explored the significant results. They revealed that the heat and mass transport at the plate surface reduced against high Lorentz forces and volume fraction. Further, increment in the mass transport was detected against Schmidt and squeeze numbers. A novel thermal transport in GO-Molybdenum-disulfide/H_2O-$C_2H_6O_2$ was reported in [20]. They conducted the analysis between parallel rotating plates and found the significant results for the thermal performance of the hybrid nanoliquid. To improve the thermal performance rate, they ingrained the Cattaneo–Christov model in the energy relation and perceived significant changes in the heat transfer rate. Further, significant heat transfer investigation in the nanofluids was detected in [21, 22].

In 2019, Shah et al. [23] performed the analysis of heat and mass transport in rotating geometry. For novel analysis, they ingrained the radiative heat flux and Hall current in the energy and momentum constitutive relations. They decorated the results against the flow parameters and explained comprehensively. Recently, Khan et al. [24] conducted the thermal transportation over a sensor surface for γ-nanofluids. For thermal enhancement of the fluid, they adopted the effective Prandtl model and revealed significant changes in the thermal and momentum transport. Recently, Gul et al. [25] revealed the nanofluid thermal transport against the upright channel through the permeable medium. They highlighted the flow behavior for different parameters and described comprehensively.

The thermal transport analysis between parallel plates is significant. From the literature visit, it is revealed that the thermal performance in the nanofluids composed by multiple nanomaterials geometries is not conducted so far. Such thermal analysis is imperative for industrial and engineering view point. Therefore, the magnetized flow of H_2O composed by multiple nanomaterials (platelets, blades, and cylinders) is organized. For the mathematical study, variation of parameters method (VPM) is adopted, and it successfully tackled the colloidal model. The results for thermal performance, shear stresses, and local heat transfer rate against multiple flow parameters are decorated and explained comprehensively. A comparative investigation is also conducted, which proved the reliability of the adopted technique. Finally, core results of the analysis are ingrained in the conclusion section.

2. Materials and Methods

2.1. Model Formulation

2.1.1. Model Description and Geometry. The unsteady and electrically conducting fluid is taken between two parallel rotating plates in the Cartesian coordinate frame. A thermal conductance correlation based on multiple nanomaterial geometries is taken to improve the thermal performance of the model. The plates are apart from $z = \pm l[1 - \alpha t]^{1/2} = \pm h(t)$, where $\alpha > 0$ and $t < 1/\alpha$. Further, dissipation effects are plugged in the energy relation. The following restrictions are imposed on the colloidal flow model.

The flow is incompressible.

The volumetric fraction of the nonmaterial and the regular substance are thermally compatible.

There are no slip effects between them.

There is no chemical reaction occurring.

Figure 1 depicts the flow configuration of the Cu-H_2O nanofluid. In the view of aforementioned restrictions, the colloidal model which governs the flow of Cu-H_2O for multiple geometries of the nanomaterial takes the following form:

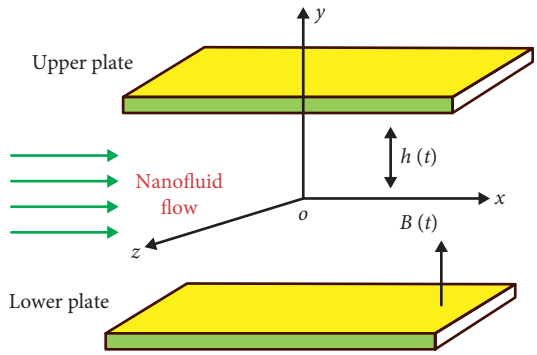


Figure 1: Geometry of the flow.

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0, \tag{1}$$

$$\rho_{nf}\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right)=-\frac{\partial p}{\partial x}+\mu_{nf}\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)-\frac{\sigma_{nf}B_0^2}{\rho_{nf}}u, \tag{2}$$

$$\rho_{nf}\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}\right)=-\frac{\partial p}{\partial y}+\mu_{nf}\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right), \tag{3}$$

$$\left(\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}\right)=\frac{k_{nf}}{\left(\rho C_p\right)_{nf}}\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}\right)+\frac{\mu_{nf}}{\left(\rho C_p\right)_{nf}}\left(4\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2\right). \tag{4}$$

And, supporting boundary conditions are

$$v=\frac{\partial u}{\partial y}=\frac{\partial T}{\partial y}=0 \text{ at } y=0, \tag{5}$$

$$v=v_w=\frac{\mathrm{d}h}{\mathrm{d}t}, T=T_H \text{ at } y=h(t). \tag{6}$$

Equation (1) is the conservation of mass, and the velocities along the coordinate axes are represented by u and v, respectively, p is the pressure, and T is the temperature. To enhance the flow characteristics, following correlations are adopted [26, 27]:

$$\rho_{nf}=(1-\phi)\rho_f+\phi\rho_s, \tag{7}$$

$$\mu_{nf}=\frac{\mu_f}{(1-\phi)^{2.5}}, \tag{8}$$

$$\left(\rho C_p\right)_{nf}=(1-\phi)\left(\rho C_p\right)_f+\phi\left(\rho C_p\right)_s, \tag{9}$$

$$\sigma_{nf}=\sigma_f\left(1+\frac{3\left(\sigma_s/\sigma_f-1\right)\phi}{\left(\sigma_s/\sigma_f+2\right)-\left(\sigma_s/\sigma_f-1\right)\phi}\right). \tag{10}$$

For thermal improvement, the following model is used [16]:

$$k_{nf}=k_f\left(\frac{k_s+(m-1)k_f-(m-1)\left(k_f-k_s\right)\phi}{k_s+(m-1)k_f+\left(k_f-k_s\right)\phi}\right), \tag{11}$$

where conductivity of the tiny particles, regular liquid, and shape factor of the nanoparticles are denoted by k_s, k_f, and m, respectively. The shape factor of the tiny particles is computed by $3/\psi$.

The appropriated dimensionless transforms are described by the following formulas in which η is similarity variables in which the fluid is squeezed for $\alpha>0$ and $t<1/\alpha$. The flow is squeezed between the plates until the time reaches the limit $1/\alpha$. Further, u, v, and T describe the velocities along horizontal, vertical, and the temperature, respectively:

$$\eta=\frac{y}{l(1-\alpha t)^{1/2}}, u=\frac{\alpha x}{2(1-\alpha t)}F'(\eta), v=-\frac{\alpha l}{2(1-\alpha t)^{1/2}}F(\eta), T=T_H\theta(\eta). \tag{12}$$

After using these variables in equations (1)–(4) and plugging the partial differentiations and self-similar variables, the following nanofluid model is attained:

$$F''''(\eta) - \mathrm{SK}_1(1-\phi)^{2.5}\left(\eta F''''^{(\eta)} + F'(\eta)F''^{(\eta)} - F(\eta)F'''(\eta) + 3F''^{(\eta)}\right) - \frac{\mathrm{MA}_1}{(1-\phi)^{-2.5}}F''(\eta) = 0, \tag{13}$$

$$\theta''(\eta) + \frac{\mathrm{PrSK}_2}{K_3}\left(\theta'(\eta)F(\eta) - \eta\theta'(\eta)\right) + \frac{\mathrm{PrEc}}{K_3(1-\phi)^{2.5}}\left(F''^2(\eta) + 4\delta^2 F'^2(\eta)\right) = 0. \tag{14}$$

The reduced dimensionless conditions at the plates are labeled in equations (15) and (16), respectively. It is noteworthy to mention that $\eta = 0$ and $\eta = 1$ denote the conditions at the lower and upper plates, respectively:

$$F(0) = 0, F''(0) = 0, F(1) = 1, F'(1) = 0, \tag{15}$$

$$\theta'(0) = 0, \theta(1) = 1. \tag{16}$$

In equations (13) and (14), squeeze, Prandtl, Eckert, and Hartmann numbers are denoted by $S = \alpha l^2/2\nu_f$, $\mathrm{Pr} = \mu_f(\rho C_p)_f/\rho_f k_f$, $\mathrm{Ec} = \rho_f/(\rho C_p)_f(\alpha x/2(1-\alpha t))^2$, $M = lB_0(\sqrt{\sigma_f(1-\alpha t)}/\mu_f)$, and $\delta = l/x$. The quantities represented by K_1, K_2, K_3, and A_1 are equal to

$$K_1 = (1-\phi) + \frac{\phi\rho_s}{\rho_f}, \tag{17}$$

$$K_2 = (1-\phi) + \frac{\phi\left(\rho C_p\right)_s}{\left(\rho C_p\right)_f}, \tag{18}$$

$$K_3 = \left(\frac{k_s + (m-1)k_f - (m-1)\left(k_f - k_s\right)\phi}{k_s + (m-1)k_f + \left(k_f - k_s\right)\phi}\right), \tag{19}$$

$$A_1 = \left(\frac{\left(\sigma_s + 2\sigma_f\right) + 2\left(\sigma_s - \sigma_f\right)\phi}{\left(\sigma_s + 2\sigma_f\right) - \left(\sigma_s - \sigma_f\right)\phi}\right). \tag{20}$$

The dimensional formulas for the shear stresses and Nusselt number are $C_F = \mu_{nf}(\partial u/\partial y)_{y=h(t)}/\rho_{nf}v_w^2$ and $\mathrm{Nu} = -lk_{nf}(\partial T/\partial y)_{y=}h(t)/\mathrm{KT}_H$.

After some calculations, these expressions become

$$C_F = K_1(1-\phi)^{2.5}F''(1), \tag{21}$$

$$\mathrm{Nu} = -K_3\theta'(1). \tag{22}$$

2.2. Mathematical Analysis. As under consideration, the nanofluid model is highly nonlinear in nature, and a closed form of the solution is very tedious or do not exist in general for such models. Therefore, we then move to tackle the model approximately. For the said purpose, the method known as variation of parameters method [28] is used. Primarily, the method is based on the Lagrange multiplier which can be calculated according to the model. After that, a recursive relation is written for the velocity and temperature equations. The initial conditions are used to calculate the initial guess, and the boundary conditions are used to compute the remaining constants appearing in the recursive relation. The detailed implementation of the method is given below.

2.2.1. Working Rules for VPM

Step 1: the initial step in VPM [28] is to reduce the nonlinear model in the following form:

$$\widetilde{\mathscr{L}}\,\widetilde{\mathscr{F}}(\eta) + \widetilde{\mathfrak{R}}\,\widetilde{\mathscr{F}}(\eta) + \widetilde{\mathfrak{N}}\,\widetilde{\mathscr{F}}(\eta) + \widetilde{g}^*(\eta) = 0, \tag{23}$$

where the highest order linear operator is $\widetilde{\mathscr{L}}$, $\widetilde{\mathfrak{R}}$ is also the linear operator with order less than $\widetilde{\mathscr{L}}$, $\widetilde{\mathscr{L}}$ presents the nonlinear operator, and $\widetilde{g}^*$ is the inhomogeneous part of the model.

Step 2: it is the calculation of the Lagrange Multiplier. For VPM, the Lagrange multiplier is calculated by adopting the following formula:

$$\lambda(s,\eta) = \frac{(-1)^{n*}(\eta - s)^{n^*-1}}{(n^*-1)!}. \tag{24}$$

In equation (24), n^* signifies the highest order derivative in the model.

Step 3: it is the calculation of initial approximation for the model. The initial approximation is calculated by means of initial conditions by adopting the following formula:

$$\mathscr{F}_0(\eta) = \sum_{i=0}^{k}\frac{\eta^i\mathscr{F}^i(0)}{i!}. \tag{25}$$

Step 4: it is the construction of the iterative scheme for VPM. The scheme is constructed by the following way:

$$\mathscr{F}_{n+1}(\eta) = \mathscr{F}_0(\eta) + \int_0^{\eta} \lambda(s,\eta)\left(-\Re\mathscr{F}(s) - \aleph\mathscr{F}(s) - g^*(s)\right)\mathrm{d}s, \quad n \geq 0. \tag{26}$$

By plugging the initial approximation and Lagrange multiplier, higher order approximations can be computed by running the recursive iterated scheme given in equation (26).

2.3. Implementation of the Proposed Technique. In VPM, equations (13) and (14) take the following form of recurrence relation:

$$F_{n+1}(\eta) = F(0) + F'(0)\eta + F''(0)\frac{\eta^2}{2!} + F'''(0)\frac{\eta^3}{3!} - \int_0^{\eta}\frac{(\eta-\check{s})^3}{3!}\left[-\mathrm{SK}_1(1-\phi)^{2.5}\left(\check{s}F'''(\check{s}) + 3F''(\check{s}) + F'(\check{s})F''(\check{s}) - F(\check{s})F'''(\check{s})\right)\right.$$
$$\left. - \mathrm{MA}_1(1-\phi)^{2.5}F''(\check{s})\right]\mathrm{d}\check{s}, \tag{27}$$

$$\theta_{n+1}(\eta) = \theta(0) + \eta\theta'(0) - \int_0^{\eta}(\eta-\check{s})\left[\frac{\mathrm{PrSK}_2}{K_3}\left(F(\check{s})\theta'(\check{s}) - \check{s}\theta'(\check{s})\right) + \frac{\mathrm{PrEc}}{K_3(1-\phi)^{2.5}}\left(F''^2(\check{s}) + 4\delta^2F'^2(\check{s})\right)\right]\mathrm{d}\check{s}. \tag{28}$$

After plugging the conditions in equations (27) and (28), the following relations are obtained:

$$F_{n+1}(\eta) = \breve{C}_1\eta + \frac{\breve{C}_2\eta^3}{6} - \int_0^{\eta}\frac{(\eta-\check{s})^3}{3!}\left[-\mathrm{SK}_1(1-\phi)^{2.5}\left(\check{s}F(\check{s}) + 3F(\check{s}) + F'(\check{s})F(\check{s}) - F(\check{s})F'''(\check{s})\right) - \mathrm{MA}_1(1-\phi)^{2.5}F''(\check{s})\right]\mathrm{d}\check{s}, \tag{29}$$

$$\theta_{n+1}(\eta) = C_3 + C_4\eta - \int_0^{\eta}(\eta-\check{s})\left[\frac{\mathrm{PrSK}_2}{K_3}\left(F(\check{s})\theta'(\check{s}) - \check{s}\theta'(\check{s})\right) + \frac{\mathrm{PrEc}}{K_3(1-\phi)^{2.5}}\left(F'^2(\check{s}) + 4\delta^2F'^2(\check{s})\right)\right]\mathrm{d}\check{s}, \tag{30}$$

where $\breve{C}_i$ for $i = 1, \ldots, 4$ are constants and computed by plugging the remaining flow conditions. Table 1 presents the comparative computation between VPM solutions and numerical computations.

3. Results and Discussion

3.1. The Velocity Field. The analysis of the imposed magnetic field on the flow of nanofluids is significant as it is used in many industrial products to purify the products. Therefore, the nanofluid motion against the Hartmann number is presented in Figures 2(a) and 2(b) for $F(\eta)$ and $F'(\eta)$, respectively. It is noticed that the fluid velocity $F(\eta)$ reduces by strengthening the magnetic field. Physically, the applied magnetic field opposes the fluid movement due to which the fluid motion declines. On the contrary, dual behavior of the fluid motion is observed for the velocity component $F'(\eta)$. Further, from the plotted results, it is obvious that the fluid motion declines and rises almost inconsequentially against the imposed magnetic field.

Figures 2 and 3 elaborate the fluid velocities $F(\eta)$ and $F'(\eta)$ for increasing values of volumetric fraction of Cu nanoparticles and squeezing parameter. From Figure 3, it is observed that the velocity components alter almost inconsequentially due to varying ϕ. Physically, by increasing the volume fraction of Cu nanoparticles, the resultant colloidal fluid becomes more denser which leads to slow movement of the fluid. The effects of the squeezing parameter on the velocities are demonstrated in Figure 4. It is noticed that the variations in the fluid motion are not prominent due to increasing squeezing effects. The values of thermophysical parameters are described in Table 2.

3.2. Temperature Field. The behavior of fluid temperature against multiple values of ϕ and Pr is plotted in Figure 5. These results are plotted for different nanoparticle shapes, namely, platelets, cylinder, and bricks. From the critical analysis of the plotted results, the temperature declines against higher ϕ. The physical reason behind the decrement in the temperature is that due to higher volume fraction, the colloidal fluid becomes more denser, and the collision between the fluid particles reduces which leads to decrement in the temperature. The results for temperature against higher Prandtl number values are shown in Figure 5(b) over the domain of interest. The temperature field $\beta(\eta)$ significantly rises for multiple values of Prandtl number. At the lower end ($\eta = 0$), the temperature significantly increases. The

Table 1: The results for $F(\eta)$ and $\theta(\eta)$.

$\eta \downarrow$	$F(\eta)$			$\theta(\eta)$		
	VPM	Numerical	Error	VPM	Numerical	Error
0.0	0.000	0.000	0.000	1.031	1.031	1.749×10^{-7}
0.1	0.140	0.140	4.178×10^{-9}	1.031	1.031	1.749×10^{-7}
0.2	0.279	0.279	3.535×10^{-9}	1.031	1.031	1.775×10^{-7}
0.3	0.413	0.413	1.804×10^{-9}	1.030	1.030	1.761×10^{-7}
0.4	0.541	0.541	2.077×10^{-9}	1.030	1.030	1.766×10^{-7}
0.5	0.661	0.661	2.763×10^{-9}	1.029	1.029	1.771×10^{-7}
0.6	0.769	0.769	2.745×10^{-9}	1.028	1.028	1.761×10^{-7}
0.7	0.861	0.861	2.208×10^{-9}	1.025	1.025	1.703×10^{-7}
0.8	0.934	0.934	1.634×10^{-9}	1.020	1.020	1.495×10^{-7}
0.9	0.982	0.982	8.843×10^{-9}	1.013	1.013	9.747×10^{-7}
1.0	1.000	1.000	0.000	0.999	1.000	3.000×10^{-10}

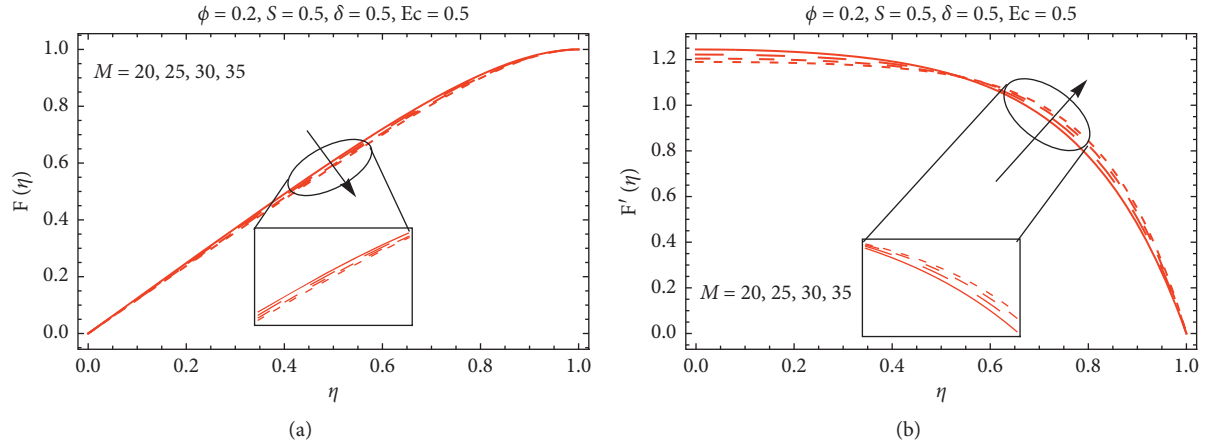


Figure 2: The velocity against M. (a) $F(\eta)$. (b) $F'(\eta)$.

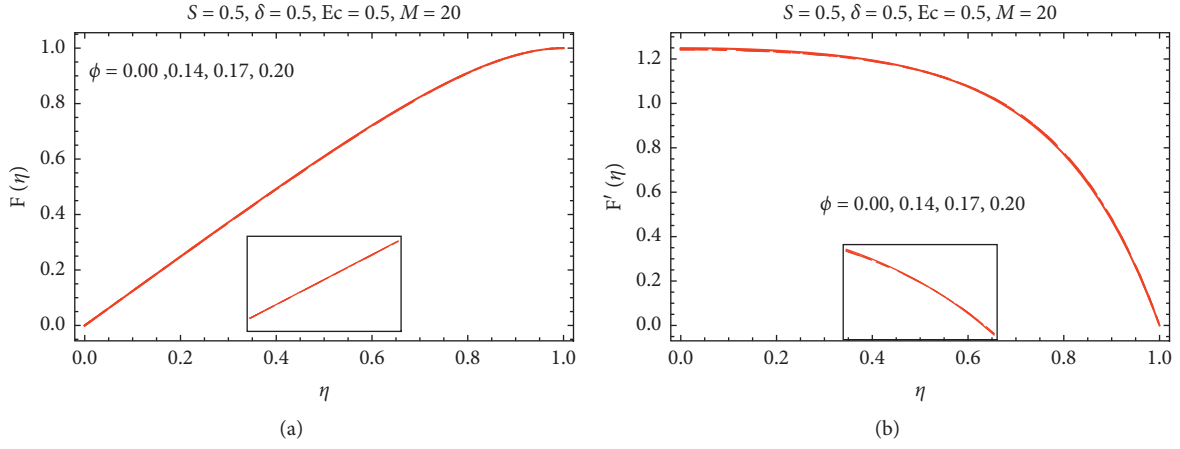


Figure 3: The velocity against ϕ. (a) $F(\eta)$. (b) $F'(\eta)$.

increment in the temperature of nanofluids gradually slows down from lower to upper end of the channel. Further, the increment in $\beta(\eta)$ prevailed for the brick-shaped nanomaterial in comparison with platelets and cylinder-shaped nanomaterial, respectively.

The temperature of Cu-H_2O for increasing S and M is plotted in Figures 6(a) and 6(b), respectively. It is worthy to mention that the plates move apart for $S > 0$. The fluid temperature declines for higher values of S. Physically, when the upper plate moves apart from the lower once, then the flowing area increases, which reduces the collisions between the fluid particles, and consequently, the temperature $\beta(\eta)$ drops. The temperature in Cu-H_2O containing platelet-shaped Cu declines abruptly than those cylinder- and brick-shaped nanofluids. This behavior of the temperature is due to the thermophysical parameters of the Cu nanoparticles.

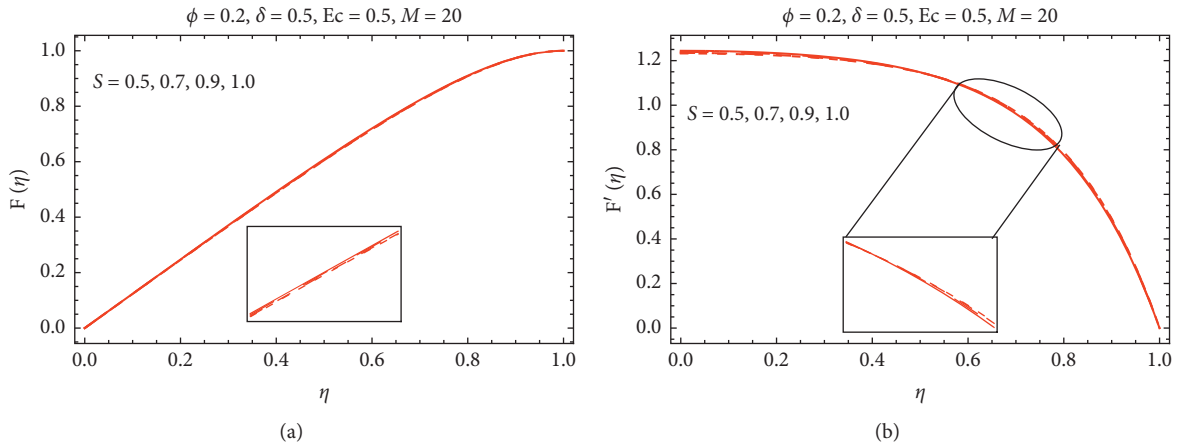


Figure 4: The velocity against S. (a) $F(\eta)$. (b) $F'(\eta)$.

Table 2: Attributes of tiny particles and regular liquid [26].

	ρ(kg/m^3)	C_p (J/kgK)	k(W/mk)	σ(Ωm)
Pure water	997.1	4179	0.613	0.05
Copper (Cu)	8933	385	401	596×10^7

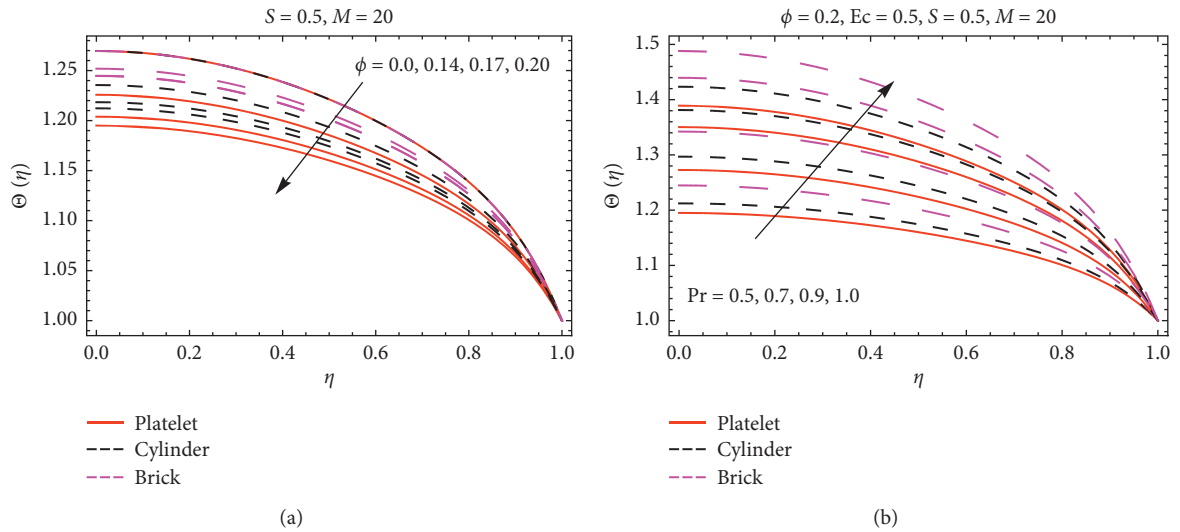


Figure 5: Thermal behavior $\theta(\eta)$ against (a) ϕ and (b) Pr.

The temperature field $\beta(\eta)$ for the applied magnetic field is presented in Figure 6(b). The motion of the nanofluids reduces due to the imposed magnetic field; as a result, the fluid temperature reduces. Physically, due to perpendicularly imposed magnetic field strength, collisions between the particles reduce; consequently, the temperature drops. In the vicinity of the plates, the temperature declines very rapidly due to stronger magnetic field effects there.

The effects of viscous dissipation on $\beta(\eta)$ are demonstrated in Figure 7(a) over the desired domain. It is examined that the temperature significantly rises due to stronger dissipation effects. The physical reason is that the internal energy of the fluid particles rises due to dissipation; as a result, the temperature rises significantly. For the nanofluid composed by brick-shaped Cu, more rapid increment in the temperature is examined. The effects are prominent near the lower plate. The temperature $\beta(\eta)$ variations against δ are plotted in Figure 7(b). It is noticed that the temperature promptly rises by increasing δ. The dominating behavior of the temperature is observed for the nanofluid containing brick-shaped Cu nanoparticles.

The impact of ϕ, magnetic parameter, and S on the shear stresses is plotted in Figures 8and 9, respectively. From the analysis, it is examined that the shear stresses drop against high volume fraction and stronger magnetic field effects. It is also observed that when the upper plate accelerates apart from the lower once, then the shear stresses decline.

Figure 10 presents the local heat transfer rate for multiple ϕ values and Eckert number. From Figure 10(a), it is noticeable that the local heat transfer rate rises against high

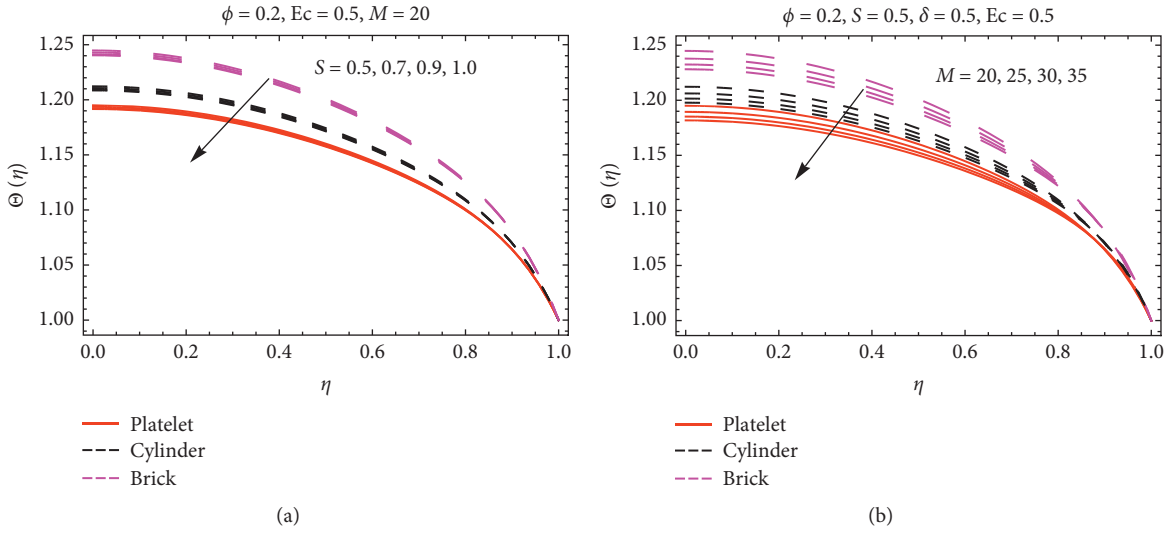


Figure 6: Thermal behavior $\theta(\eta)$ against (a) S and (b) M.

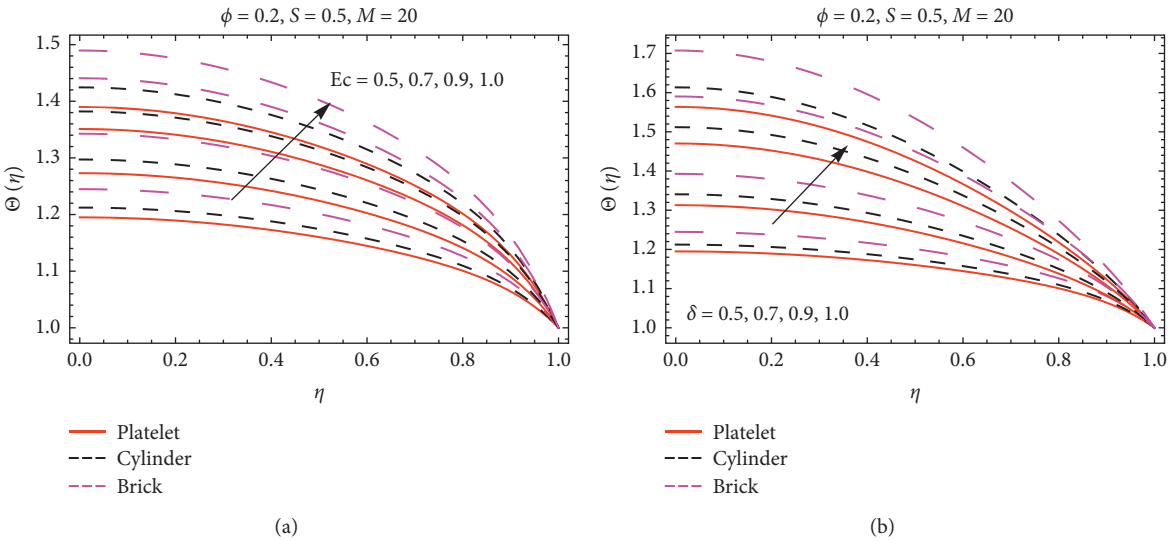


Figure 7: Thermal behavior $\theta(\eta)$ against (a) Ec and (b) δ.

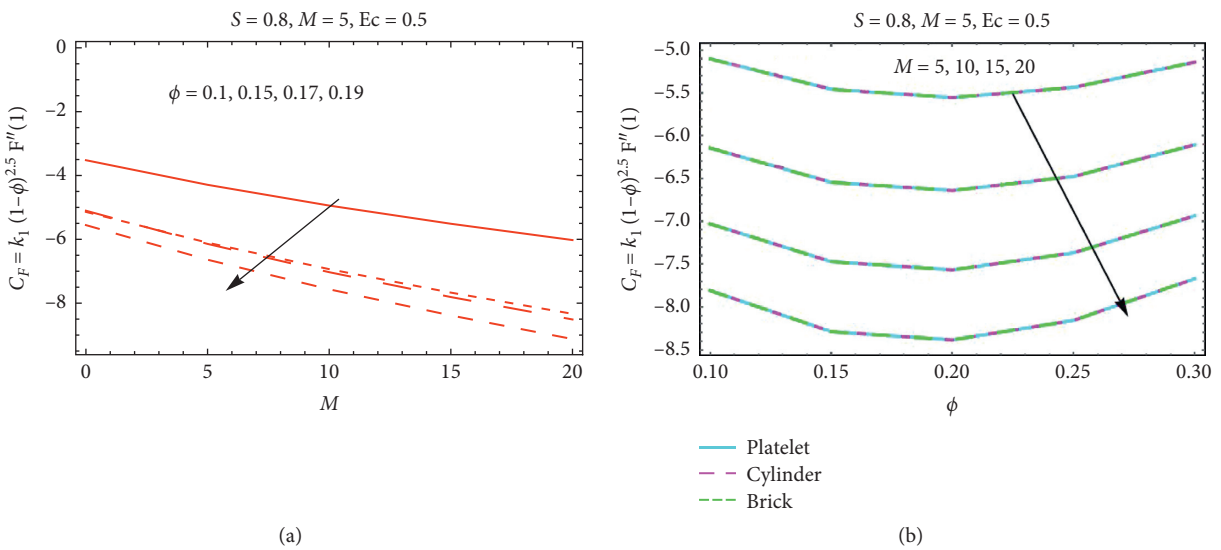


Figure 8: Skin friction against (a) ϕ and (b) M.

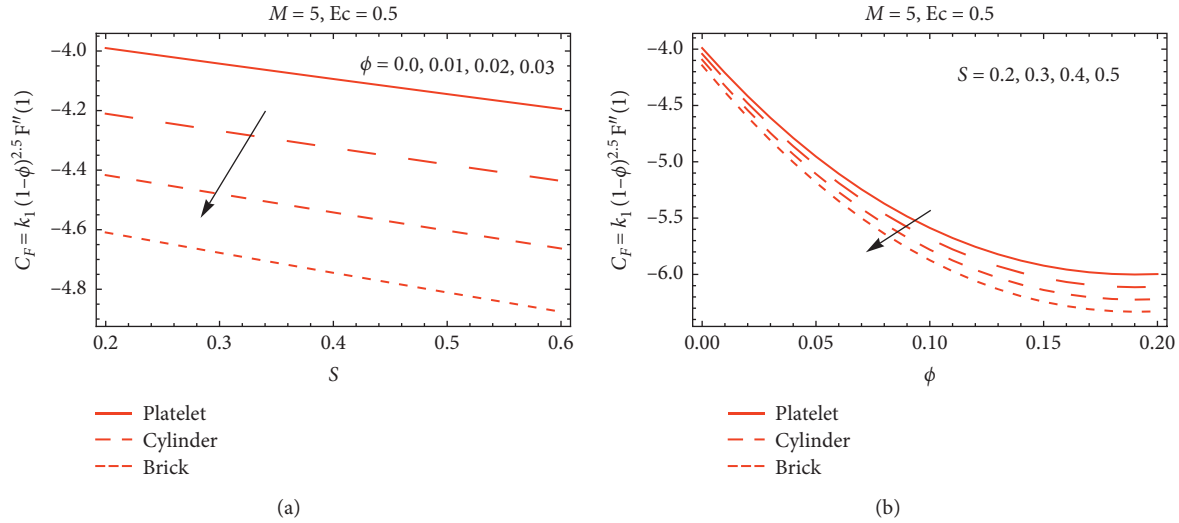


Figure 9: Skin friction against (a) ϕ and (b) S.

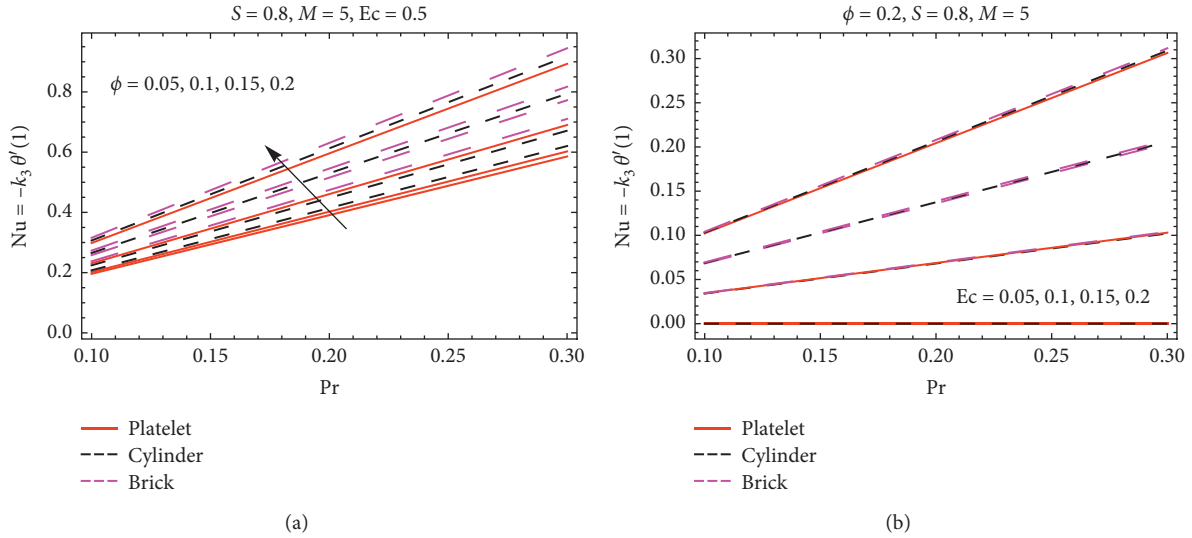


Figure 10: Nusselt number against (a) ϕ and (b) Ec.

volume fraction. Physically, the volume fraction enhances the thermal conductance, which significantly plays the role in the local heat transfer rate. Due to high thermal conductance of bricks nanomaterial, the local heat transfer rate prevailed in comparison with platelets and cylinder nanomaterial-based nanofluids. Figure 10(b) describes the local heat transfer for increasing Ec. It is pertinent that due to high dissipation, internal energy in the nanofluid rises due to which the local thermal performance rate rises. The local heat transfer rate is slow for platelets' nanomaterial-based nanofluids, while larger amount of heat transfer is noticed against bricks' nanomaterial-based nanofluids. Similarly, Figure 11 shows the increasing behavior of the local heat transfer against Prandtl variations.

Table 3 presents the comparative analysis between the present and existing science literature. From the comparison, it is revealed that the presented results are aligned with existing results which proves the reliability of the implemented technique and shows that the presented results are acceptable.

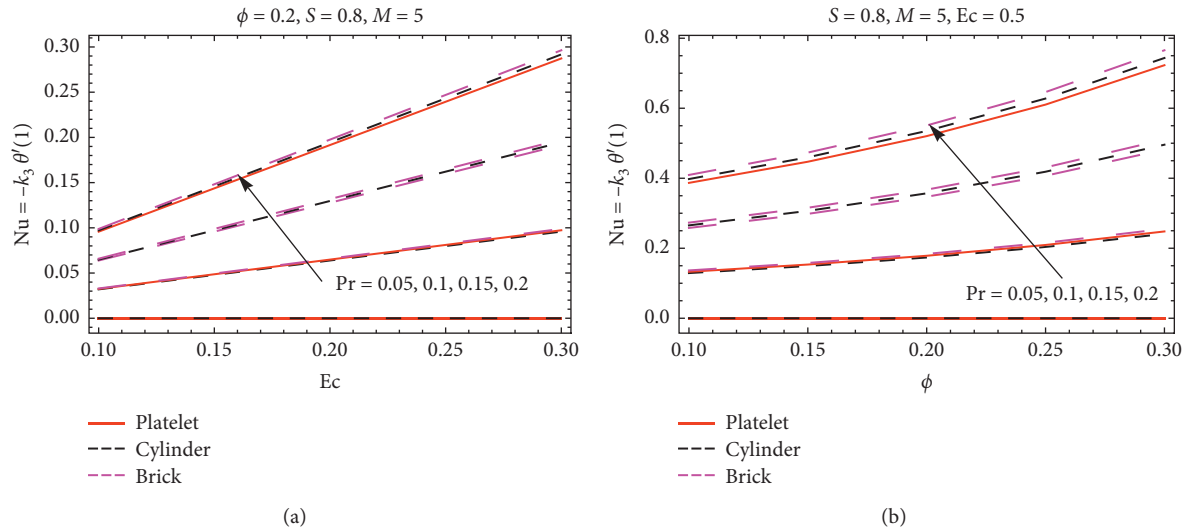


Figure 11: Nusselt number against (a) Pr and (b) ϕ.

Table 3: Comparison of $F(\eta)$ with existing scientific literature for $S = 1, \mathrm{Pr} = 6.2, \delta = 0.01, M = 0, \mathrm{Ec} = 0.01$, and $\phi = 0.02$.

η	Present		Sheikholeslami et al. [11]	
	VPM	Numerical	ADM	Numerical
0.0	0.000	0.000	0.000	0.000
0.1	0.141	0.141	0.141	0.141
0.2	0.280	0.280	0.280	0.280
0.3	0.415	0.415	0.415	0.415
0.4	0.544	0.544	0.544	0.544
0.5	0.663	0.663	0.663	0.663
0.6	0.771	0.771	0.771	0.771
0.7	0.863	0.863	0.863	0.863
0.8	0.935	0.935	0.935	0.935
0.9	0.982	0.982	0.982	0.982
1.0	1.000	1.000	1.000	1.000

4. Conclusions

An analytical investigation of the Cu nanomaterial-based nanofluid between parallel rotating sheets is conducted. The colloidal model is tackled analytically, and the results are plotted against the pertinent flow parameters. From the critical analysis of attained results, it is noticed that the temperature of the nanofluid declines against higher volumetric fraction and abruptly increases for the Prandtl number. These effects are prominent near the surface and gradually become slow apart from the plates. Moreover, the stronger viscous dissipation effects lead to the increment in the temperature of the nanofluid. The local rate of heat transfer significantly rises for brick-shaped nanoparticles. Finally, it is observed that the nanofluid is better for industrial uses due to its prominent heat transfer properties. The analysis is confined to the Newtonian host liquid; otherwise, the model will be not valid. A comparative analysis aligned the presented study with existing science literature.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] M. J. Stefan, "Versuch Uber die scheinbare adhesion," *Mathematisch-Naturwwissenschaftliche*, vol. 69, 1874.

[2] M. Mustafa, T. Hayat, and S. Obaidat, "On heat and mass transfer in the unsteady squeezing flow between parallel plates," *Meccanica*, vol. 47, no. 7, pp. 1581–1589, 2012.

[3] G. Domairry and A. Aziz, "Approximate analysis of MHD squeeze flow between two parallel disks with suction or injection by homotopy perturbation method," *Mathematical Problems in Engineering*, vol. 2009, Article ID 603916, 2009.

[4] U. Khan, N. Ahmed, and S. T. Mohyud-Din, "Analysis of magnetohydrodynamic flow and heat transfer of Cu-water nanofluid between parallel plates for different shapes of nanoparticles," *Neural Computing and Applications*, vol. 29, no. 10, pp. 695–703, 2018.

[5] S. Mohyud-Din, U. Khan, N. Ahmed, and S. Hassan, "Magnetohydrodynamic flow and heat transfer of nanofluids in stretchable convergent/divergent channels," *Applied Sciences*, vol. 5, no. 4, pp. 1639–1664, 2015.

[6] S. U. S. Choi, "Enhancing thermal conductivity of fluids with nanoparticles," in *Developments and applications of non-Newtonian flows. ASME FED, 231/MD66*, D. A. Siginer and H. P. Wang, Eds., pp. 99–105, 1995.

[7] S. U. S. Choi, Z. G. Zhang, W. Yu, F. E. Lockwood, and E. A. Grulke, "Anomalous thermal conductivity enhancement in nanotube suspensions," *Applied Physics Letters*, vol. 79, no. 14, pp. 2252–2254, 2001.

[8] H. Ş. Aybar, M. Sharifpur, M. R. Azizian, M. Mehrabi, and J. P. Meyer, "A review of thermal conductivity models for nanofluids," *Heat Transfer Engineering*, vol. 36, no. 13, pp. 1085–1110, 2015.

[9] S. Göktepe, K. Ertürk, and H. Erturk, "Comparison of single and two-phase models for nanofluid convection at the entrance of a uniformly heated tube," *International Journal of Thermal Sciences*, vol. 80, pp. 83–92, 2014.

[10] N. Hedayati and A. Ramiar, "Investigation of two phase unsteady nanofluid flow and heat transfer between moving parallel plates in the presence of magnetic field using GM," *Nano Micro Scale Sci. Tech.*vol. 4, no. 2, pp. 47–53, 2016.

[11] M. Sheikholeslami, D. D. Ganji, and H. R. Ashorynejad, "Investigation of squeezing unsteady nanofluid flow using ADM," *Powder Technology*, vol. 239, pp. 259–265, 2013.

[12] M. Sheikholeslami, T. Hayat, and A. Alsaedi, "MHD free convection of Al2O3-water nanofluid considering thermal radiation: a numerical study," *International Journal of Heat and Mass Transfer*, vol. 96, pp. 513–524, 2016.

[13] M. Sheikholeslami and M. M. Rashidi, "Non-uniform magnetic field effect on nanofluid hydrothermal treatment considering Brownian motion and thermophoresis effects," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 38, no. 4, pp. 1171–1184, 2016.

[14] M. Sheikholeslami and A. J. Chamkha, "Flow and convective heat transfer of a ferro-nanofluid in a double-sided lid-driven cavity with a wavy wall in the presence of a variable magnetic field," *Numerical Heat Transfer, Part A: Applications*, vol. 69, no. 10, pp. 1186–1200, 2016.

[15] A. S. Dogonchi, K. Divsalar, and D. D. Ganji, "Flow and heat transfer of MHD nanofluid between parallel plates in the presence of thermal radiation," *Computer Methods in Applied Mechanics and Engineering*, vol. 310, pp. 58–76, 2016.

[16] M. Sheikholeslami and D. D. Ganji, "Nanofluid flow and heat transfer between parallel plates considering Brownian motion using DTM," *Computer Methods in Applied Mechanics and Engineering*, vol. 283, pp. 651–663, 2014.

[17] M. Sheikholeslami, M. M. Rashidi, D. M. Al Saad, F. Firouzi, H. B. Rokni, and G. Domairry, "Steady nanofluid flow between parallel plates considering thermophoresis and Brownian effects," *Journal of King Saud University-Science*, vol. 28, no. 4, pp. 380–389, 2016.

[18] S. T. Mohyud-Din, Z. A. Zaidi, U. Khan, and N. Ahmed, "On heat and mass transfer analysis for the flow of a nanofluid between rotating parallel plates," *Aerospace Science and Technology*, vol. 46, pp. 514–522, 2015.

[19] K. Singh, S. K. Rawat, and M. Kumar, "Heat and mass transfer on squeezing unsteady MHD nanofluid flow between parallel plates with slip velocity effect," *Journal of Nanoscience*, vol. 2016, Article ID 9708562, 11 pages, 2016.

[20] S. T. Mohyud-Din, Adnan, U. Khan et al., "Thermal transport investigation in magneto-radiative GO-MoS2/H2O-C2H6O2 hybrid nanofluid subject to cattaneo-christov model," *Molecules*, vol. 25, no. 11, Article ID 2592, 2020.

[21] S. Z. A. Adnan, S. Z. A. Zaidi, U. Khan et al., "Impacts of freezing temperature based thermal conductivity on the heat transfer gradient in nanofluids: applications for a curved riga surface," *Molecules*, vol. 25, no. 9, Article ID 2152, 2020.

[22] N. Ahmed, Adnan, U. Khan et al., "Radiative colloidal investigation for thermal transport by incorporating the impacts of nanomaterial and molecular diameters (dNanoparticles, dFluid): applications in multiple engineering systems," *Molecules*, vol. 25, no. 8, Article ID 1896, 2020.

[23] Z. Shah, S. Islam, H. Ayaz, and S. Khan, "Radiative heat and mass transfer analysis of micropolar nanofluid flow of casson fluid between two rotating parallel plates with effects of Hall current," *Journal of Heat Transfer*, vol. 141, no. 2, 2019.

[24] U. Khan, Adnan, N. Ahmed et al., "γ-Nanofluid thermal transport between parallel plates suspended by micro-cantilever sensor by incorporating the effective Prandtl model: applications to biological and medical sciences," *Molecules*, vol. 25, no. 8, p. 1777, 2020.

[25] T. Gul, M. Z. Ullah, A. K. Alzahrani, and I. S. Amiri, "Thermal performance of the graphene oxide nanofluids flow in an upright channel through a permeable medium," *IEEE Access*, vol. 7, pp. 102345–102355, 2019.

[26] N. S. Akbar and A. W. Butt, "Ferromagnetic effects for peristaltic flow of Cu-water nanofluid for different shapes of nanosize particles," *Applied Nanoscience*, vol. 6, no. 3, pp. 379–385, 2016.

[27] R. Ellahi, M. Hassan, A. Zeeshan, and A. A. Khan, "The shape effects of nanoparticles suspended in HFE-7100 over wedge with entropy generation and mixed convection," *Applied Nanoscience*, vol. 6, no. 5, pp. 641–651, 2016.

[28] W. Sikandar, U. Khan, N. Ahmed, and S. T. Mohyud-Din, "Variation of parameters method with an auxiliary parameter for initial value problems," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 1959–1963, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 8353391, 9 pages
https://doi.org/10.1155/2021/8353391

Research Article

Data Analysis of Physical Fitness Monitoring Based on Mathematical Models

Cuixiang Guo,[1] **Junwu Suo**,[2] **Chunguang Xu**,[1] **Xinhua Yang**,[1] **and Liping Zhang**[1]

[1]*Shandong Polytechnic, Jinan 250104, China*
[2]*Shandong Jianzhu University, Jinan 250101, China*

Correspondence should be addressed to Cuixiang Guo; guocuixiang@sdp.edu.cn and Junwu Suo; suojunwu@sdjzu.edu.cn

Received 22 May 2021; Revised 5 June 2021; Accepted 9 June 2021; Published 22 June 2021

Academic Editor: Gengxin Sun

Physical monitoring and analysis are of great significance to improve national physical fitness. The main work of this paper is that the physical health level of college students is studied and analysed by using a statistical model and mathematical model. According to the characteristics of the collected data, different mathematical models are established. Firstly, the grey correlation model is used to analyse the correlation between pull-up and other physical fitness indexes. Then, based on the classification of college students and the influence and flow law of interclass crowd, a differential equation system is established based on the LMC model. By analysing the existence and stability of the equilibrium point of the system, as well as the possible folding bifurcation and backward bifurcation at the equilibrium point, this paper makes qualitative and quantitative research on the trend of college students' physical exercise on campus.

1. Introduction

Relying on data analysis technology and mathematical model to improve the development of national fitness monitoring can give full play to the application of data analysis technology in the field of economic and social life and effectively improve the information technology level of national fitness monitoring, so as to provide objective and scientific data support for the improvement of national fitness. College students have the characteristics of concentration and easy management, which is the breakthrough to effectively improve the national physical quality. However, what are the factors that affect college students' physical exercise and how to improve their participation in physical exercise are of great significance in improving the national physical quality. In the research of college students' physique, how to improve the number of college students participating in physical exercise is a problem to be solved. The study found that college students do not participate in or less participate in physical exercise activities, not because of lack of time, but with receiving sports information media publicity. And the publicity information mainly comes from regular classroom learning and through interaction between students.

The earliest physical monitoring can be traced back to the 1860s. In the United States, Hitchcock M. D. took the lead in making annual periodic measurements of all students' age, height, weight, length, circumference, vital capacity, and muscle strength in 1861. In 1954, the Kraus–Weber test, which involved six minimum levels of strength and flexibility of low back muscles, was developed and tested on American and European children. Subsequently, the National Federation of health, sports, rehabilitation, and dance designed a physical fitness test method covering seven indicators, including 50 m running, 600 m running, standing long jump, softball throwing, round-trip running, pull-up, and sit-ups.

At present, the research on the monitoring and measurement of college students' physique data mainly includes the following methods: the physical fitness level of college students is studied by using the methods of literature, questionnaire interview, physical test, and street survey [1–3], the influencing factors of college students' physique are explored by using correlation analysis and high-tech

instruments [4, 5], improving the model of college students' physique score [6, 7], combining machine-learning algorithm to predict college students' performance in body measurement [8–10]. Although these studies show the physical condition of college students from many fields, they can also classify and predict the physical condition of college students, but they cannot reflect the dynamic change process of college students' physical fitness behavior.

In order to improve the level of college students' daily physical exercise, researchers also studied the behavior and motivation of daily physical exercise. Bellar et al. [11] explored the level of college students' sports activities through a cross-sectional online survey. Feng et al. [12] used the generalized estimation model to analyse the influence of fitness consciousness and fitness behavior on college students' physique and found that strengthening daily exercise consciousness can effectively improve college students' physique level.

In order to find the hidden law of things, propagation dynamics [13, 14] can be used to build more realistic models. Different from the traditional statistical method, Choszcz et al. [15] applied population dynamics to sports activities, classified the objects into three categories, and established the mathematical model of human fixed group and nonfixed group to carry out competitive sports. Posluszny et al. [16] made a qualitative analysis of the fixed group model and gave the ordinary differential equation model and partial differential equation model of the nonfixed group. Wang et al. [17] proposed a group competitive sports activity model with a delay-dependent effect; after the qualitative research on the model, it is found that the number of times of college students participating in exercise every week and the time of each exercise are low.

Although the existing research shows the physical condition of college students from many aspects and it can also classify and predict the physique of college students, it cannot reflect the dynamic process of sports behavior. We found that according to different classification criteria, such as weekly exercise frequency, time, or place of each exercise, college students could be classified. Therefore, this paper classifies college students according to exercise frequency. Secondly, the will of physical exercise mainly comes from classroom learning and peer communication, based on the rule that the consciousness of physical exercise is affected; the linear and standard influence rates are used to establish the propagation dynamics model of sports population flow. Compared with the traditional statistical method, the dynamic method can make us understand the overall situation better from the flow direction of college students. By combining with computer simulation and other methods, we can have a more in-depth and comprehensive understanding of the changing trend of college students' participation in physical exercise.

2. Analysis of College Students' Pull-Up Performance Based on Grey Correlation Model

As a sport to test the development level of human upper limb muscle strength, pull-up has been accepted as a physical fitness test item in many countries. The data of this paper comes from the physical fitness test results of 7921 male students in a university in China in 2019. Physical fitness test indexes include pull-up (times/min), BMI (kg/m^2), vital capacity (ml), 50 meters (s), standing long jump (cm), sitting forward (cm), and 1000 meters (s).

Statistical analysis shows that the distribution of each variable follows a normal distribution. 3σ method is used to preprocess the data, and the outliers caused by improper records are removed. The basic information of the sample is shown in Table 1.

2.1. Correlation Analysis between Pull-Up and Other Indexes. Pearson correlation analysis [18] is often used to mine the correlation between variables in statistics to describe the linear correlation between variables. Pearson correlation coefficient can be used to describe the close relationship between two distance variables. The calculation formula is as follows:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 (y_i - \overline{y})^2}}, \quad (1)$$

where $\overline{x}$ and $\overline{y}$ represent average values of two samples and r is Pearson correlation coefficient; the larger the value, the higher the linear correlation between variables.

According to the sample data set, the Pearson correlation coefficient between the pull-up and each measurement index is shown in Table 2.

Table 2 shows that there is a significant correlation between pull-up and various physical measurement indexes, in which standing long jump and sitting forward are positively correlated with pull-up, and other indexes are negatively correlated with pull-up. BMI, standing long jump, and 1000 meters are highly correlated with pull-up, but the values of r are all lower than 0.5, so it is necessary to seek other methods to verify the correlation.

Correlation analysis is the comparison of the geometric relationship of the statistical data. The closer the geometric curve of the sequence is, the greater the correlation degree is. In this paper, the improved grey slope correlation degree is chosen to describe the correlation between pull-up and other indexes. The improved grey slope correlation degree can reflect the positive and negative of sequence correlation and eliminate the influence of dimensionless transformation of the original sequence on the correlation coefficient and correlation degree value. The calculation steps of the improved slope correlation degree method are as follows:

Table 1: The basic information of body measurement sample.

Index	BMI	Vital capacity	50 meters	Standing long jump	Sitting forward	1000 meters	Pull-up
Maximum value	14.31	4216	20	298	38.7	438	57
Minimum value	41.28	552	5.25	128	11.9	126	0
Average value	21.92	4009.2	7.83	227.3	20.6	251.9	5.8

Table 2: Pearson correlation coefficient between pull-up and each measurement index.

	BMI	Vital capacity	50 meters	Standing long jump	Sitting forward	1000 meters
r	-0.336^{**}	-0.092^{**}	-0.153^{**}	0.308^{**}	0.112^{**}	-0.276^{**}
Significance	0.000	0.000	0.000	0.000	0.000	0.000

**Significant correlation at 0.01 level.

Step 1: dimensionless transformations of reference sequence $s_0 = \{s_0(1), s_0(2), \ldots, s_0(n)\}$ and comparison sequence $s_i = \{s_i(1), s_i(2), \ldots, s_i(n)\}$ are carried out:

$$y_i(k) = \frac{s_i(k)}{s_i(1)}, \quad i = 0, \ldots, m; k = 1, \ldots, n. \tag{2}$$

The slope of each time point is calculated as follows:

$$\Delta y_i(k) = y_i(k+1) - y_i(k). \tag{3}$$

Then, the grey slope correlation coefficient of each time point is

$$\varepsilon_i(k) = \operatorname{sgn}\left(\Delta y_0(k), \Delta y_i(k)\right). \frac{1 + \left|\left(\Delta y_0(k)/\overline{\overline{y}_0}\right)\right|}{1 + \left|\left(\Delta y_0(k)/\overline{\overline{y}_0}\right)\right| + \left|\left(\Delta y_0(k)/\overline{\overline{y}_0}\right) - \left(\Delta y_i(k)/\overline{\overline{y}_i}\right)\right|}, \tag{4}$$

where $\overline{y}_0 = (1/n)\sum_{k=1}^{n} y_0(k)$, $\overline{y}_i = (1/n)\sum_{k=1}^{n} y_i(k)$, and $\operatorname{sgn}(\Delta y_0(k), \Delta y_i(k))$ is used to describe the positive and negative correlation of variables, when $\Delta y_0(k) \cdot \Delta y_i(k) \geq 0$. Its value is equal to 1; otherwise, its value is -1.

Step 2: by calculating the grey correlation degree and sorting, the correlation between the comparison sequence and the reference sequence can be obtained:

$$\gamma_i = \frac{1}{n-1}\sum_{k=1}^{n-1} \varepsilon_i(k), \tag{5}$$

where γ_i represents the correlation strength of index i.

The sampling data is divided according to the enrollment year, and the mean value of the physical measurement data of the students' enrollment in the same year is calculated to obtain a group of time series of the enrollment year from 2012 to 2020, shown as Table 3.

Taking pull-up as the reference sequence and the other six indexes as the comparison sequence, according to the above steps, the correlation coefficient and correlation degree between the comparison sequence and the reference sequence are calculated, and the results are shown in Tables 4 and 5 .

2.2. Prediction of Pull-Up Performance Trend. Grey prediction [19] refers to establishing the corresponding differential equation by mining the rules of the system change through the processing of the original data, so as to predict future development. From the Pearson correlation coefficient, it can be seen that the linear correlation between each variable and pull-up is not strong, so it is impossible to establish the regression model with standard accuracy. Therefore, the grey prediction model with small sample demand is selected to predict the development trend of college students' pull-up performance in the future. In this paper, the data of the first eight years from 2012 to 2020 are selected to establish the model, the data of the ninth year are used to test the prediction accuracy of the model, and the optimal prediction model is obtained by optimizing the initial value and background value.

In the process of model construction, firstly, y_t is calculated as an accumulation sequence of the original sequence x_t:

$$Y_t = \sum_{k=1}^{t} X_t, \quad t = 1, 2, \ldots, n. \tag{6}$$

Then, a weighted adjacent value sequence Z_t is generated as follows:

$$Z_t = \vartheta Y_t + (1-\vartheta) Y_{t-1}, \quad 0 < \vartheta < 1. \tag{7}$$

When traditional $GM(1,1)$ is used, $\vartheta = 0.5$, the mean value sequence $Z_t = 0.5(Y_t + Y_{t-1})$.

The constructed model is defined as follows:

Table 3: The data of students' enrollment in 1992–2000.

Year	BMI	Vital capacity	50 meters	Standing long jump	Sitting forward	1000 meters	Pull-up
2012	21.68	4142.8	7.47	232.8	21.6	245.5	6.2
2013	22.09	4152.1	7.81	228.6	21.4	249.5	5.9
2014	22.12	4094.2	7.92	229.2	20.9	250.2	5.3
2015	21.91	4133.9	7.87	228.3	20.5	252.6	5.1
2016	21.98	4032.8	7.82	228.5	19.8	252.7	4.8
2017	21.87	3982.7	7.85	227.5	18.9	252.9	4.6
2018	21.83	3996.2	7.89	224.8	18.6	252.1	4.2
2019	21.81	4048.6	7.94	221.4	17.7	249.3	3.6
2020	21.62	4067.9	8.02	222.6	19.9	251.7	3.3

Table 4: Grey correlation coefficient between pull-up and each measurement index.

Correlation coefficient	$\varepsilon_1(k)$	$\varepsilon_2(k)$	$\varepsilon_3(k)$	$\varepsilon_4(k)$	$\varepsilon_5(k)$	$\varepsilon_6(k)$
$k = 1$	−0.9252	−0.9407	0.9536	0.9513	−0.9262	−0.9268
$k = 2$	−0.9012	0.9116	−0.8931	−0.8977	0.9173	−0.8977
$k = 3$	0.9501	−0.9378	0.9519	0.9498	0.9872	−0.9376
$k = 4$	0.9807	0.9946	0.9831	−0.9762	0.9971	−0.9772
$k = 5$	0.9452	0.9501	−0.9412	0.9477	0.9821	−0.9418
$k = 6$	−0.9222	−0.9205	−0.9163	0.9332	−0.9225	0.9261
$k = 7$	0.9045	−0.8932	−0.8951	0.9162	0.9413	0.9107
$k = 8$	0.9478	−0.9359	−0.9301	−0.9327	−0.8578	−0.9301

Table 5: The correlation degree and ranking of pull-up and various physical measurement indexes.

	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
Correlation degree	0.2473	0.2212	−0.2114	0.2265	0.2638	−0.4017
Ranking	3	5	6	4	2	1

$$X_t + \mu Z_t = \varphi, \tag{8}$$

where X_t is grey derivative, μ is developing coefficient, φ is grey action, and Z_t is albino background value.

The parameters μ and φ are obtained by the least square method, and the estimated parameters are substituted into the albino differential equation:

$$\frac{dY_t}{dt} + \mu Y_t = \varphi. \tag{9}$$

The time response formula can be obtained by solving the above formula:

$$\widehat{Y}_{t+1} = \left(X_t - \frac{\varphi}{\mu}\right)e^{-\mu t} + \frac{\varphi}{\mu}. \tag{10}$$

The prediction expression of X_t can be obtained by subtracting $\widehat{Y}_t$:

$$\widehat{X}_t = \widehat{Y}_t - \widehat{Y}_{t-1}, \quad t = 2, 3, \ldots, n, \tag{11}$$

where $\widehat{X}_1 = Y_1 = X_1$.

The original sequence of the pull-up is the eight-dimensional sequence from 2012 to 2019. It is found that different initial values will also affect the prediction results of the prediction model. When the dimensions of the model are 4, 5, 6, 7, and 8, respectively, the model test results are shown in Table 6.

The prediction models with different initial values have passed the variance ratio test and small probability error P value test. When the model dimension is 5, the variance ratio (0.0583) is the smallest, and the average relative error is 0.98%, so the prediction performance of the model is the best.

Background value is used to estimate parameters μ and φ of albino differential equation, so the optimization of background value is an important means to improve the prediction accuracy of the model. Setting different value for ϑ can optimize the background; for different value of ϑ, the 5-dimensions model test results are shown in Table 7.

As shown in Table 7, when $\vartheta = 0.52$, the average relative error (1.25%) and variance ratio (0.0826) of the model are minimum.

Using the above optimal model, it can be predicted that the pull-up scores of students who will enroll in 2021 and 2022 are 2.8 and 2.6, respectively. The prediction results of the optimal model are shown in Table 8.

According to the trend of the actual value series and the prediction value series, with the increase of the student's enrollment year, the pull-up score shows a decreasing trend.

3. Analysis on College Students' Physical Exercise Trend Based on LMC Model

3.1. LMC Model Definition. The stage behavior change mode theory [20] divides the process of human behavior change into five main stages: preconsciousness stage, consciousness stage, preparation stage, behavior stage, and maintenance stage. Based on the stage behavior change mode theory, this

Table 6: Model test results of different dimensions.

Model dimensions	Actual value in 2020	Prediction value	Relative error (%)	Average relative error (%)	Variance ratio	P value test
4	3.27	3.22	1.5	1.06	0.0865	1
5	3.27	3.31	1.2	0.98	0.0583	1
6	3.27	3.45	5.5	2.75	0.1859	1
7	3.27	3.58	9.5	4.87	0.2136	1
8	3.27	3.55	8.6	3.77	0.1912	1

Table 7: Model test results of different ϑ in 5-dimensions model.

ϑ	Actual value in 2020	Prediction value	Relative error (%)	Average relative error (%)	Variance ratio	P value test
0.35	3.27	3.16	3.4	2.75	0.1298	1
0.50	3.27	3.21	1.8	1.87	0.0875	1
0.52	3.27	3.22	1.5	1.25	0.0826	1
0.60	3.27	3.21	1.8	1.41	0.0865	1
0.65	3.27	3.25	0.6	1.37	0.1557	1

Table 8: The prediction results of the optimal model.

Enrollment year	2017	2018	2019	2020	2021	2022
Actual value	4.6	4.2	3.6	3.3	–	–
Prediction value	4.5	4.2	3.5	3.2	2.8	2.6

paper divides the college students into three different stages according to the exercise frequency: little, medium, and completely enough. The year is taken as a unit; based on the LMC model, the law of the number of students in the three states is studied. The basic dynamic assumptions of the LMC model are as follows:

Little state: $L(t)$ represents the number of students who exercise no more than once a week at time t. In the preconsciousness stage, consciousness stage, and preparation stage, all the changes are only in the consciousness level, which is difficult to distinguish from the external performance. At the same time, they have a common characteristic; that is, they have no exercise behavior. In the process of participating in physical exercise behavior, the attitude of students in the little state to physical exercise will change with the increase of the breadth and depth of knowledge. Recognizing the benefits of exercise or the acquisition of sports skills will promote the awareness of students to begin to exercise.

Medium state: $M(t)$ represents the number of students who exercise between 2 and 3 times a week at time t. Students in the medium state realize the importance of physical exercise and begin to have physical exercise behavior, but the low-frequency behavior is not stable, which is a potential group to form the habit of full physical exercise.

Completely enough state: $C(t)$ represents the number of students who exercise more than 3 times a week at time t. Students in the completely enough state have a strong sense of exercise and good executive ability. They have developed their sports behavior into a part of daily life, which will not easily change and form a greater appeal to the surrounding students.

The total number of students in a university is $N(t)$, $N(t) = L(t) + M(t) + C(t)$. The change rules [21] between states are assumed as follows:

(1) Considering a university as the whole system, the number changes among the three states in terms of years will be explored. Therefore, there are no population dynamics factors such as birth, death, and migration, but there are new students and graduates who leave school every year; the enrollment rate is denoted as γ, and the graduation rate is denoted as β.

(2) For freshmen entering the University, the high school physical education curriculum ensures that all freshmen exercise at least once a week. At the same time, a certain proportion $(1 - q)$ of students will use their spare time for extra sports activities. So, at t time, the number of students in the little state is $q\ \gamma\ N(t)$, and the number of students in the medium state is $(1 - q)\ \gamma\ N(t)$.

(3) At t time, the number of students who take part in physical exercise is directly proportional to the number of students in class. Thus, the number of students moving out of the little state in unit time is proportional to $L(t)$, and the proportional coefficient is α. In unit time, $\alpha L(t)$ students enter the medium state.

(4) Suppose that the influence of peer communication is standardized and the flow of students between the three states is gradual. At t time, the number of students in the little state that can be influenced by a student with enough exercise is directly proportional to the proportion of students in the little state of the system, and the proportion coefficient is ω. The number of students with a medium state added in unit time is $\omega L(t)C(t)/N(t)$. At t time, the number of students in the medium state that can be affected by a student with sufficient exercise is directly proportional to the proportion of students in the medium state in the system, and the proportion coefficient is ε. The number of students with the

completely enough state added in time unit time is $\varepsilon M(t)C(t)/N(t)$.

Under the above four basic assumptions, the flow chart of the transition between the three states of the proposed LMC model is shown in Figure 1.

As shown in Figure 1, the students in the little state could transit to the medium state with probability $(\alpha L + \omega LC/N)$, and the students in the medium state could transit to the completely enough state with probability $\omega MC/N$.

Let $x = L/N$, $y = M/N$, $z = C/N$, if the total population of the model remains unchanged; LMC model can be transformed into the following equivalent system:

$$\begin{cases} \dfrac{dx}{dt} = q\gamma - \alpha x - \beta x - \omega xz, \\ \dfrac{dy}{dt} = (1-q)\gamma + \alpha x + \omega xz - \varepsilon yz - \beta y, \\ \dfrac{dz}{dt} = \varepsilon yz - \beta z. \end{cases} \tag{12}$$

3.2. Numerical Simulation and Sensitivity Analysis. In order to more intuitively observe the dynamic process of students' changes in each stage, numerical simulations are carried out in this section. The initial values are as follows: $x(0) = 0.6$, $y(0) = 0.3$, and $z(0) = 0.1$. Other parameters are $\omega = 0.9$, $\varepsilon = 0.5$, $q = 0.5$, and $\gamma = \beta = 0.25$ [21]. The fitting diagram of the system is shown in Figure 2.

It can be seen from Figure 2 that when the existing condition of positive equilibrium satisfies the inequality of α, the system tends to a nonzero equilibrium state. When the conditions are not satisfied, the system tends to the boundary equilibrium point:

$$R = \frac{((\gamma/\beta) - (\beta/\varepsilon))}{(q\gamma/(\alpha+\beta))}. \tag{13}$$

The value of threshold R has a great influence on the existence and stability of the equilibrium point. Through evaluating the influence of various parameters on the threshold, we can find out the measures to increase the number of students who exercise enough. The linear influence rate α of threshold R on classroom teaching, the proportion q of freshmen without extra exercise before admission to college, and the sensitivity of influence ε of the completely enough state on the medium state are discussed.

Considering that there is little change in the number of students enrolled and graduated each year [21], it is assumed that the annual enrollment rate and graduation rate of a university are equal $(\gamma = \beta)$. The sensitivity index of threshold R to variable β is defined as elasticity index:

$$\frac{ER}{E\beta} = \frac{\partial R}{\partial \beta} \times \frac{\beta}{R}. \tag{14}$$

Similarly, the elasticity index of threshold R for parameters ε, α, and q are as follows:

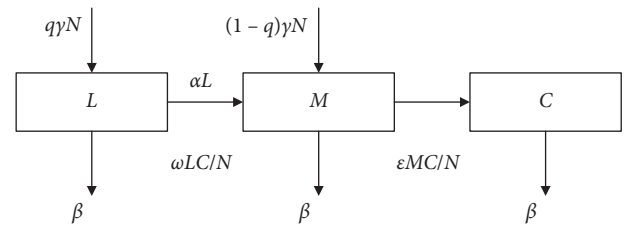


Figure 1: Flow chart of the transition between the three states of LMC model.

$$\begin{aligned} \frac{ER}{E\varepsilon} &= \frac{\partial R}{\partial \varepsilon} \times \frac{\varepsilon}{R} = \frac{\gamma\varepsilon}{\gamma\varepsilon - \beta^2} - 1, \\ \frac{ER}{E\alpha} &= \frac{\partial R}{\partial \alpha} \times \frac{\alpha}{R} = \frac{\alpha}{\alpha + \beta}, \\ \frac{ER}{Eq} &= \frac{\partial R}{\partial q} \times \frac{q}{R}. \end{aligned} \tag{15}$$

According to setting $\varepsilon = 0.5$, $q = 0.5$, $\gamma = \beta = 0.25$ [21], it can be concluded that

$$\begin{pmatrix} \dfrac{ER}{E\varepsilon} < 0, & 0 < \varepsilon < 0.25, \\ \dfrac{ER}{E\varepsilon} > 1, & 0.25 < \varepsilon < 0.5, \\ 0 < \dfrac{ER}{E\varepsilon} < 1, & 0.5 < \varepsilon < 1, \\ \lim\limits_{\varepsilon \longrightarrow 0.25} \left|\dfrac{ER}{E\varepsilon}\right| = +\infty, & \\ 0 < \dfrac{ER}{E\alpha} < 1, & 0 < \alpha < 1, \\ \left|\dfrac{ER}{Eq}\right| = 1, & 0 < q < 1. \end{pmatrix} \tag{16}$$

It means that the effect of ε on threshold R is complex. Firstly, when $0 < \varepsilon < 0.25$, with the increase of ε value, R becomes smaller and negative. According to the hypothesis, students in the completely enough state mainly come from the transformation of students in the medium state. In order to get a nonzero equilibrium point, the number of new students in the completely enough state in unit time at t time should be greater than the number of students who leave in unit time. That is, the value of ε should be greater than 0.25. When ε approaches 0.25 from the left, $ER/E\varepsilon$ tends to be negative infinity, and when ε approaches 0.25 from the right, $ER/E\varepsilon$ tends to be positive infinity. Secondly, when $0.25 < \varepsilon < 0.5$, with the increase of ε value, the value of $ER/E\varepsilon$ is smaller and smaller, but larger than 1. At this time, ε has a very strong influence on R. When $0.5 < \varepsilon < 1$, with the increase of ε value, $ER/E\varepsilon$ becomes smaller and smaller and

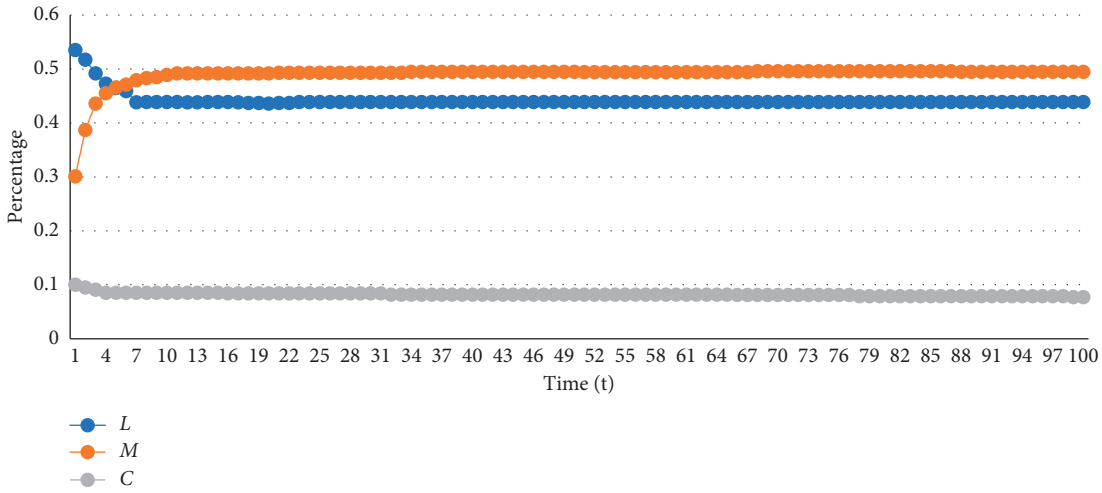


Figure 2: Fitting diagram of the system.

less than 1. When $0.5 < \alpha < 1$, $ER/E\alpha$ increases with the increase of ε and is less than 1. When $0 < q < 1$, ER/Eq is always -1.

Therefore, the following conclusions can be drawn: if the influence of students in the completely enough state on students in the medium state is less than the graduation rate, there is no positive equilibrium point, and the number of students in the completely enough state will eventually tend to zero. When students in the completely enough state have less influence on students in the medium state, attention should be paid. At this time, increasing ε can significantly increase the value of R; thus, the state of the equilibrium point can be changed. With the increasing influence of students in the completely enough state on students in the medium state, the role improvement of classroom education can have a greater effect.

In order to further determine which parameters have a great influence on the simulation results of the proposed model, the partial rank correlation coefficient (PRCC) [22] based on Latin hypercube sampling (LHS) [23] is used to analyse the parameter sensitivity of the model. PRCC measures the nonlinear but monotonic relationship between two variables and provides a measure of monotonicity between parameters. LHS is a stratified Monte Carlo sampling method, which divides the range of each parameter into n equal intervals according to the distribution function of parameters and randomly selects a sample from each interval.

In this paper, the sensitivity of the simulation results $L(t)$ and $M(t)$ of the model for four parameters q, α, ω, and ε is analysed. The sample size n is set to 1000, and the distribution and value range of each parameter are shown in Table 9.

The specific influence of parameters on different states is shown in Table 10.

The results in Table 10 show that the influence of each parameter on the final scale is very significant ($P < 0.01$) for both students in the little state and students in the completely enough state. The parameter q has the greatest influence on students in the little state. The parameter ω has the greatest influence on students in the completely enough state. Table 10 shows that the relative importance of parameters can be sorted differently among different output variables. For example, the effect of ω on $L(t)$ is relatively small, while that on $C(t)$ is the largest. The correlation between parameters and different output variables is also different. q was positively correlated with $L(t)$ and negatively correlated with $C(t)$. α, ω, and ε were negatively correlated with $L(t)$ and positively correlated with $C(t)$.

Based on theoretical analysis, numerical simulation, and sensitivity analysis, we find that proportion of students entering the little state each year q, the linear influence rate of classroom teaching α, the influence of students in the completely enough state on students in the little state ω, and the influence of students in the completely enough state on students in the medium state ε would affect the final stable state of the system.

Table 9: PRCC analysis parameters and sampling interval.

Parameter	Baseline value	Variance	Distribution	Sampling interval
q	0.8	0.02	Normal distribution	(0, 1)
α	0.5	0.02	Normal distribution	(0, 1)
ω	0.9	0.02	Normal distribution	(0, 1)
ε	0.5	0.02	Normal distribution	(0, 1)

Table 10: Specific influence of parameters on different states.

Order	Little state ($L(t)$)			Completely enough state ($C(t)$)		
	Parameter	PRCC	P value	Parameter	PRCC	P value
1	q	0.9369	0	ω	0.9918	0
2	α	−0.8875	0	q	−0.8847	$9.8523e-306$
3	ω	−0.8869	0	α	0.8295	$2.3092e-245$
4	ε	−0.5538	$2.9161e-83$	ε	0.3902	$9.6109e-43$

4. Conclusions

In this paper, in order to explore the current situation of college students' constitution and how to improve their physical level, two mathematical models are established based on different types of physical health data. First of all, the grey model of college physical fitness test results is established. Through correlation analysis and prediction, the factors that affect college students' pull-up performance are analysed. Then, the college students are classified according to the usual exercise frequency, and the LMC model is established to dynamically analyse the trend of college students' participation in physical exercise. The research, combined with computer simulation and mathematical model, can make people have a more in-depth and comprehensive understanding of the trend of college students' participation in physical exercise.

Data Availability

The basic data used in this paper were downloaded from the online public data set: Micro Data of College Students' Physical Health (https://bbs.pinggu.org/a-3280384.html).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by a grant from Shandong Family Education Association of China (no. 201833).

References

[1] J. M. Sacheck, J. F. Kuder, and C. D. Economos, "Physical fitness, adiposity and metabolic risk factors in young college students," *Medicine & Science in Sports & Exercise*, vol. 42, no. 6, pp. 1039–1044, 2009.

[2] B. K. Yoon, "The level of physical fitness is an independent indicator of metabolic disease in female college students," *International Journal of Human Movement Science*, vol. 8, no. 2, pp. 85–95, 2014.

[3] X. Lu, C. Yang, C. Yang et al., "Test method for health-related physical fitness of college students in mobile internet environment," *Mathematical Biosciences and Engineering*, vol. 16, no. 4, pp. 2189–2201, 2019.

[4] B. Peng, "Analysis and intervention on the influencing factors of college students' physical fitness," *Revista Brasileira de Medicina do Esporte*, vol. 27, no. 9, pp. 11–13, 2021.

[5] H. Rong, X. Cheng, J. M. Garcia et al., "Survey of health literacy level and related influencing factors in military college students in Chongqing, China: a cross-sectional analysis," *PLoS One*, vol. 12, no. 5, Article ID e0177776, 2017.

[6] T. Noguchi, S. I. Demura, and M. Shimada, "Effect of sports club activities on the physique and physical fitness of young Japanese males," *World Journal of Education*, vol. 3, no. 6, pp. 27–32, 2013.

[7] E. K. White and C. S. Warren, "The influence of social anxiety on the body checking behaviors of female college students," *Body Image*, vol. 11, no. 4, pp. 458–63, 2014.

[8] X. Wang, "Course-taking patterns of community college students beginning in STEM: using data mining techniques to reveal viable STEM transfer pathways," *Research in Higher Education*, vol. 57, no. 5, pp. 544–569, 2016.

[9] Z. Kang, "Using machine learning algorithms to predict first-generation college students' six-year graduation: a case study," *International Journal of Information Technology and Computer Science*, vol. 11, no. 9, pp. 1–8, 2019.

[10] B. Sudha, A. J. Samuel, and K. Narkeesh, "Feasibility online survey to estimate physical activity level among the students studying professional courses: a cross-sectional online survey," *Journal of Exercise Rehabilitation*, vol. 14, no. 1, pp. 58–63, 2018.

[11] D. Bellar, L. W. Judge, and J. Petersen, "Exercise and academic performance among nursing and kinesiology students at US colleges," *Journal of Education & Health Promotion*, vol. 3, no. 1, 9 pages, 2014.

[12] B. Feng, P. Zhou, and H. Dong, "Influence of long-distance running on mental health and positive sports emotions of college students," *Revista Argentina de Clinica Psicologica*, vol. 29, no. 2, pp. 685–691, 2020.

[13] G. Sun and S. Bin, "Router-level internet topology evolution model based on multi-subnet composited complex network model," *Journal of Internet Technology*, vol. 18, no. 6, pp. 1275–1283, 2017.

[14] S. Bin, G. Sun, N. Cao et al., "Collaborative filtering recommendation algorithm based on multi-relationship social

network," *Computers, Materials & Continua*, vol. 60, no. 2, pp. 659–674, 2019.

[15] D. Choszcz, R. Podstawski, and M. Wysocka-Welanc, "Measurement of motor fitness of students using the rowing ergometer," *Human Movement*, vol. 10, no. 1, pp. 46–52, 2009.

[16] J. Posluszny and L. A. Rubel, "The motions of an ordinary differential equation," *Journal of Differential Equations*, vol. 34, no. 2, pp. 291–302, 1979.

[17] Y. Wang, F. Brauer, J. Wu, and J. M. Heffernan, "A delay-dependent model with HIV drug resistance during therapy," *Journal of Mathematical Analysis and Applications*, vol. 414, no. 2, pp. 514–531, 2014.

[18] S. Bin and G. Sun, "Optimal energy resources allocation method of wireless sensor networks for intelligent railway systems," *Sensors*, vol. 20, no. 2, 482 pages, 2020.

[19] C.-H. Lin, J.-L. Chen, and P.-Z. Huang, "Dissolved gases forecast to enhance oil-immersed transformer fault diagnosis with grey prediction-clustering analysis," *Expert Systems*, vol. 28, no. 2, pp. 123–137, 2011.

[20] A. Biehl, A. Ermagun, and A. Stathopoulos, "Utilizing multi-stage behavior change theory to model the process of bike share adoption," *Transport Policy*, vol. 77, no. 5, pp. 30–45, 2019.

[21] A. Xw, A. Lq, and T. B. Pan, "Research on college students' physical exercise trend based on compartment model," *Mathematics and Computers in Simulation*, vol. 180, pp. 24–42, 2021.

[22] G. Sun, C.-C. Chen, and S. Bin, "Study of cascading failure in multisubnet composite complex networks," *Symmetry*, vol. 13, no. 3, 523 pages, 2021.

[23] G. L. Tian, S. Zhou, G. X. Sun, and C. C. Chen, "A novel intelligent recommendation algorithm based on mass diffusion," *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 4568171, 9 pages, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 1912442, 10 pages
https://doi.org/10.1155/2021/1912442

Research Article

Mathematical Model Construction of the Production Workshop Based on the Complex Network and Markov Theory

Gang Guo [1,2] **and Fengjing Shao** [1]

[1]*Institute of Complexity Science, Qingdao University, Qingdao 266071, China*
[2]*Qingdao Open University, Qingdao 266199, China*

Correspondence should be addressed to Fengjing Shao; sfj@qdu.edu.cn

Received 10 April 2021; Revised 7 May 2021; Accepted 25 May 2021; Published 8 June 2021

Academic Editor: Libor Pekař

Because of the advantages of the complex network in describing the interaction between nodes, the complex network theory is introduced into the production process of the modern workshop in this paper. According to the characteristics of the workshop, based on extracted key nodes, the complex network model of the workshop is constructed to realize the mathematical description of the production process of the workshop. Aiming at the multidisturbance factors in the production process of the workshop, the key disturbance factors are predicted based on the Markov method, and the propagation dynamics model close to the actual production of the workshop is established. Finally, the bottleneck prediction model of the workshop under the disturbance environment is established. The simulation results show that the proposed prediction model is in good agreement with the actual data, and the coincidence rate is as high as 93.7%.

1. Introduction

Production workshop is the basic unit of manufacturing enterprises, which consists of equipment, personnel, departments, and various workpieces. In the production workshop, the workpiece is processed to the final product according to the established process route, and the process is accompanied by many uncertain factors. These disturbance factors lead to the generation, transfer, and disappearance of the bottleneck in the workshop, and the bottleneck unit restricts the effective output of the production system. Therefore, it is necessary to study the dynamic prediction of the bottleneck in the workshop under the disturbance environment for improving the management of the production system.

In the manufacturing process, how to establish a reasonable and accurate mathematical model to identify and predict bottlenecks can provide a theoretical basis for enterprises to make production plans and management. Wu et al. [1] thought that, from the point of view of complex network statistical parameters, the cooperation, supply and demand, information transmission, and other aspects between production enterprises show self-similar characteristics. Hao et al. [2] used simulation software to simulate and analyze the production process of the production system. The research showed that the evolution process of the production system was extremely sensitive to the system parameters. Sun and Bin [3] analyzed that all kinds of manufacturing workshop cell network, mould manufacturing network, and all kinds of supply chain network show small-world characteristics.

In recent years, more and more researchers have applied the complex network theory to the manufacturing industry in order to make new breakthroughs in complex product development, supply chain optimization, and enterprise production management and optimization. Ma et al. [4] proposed the concept of collaborative manufacturing chain based on the complex network theory for the effective utilization of resources in the complex environment and finally realized the optimal utilization of resources. Cao et al. [5] modelled the production process of the product with the modular network, analyzed the evolution of the product

module by using the brittleness theory of the complex network, and improved the antirisk ability of the modular product network. Funke and Becker [6] analyzed the application of the complex network theory in the manufacturing system and provided a new method for network modelling of the manufacturing system. Giret et al. [7] applied the complex network theory to the service-oriented manufacturing network and analyzed the network structure and statistical parameters of the service-oriented manufacturing network.

Nowadays, artificial intelligent algorithms had been applied in the construction for different applications. Kea et al. [8] proposed a new model to estimate the disc cutter life by integrating a group method of data handling- (GMDH-) type neural network with a genetic algorithm. Wei et al. [9] proposed an analytical method for estimating the horizontal transition probability matrix, which is one of the important input parameters for the coupled Markov chain model. Shen et al. [10] presented an automatic pumping-recharge system to maintain groundwater balance during dewatering. Lin et al. [11] presented an approach for eutrophication evaluation based on the technique for order preference. Zhang et al. [12] proposed an artificial intelligence model to predict ground settlement. Ssl et al. [13] proposed an intelligent framework for predicting the advancing speed during earth pressure balance.

The definition and classification of bottlenecks are different according to different research objects, different methods, and different observation angles. Wang et al. [14] regarded the queuing number of parts in the processing area as the bottleneck identification index, and the bottleneck is the one with the largest queuing number. Samouei et al. [15] studied the bottleneck in the workshop from various aspects of the equipment parameters and identified the bottleneck unit by building a time-based mathematical model. Azadeh and Shoja [16] proposed a network model of bottleneck identification, which converted the main units of the workshop into a network model to identify the bottleneck location from the perspective of the system. Rui and Cheng [17] studied the workshop of the assembly line production mode, constructed a mathematical model based on the balance rate of the production line, and identified the bottleneck in real time. Based on the TOC constraint theory, Bin and Sun [18] combined with the intelligent algorithm and improved some processes of the previous genetic algorithm, so as to optimize the bottleneck identification method. According to the linear programming method, Brucker et al. [19] constructed a programming model for the objective function and constraints of the production workshop to identify the bottleneck unit of the production system. Masoud et al. [20] used the model graph and network model construction algorithm to identify the real-time bottleneck in the production workshop. Thurer and Stevenson [21] used simulation software to simulate the main production factors of the production system and compared the simulation data with the theoretical model and workshop data to identify the real-time bottleneck. Sweeney et al. [22] carried out orthogonal experiments on the factors that affect the output target of the system in the production workshop, compared with the mobile bottleneck generator, and verified the superiority of the orthogonal experiment bottleneck identification method.

The existing bottleneck research methods are limited to a specific parameter index, such as equipment, station, personnel, and materials, which lack systematic thinking. The bottleneck identification method in this paper is based on the complex network theory, the Markov model is simple, and it can be used to describe complex random phenomena. Markov process is used to analyze the disturbance factors, and the state transition probability and steady distribution of the Markov chain are obtained; it provides a way for the quantitative analysis of the disturbance factors in the production workshop. From the perspective of the complex network, all production factors are considered to identify and predict the bottleneck. And then, the accuracy and effectiveness of the proposed method are verified by the actual data of the workshop and simulation software.

2. Disturbance Factors' Analysis of the Production Workshop Based on the Markov Model

In this paper, Markov model is used to evaluate and predict the intensity of disturbance factors. The method has a good effect on process state prediction, and it can be used for production site state prediction. However, it is not suitable for medium- and long-term system prediction. From the perspective of the production process, the disturbance factors of the workshop are analyzed, the disturbance factor intensity matrix is constructed, the relationship between the matrix and disturbance factor intensity is simulated, and the Markov chain prediction model is established to predict the change of disturbance factor intensity, so as to determine the key disturbance factors and their occurrence probability. The flowchart indicating the bottleneck prediction of the production workshop is shown in Figure 1.

In each link of the workshop, there will be a variety of disturbance factors, such as demand change, emergency order insertion, equipment failure, machining accuracy, and personnel absence. At the same time, there will be a lot of inaccurate information, such as material arrival time, manual clamping time, and auxiliary processing time. The existence of these disturbance factors will make the workshop conditions change dynamically and seriously affect the normal production activities.

The disturbance in the workshop mainly refers to the factors that affect the effective output of production units such as equipment, personnel, process, and department. Disturbance factors can be divided into the following four categories: internal production environment, external environment, monitoring technology, and human factors. Disturbance factors caused by the internal production environment: although the production in the manufacturing workshop has been oriented to standardization and precision, the microdifferences and randomness in time cannot be eliminated. These uncertainties are mostly related to the internal production

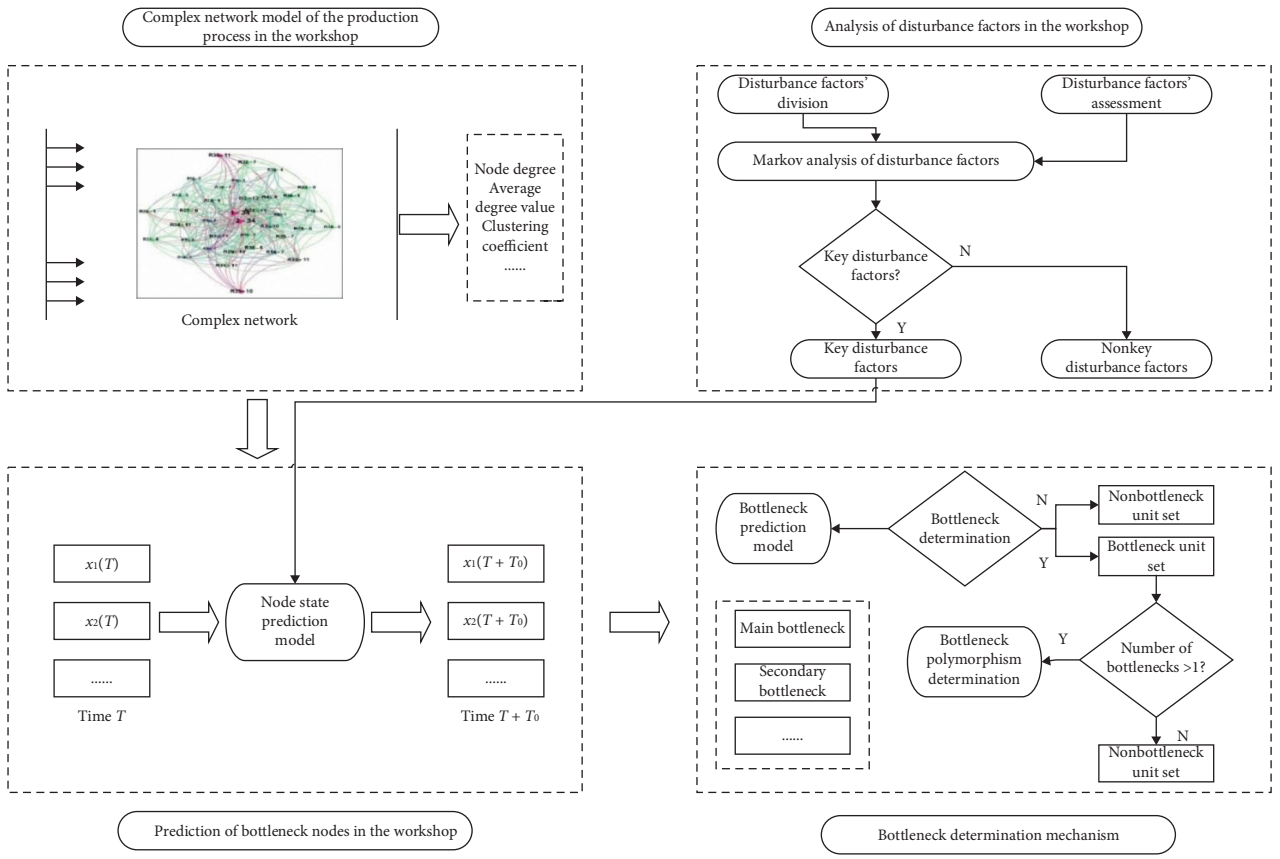


Figure 1: The flowchart indicating the bottleneck prediction of the production workshop.

environment, such as equipment random failure and equipment accuracy error. Disturbance factors caused by the external environment: they are materials not transported in place according to regulations, order changes, etc. The disturbance factors caused by the external environment will lead to the cancellation of existing production tasks or the change of production schedule, which is a kind of disturbance factors with great influence on the determination of results. The disturbance factors caused by monitoring technology include detection method, detection time, and detection environment. The influence of such disturbance factors is relatively weak. Disturbance factors caused by human factors mainly included workers absence, workers' proficiency, and workers' mood. These disturbance factors affect the quality and production schedule of the product directly.

The intensity of the disturbance factor is a quantitative description of the disturbance factor. The expression of disturbance factor intensity is different in different fields. The intensity of the disturbance factor reflects an inherent dynamic characteristic of the disturbance factor, which is used to describe the direct influence of the disturbance factor on the workshop network. The influence of different disturbance factors on the workshop is different; Chin et al. [23] classified the disturbance factors into four types: external disturbance, internal disturbance, monitoring disturbance, and human disturbance. The bottleneck influencing factor matrix is shown in Table 1.

There are various uncertain disturbance factors in the process of production and processing. The occurrence of disturbance factors may lead to the generation, transfer, increase, or decrease of bottlenecks in the production workshop. In this paper, the intensity of the disturbance factor is defined to express the comprehensive action degree of various disturbance factors in the process of production and processing. The intensity of the disturbance factor is expressed as follows:

$$I = \sum_{i=1}^{n} w_i d_i, \quad i = 1, 2, \ldots, n, \tag{1}$$

where w_i represents the importance of each disturbance factor and d_i represents the measurement indexes of the intensity of each human disturbance factor to the disturbance factor.

Because the intensity of disturbance factors in the production workshop is a comprehensive quantitative index of various disturbance factors and the intensity of disturbance factors is a fuzzy quantity, the intensity of disturbance factors in the production workshop can be measured by the fuzzy evaluation method at all levels, analytic hierarchy process, expert scoring, and other comprehensive methods. The specific steps are as follows:

(i) Step 1: the set of disturbance factors in the workshop is established. $D = \{d_1, d_2, d_3, d_4\}$ is the disturbance factors' set. Analytic hierarchy process

(AHP) is used to measure the disturbance intensity, $W = \{w_1, w_2, w_3, w_4\}$, $\sum_{i=1}^{4} w_i = 1$, $w_i \geq 0$. The weight of fuzzy subsets in D is $w_i = \{w_{i1}, w_{i1}, \ldots, w_{im}\}$, $\sum_{i=1}^{m} w_{im} = 1$, $w_{im} \geq 0$.

(ii) Step 2: the workshop influence evaluation set is established. By using the expert scoring method, the manufacturing workshop bottleneck research experts score the intensity of disturbance factors in the production workshop and grade different disturbance factors. Each disturbance factor is divided into five levels according to the influence level, and the level from low to high represents the influence degree. It is assumed that the evaluation set is $C = \{c_1, c_2, c_3, c_4, c_5\}$; c_i represents the influence of each disturbance factor on the intensity of the disturbance factor in the workshop.

(iii) Step 3: the evaluation of the disturbance factor to disturbance factor intensity is obtained by expert scoring, fuzzy function $f: D \longrightarrow F(V)$ is established, $F(V)$ is the fuzzy complete set of subset V, and $d_i \longrightarrow f(d_i) = (d_{i1}, d_{i2}, d_{i3}, d_{i4}, d_{i5})$ is the feedback of disturbance factor d_i in the workshop to the expert scoring system and evaluation set. It is assumed that the feedback vector of disturbance factors to bottleneck set V is $R_i = \{r_{i1}, r_{i2}, \ldots, r_{i5}\}$; after fuzzy transformation, the following formula can be obtained:

$$B_i = w_i \cdot R_i = (b_1, b_2, b_3, b_4, b_5). \tag{2}$$

The obtained values are imported into the upper layer by the analytic hierarchy process and then evaluated. Finally, the overall disturbance factor intensity is calculated as

$$D = W \cdot R = (D_1, D_2, D_3, D_4, D_5). \tag{3}$$

The bottleneck of the production workshop is produced in the production activities, and it is the result of the comprehensive effect of various disturbance factors in the production activities. The intensity of disturbance factors caused by human disturbance factors is divided into states. Markov model is used to predict the divided states. Finally, the probability of each state is obtained to predict the key disturbance factors.

In this paper, Markov chain is used to model the disturbance intensity caused by disturbance factors in the workshop. $S = \{s_1, s_2, \ldots, s_n\}$ represents the state set; it is the division of the disturbance factor intensity and the probability measure of the disturbance factor intensity under the comprehensive action of each disturbance factor in the production workshop. S can be preliminarily determined according to the historical data of disturbance factors in the production workshop. The state of disturbance factors is defined in Table 2.

The initial probability of the state set is determined by fitting the curve with the historical data of disturbance factors in the workshop. The probability of the initial state is expressed by vector a:

$$a = [a_1, a_2, \ldots, a_k], \tag{4}$$

where a_k represents the occurrence probability of state k.

In the Markov process, x_k is the state in time step t_k, $x_i(k) = P(x_k = i)$, the probability in matrix P_{ij} represents the probability that the state is i at this moment and j at the next moment. The matrix expression is as follows:

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}, \tag{5}$$

where p_{ij} is defined as follows:

$$p_{ij} = \frac{N_{ij}}{\sum_{j=1}^{n} N_{ij}}, \tag{6}$$

and N_{ij} is the statistical historical data; it represents the number of times that the node state changes from i to j in the production activity of the workshop.

According to the Markov theory, the probability of the state of disturbance intensity in the workshop can be calculated as follows:

$$\widehat{p} = a \cdot p. \tag{7}$$

3. Bottleneck Prediction of the Production Workshop Based on the Complex Network under the Disturbance Environment

In a workshop, N workpieces are produced and processed on M sets of equipment, and workpieces of the same specification have their own process routes. It is assumed that the same equipment can only process one workpiece at the same time, and the processing sequence is first come first process. There are time constraints and process constraints in the processing of workpieces. Different equipment processes different workpieces with different times and processes. The variables are defined as follows.

The workpiece sequence is represented by $J = \{J_j | j = 1, 2, \ldots, N\}$, J_j represents the j-th workpiece, the equipment sequence is represented by $Q = \{Q_i | i = 1, 2, \ldots, M\}$, Q_i represents the i-th equipment, and the process sequence is represented by P_p; then, the process matrix of N workpieces produced and processed on M sets of equipment is defined as

$$P_p = \begin{bmatrix} P_{p1(1)} & \cdots & P_{p1(n)} \\ \vdots & \ddots & \vdots \\ P_{pj(1)} & \cdots & P_{pj(n)} \end{bmatrix}, \tag{8}$$

where n represents the number of processes and $P_{ij(n)}$ represents the n-th process of the j-th workpiece produced on the i-th equipment.

Machining time series of the workpiece is represented by T_i, which is defined as follows:

Table 1: The bottleneck influencing factor matrix.

Disturbance factors		Disturbance intensity level				
		Level 1 0-0.1	Level 2 0.1–0.3	Level 3 0.3–0.5	Level 4 0.5–0.8	Level 5 0.8–1.0
External disturbance factors	Order change				*	
	Material supply					*
	Production plan				*	
	Policy changes			*		
Internal disturbance factors	Equipment failure				*	
	Route change					*
	Machining accuracy	*				
	Dispatcher change		*			
Human factors	Personnel absence					*
	Proficiency				*	
Monitoring factors	Monitoring method		*			
	Monitoring technology	*				
	Monitoring environment			*		

$$T_i = \begin{bmatrix} T_{i1(1)} & \cdots & T_{i1(n)} \\ \vdots & \ddots & \vdots \\ T_{ij(1)} & \cdots & T_{ij(n)} \end{bmatrix}, \tag{9}$$

where $T_{ij(n)}$ represents the machining time of the n-th process of the j-th workpiece on the i-th equipment.

Each resource node (such as workshop department, equipment personnel, and tools) involved in the production process of a production workshop is regarded as a network node. The possible process routes and logistics paths between nodes are regarded as the connecting edges in the network. The direction of the connecting edges between nodes is determined by the priority relationship of processes; as the weight on the connected edge of the network, the device load is used to measure the closeness of the relationship between nodes. Therefore, each production process constitutes a complex multitask weighted directed network model. An example of a workshop network model is shown in Figure 2.

Figure 2 represents a workshop production process which consists of two workshops, node P represents the resources (equipment, personnel, departments, etc.) in the manufacturing system, the direction of the edge represents the flow direction of the logistics process, and the node state is represented by the load of the node. Thus, the disturbance factors can be described as follows:

$$\begin{cases} \overline{P_p} & = P_p + \Delta P_p, \\ \overline{x_i}(t) & = x_i(t) + \Delta x_i(t), \\ \overline{\vartheta} & = \vartheta + \Delta\vartheta, \end{cases} \tag{10}$$

where $\overline{P_p}$ represents changes in disturbance factors, ΔP_p represents the process matrix change increment, $\Delta x_i(t)$ represents the fluctuation of the node state caused by disturbance factors, and $\Delta\vartheta$ represents the influence of disturbance factors on network coupling strength.

Considering the interaction between nodes and the dynamic characteristics of nodes in the whole network, a coupled map lattice node state prediction model with the network scale of n is constructed in this paper.

$$x_i(t+1) = \left| (1-\overline{\vartheta}) f(x_i(t)) + \frac{\overline{\vartheta} \sum_{j=1, j\neq i}^{N} \left(a_{i,j} + r_{i,j}\right)}{k(i)} \right|, \tag{11}$$

where $x_i(t)$ represents the state value of node i at time t, $x_i(t+1)$ represents the calculated value of the next time step of the node, $k(i)$ represents the degree of node i in the complex network, $a_{i,j}$ represents the network connection matrix, ϑ represents the coupling strength between nodes, function f represents the dynamic behaviour of nodes, and $r_{i,j}$ represents the fluctuation of the connection matrix after disturbance.

In this paper, from the perspective of the complex system, considering the network characteristics such as node degree value and clustering coefficient, the bottleneck judgment standard is given. Finally, according to the judgment standard, the bottleneck classification is implemented, such as primary and secondary bottlenecks and nonbottlenecks [24]. Suppose the bottleneck criterion is τ; the calculation formula of τ is as follows:

$$\tau = \alpha C + \beta K + \gamma L, \tag{12}$$

where C represents the network clustering coefficient, K represents the node degree, L represents the node load, and α, β, and γ represent weights, and $\alpha + \beta + \gamma = 1$.

Therefore, the bottleneck identification formula in the network is defined as follows:

$$\begin{aligned} \mathrm{BS}_N &= \{X_i(t) \geq \tau\}, \\ \mathrm{BS}_{\mathrm{non}-N} &= \{0 < X_i(t) < \tau\}, \end{aligned} \tag{13}$$

where BS_N represents the bottleneck sequence and $\mathrm{BS}_{\mathrm{non}-N}$ represents the nonbottleneck sequence.

In order to analyze the fluctuation and influence of disturbance factors on the production process of the workshop, the load change rate F_i is defined to describe the fluctuation caused by disturbance factors:

Table 2: State of disturbance factors.

Frequency of disturbance factors	State
$F > 7$	s_5
$4 \le F < 7$	s_4
$2 \le F < 4$	s_3
$0 \le F < 2$	s_2
$F \ge 0$	s_1

$$F_i = \sum_{t=1}^{T} \lim_{\Delta t \longrightarrow 0} \frac{\left| x_i(t+1) - x_i(t) \right|}{\Delta t}, \tag{14}$$

where T represents the total simulation time.

4. Simulation Results and Discussion

4.1. Analysis of Disturbance Factors. In the experiment, the actual production workshop data of an automobile manufacturing enterprise [25] are taken as an example; Markov process is used to analyze the disturbance factors. Firstly, the analytic hierarchy process is used to analyze many disturbance factors and determine their respective weights, as shown in Table 3.

After classifying and subdividing the disturbance factors and calculating the weight, the occurrence frequency is fitted according to the historical data of the disturbance factors in the production workshop. The state probability transition matrix is determined according to Table 3, in which the state probability transition diagram and probability matrix are as follows:

$$P = \begin{bmatrix} 0.382 & 0.387 & 0.288 & 0 & 0 \\ 0.522 & 0.330 & 0.120 & 0.120 & 0 \\ 0.450 & 0.270 & 0.350 & 0.120 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{15}$$

The production process parameters of the workshop are shown in Table 4.

According to Table 4, the node state transition frequency matrix in the process of production activities is calculated:

$$N = \begin{bmatrix} 6 & 6 & 4 & 0 & 0 \\ 6 & 3 & 2 & 1 & 0 \\ 3 & 2 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{16}$$

According to the expert scoring method, the initial state probabilities of ten disturbance factors to the disturbance factor intensity are determined as follows:

$$a = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 & 0.0 \end{bmatrix}. \tag{17}$$

In this way, the prediction probability of each node state S_i will be calculated as follows:

$$\widehat{p} = a \cdot p = \begin{bmatrix} 0.382 & 0.380 & 0.201 & 0.027 & 0 \end{bmatrix}. \tag{18}$$

It can be seen that, in the production process, in each state of 23 production process indicators, the key disturbance factors are equipment failure, material supply, order change, and so on, and the influence of the occurrence probability of these disturbance factors accounted for 86.2%.

4.2. Bottleneck Prediction Model Simulation. Aiming at an auto parts production workshop, the data of the production process are collected in hours by sensors on the production line. According to the working characteristics, resource flow, process constraints, information flow, and personnel allocation of the production workshop, the data are collected, and the key nodes are extracted, and the complex network model is constructed to realize the network mathematical description of the production process of the production workshop. The specific production data of 10 sets of equipment and 6 workpieces' sequence are shown in Table 5.

The layout of the workshop is shown in Figure 3.

After the initial state of each node is set, the subsequent state of each node is calculated according to the prediction model. The node state values without disturbance are shown in Table 6.

Table 6 shows that the state values of each node are basically stable without disturbance, and there is no bottleneck unit.

When the disturbance factor is added, the node state values of the workshop after 3 hours of disturbance are shown in Table 7.

According to the criterion of bottleneck judgment [26], the state value of each node is counted to judge the bottleneck. The results are shown in Table 8.

As shown in Tables 7 and 8, before the disturbance, the state value of the workshop nodes fluctuates little, and there is no bottleneck unit. After the disturbance factor is added, the state value of each node begins to be affected. Through simulation calculation, it is found that equipment P_6 is the primary bottleneck and equipment P_5 is the secondary bottleneck. Through the analysis of Table 8, it is found that the degree value and clustering coefficient of the bottleneck unit are larger than other nodes, and the bottleneck unit appears on the equipment directly affected by order insertion.

In order to verify the rationality and correctness of the model, the production process of the workshop is simulated and analyzed. Through the simulation analysis, when there is no disturbance factor, the production activities of the

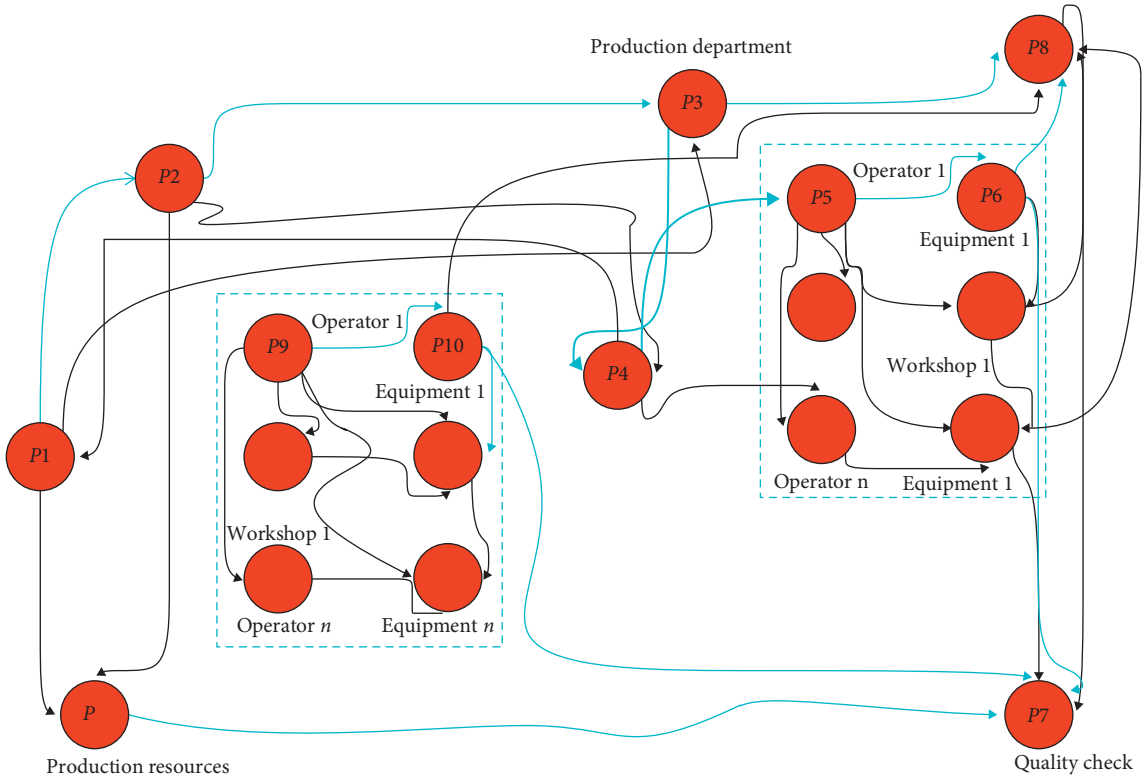


Figure 2: An example of a workshop production network model.

Table 3: Disturbance factors and their weights in the workshop.

Symbol	a_1	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
Disturbance factors	External disturbance	Order change	Material supply	Demand change	Policy change	Environment change
Weight	0.5637	0.1947	0.4628	0.1889	0.0725	0,.0733
Symbol	a_2	a_{21}	a_{22}	a_{23}	a_{24}	
Disturbance factors	Internal disturbance	Equipment failure	Process change	Machining accuracy	Workshop scheduling	
Weight	0.2631	0.2751	0.0643	0.5502	0.1236	
Symbol	a_3	a_{31}	a_{32}	a_{33}		
Disturbance factors	Human factor	Personnel absence	Skill level	Quality defects		
Weight	0.1202	0.0927	0.5201	0.3896		
Symbol	a_4	a_{41}	a_{42}	a_{43}		
Disturbance factors	Monitoring	Monitoring method	Monitoring technology	Monitoring environment		
Weight	0.0565	0.6386	0.1037	0.2579		

workshop are normal, and there is no bottleneck. With the input of the disturbance factor, the process matrix changes, resulting in the changes of the network topology, clustering coefficient, node degree value, node state, and process matrix. Accuracy index is used to verify the accuracy of the prediction model, and it is defined as follows:

$$\text{accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})}, \tag{19}$$

where TP represents true positive, FP represents false positive, FN represents false negative, and TN represents true negative.

The fluctuation caused by disturbance is positively correlated with the degree value, but not with betweenness and clustering coefficient. After the disturbance, the proposed prediction method in this paper is in good agreement with the actual data, and the coincidence rate is as high as 93.7%. The coincidence rate with other prediction methods

Table 4: The production process parameters of the workshop.

A_i	D_i	F_i	S_i
A_1 (order information)	0.2652	1	S_1
A_2 (material purchasing plan)	0.3275	0	S_1
A_3 (demand risk analysis)	0.1861	1	S_1
A_4 (initial data)	0.3785	1	S_2
A_5 (material arrival time)	0.3952	0	S_2
A_6 (manual clamping time)	0.3562	1	S_1
A_7 (auxiliary processing time)	0.3283	0	S_1
A_8 (industry policy changes)	0.3183	1	S_2
A_9 (market environment changes)	0.5902	3	S_3
A_{10} (equipment failure rate)	0.3392	0	S_2
A_{11} (aging degree of equipment)	0.3287	1	S_1
A_{12} (importance of equipment)	0.3916	1	S_3
A_{13} (route layout)	0.3285	1	S_2
A_{14} (process constraints)	0.3852	0	S_2
A_{15} (accuracy of machining different parts)	0.3907	1	S_2
A_{16} (selection of the production scheduling mode)	0.5589	2	S_3
A_{17} (scheduling objectives)	03205	0	S_1
A_{18} (staff leave rate)	0.5960	3	S_3
A_{19} (processing pass rate)	0.2298	2	S_3
A_{20} (time required for monitoring)	0.3805	5	S_4
A_{21} (monitoring environmental impacts)	0.3607	3	S_4
A_{22} (production line balance rate)	0.2385	1	S_2
A_{23} (change of delivery date)	0.3762	2	S_3

Table 5: Original data of workpiece production.

Workpiece	Process route/arrival rate per unit time/processing rate per unit time					
J_1	*P*1/35/120	*P*2/40/110	*P*3/40/90	*P*4/35/100	*P*5/35/130	—
J_2	*P*10/35/100	*P*9/40/100	*P*8/30/90	*P*7/40/80	*P*6/40/110	—
J_3	*P*10/25/100	*P*3/35/80	*P*6/30/90	—	—	—
J_4	*P*1/30/110	*P*8/30/120	*P*5/30/80	—	—	—
J_5	*P*1/45/120	*P*4/45/150	*P*6/40/160	—	—	—
J_6	*P*10/45/110	*P*2/55/100	*P*8/55/120	*P*6/45/130	*P*5/55/110	—

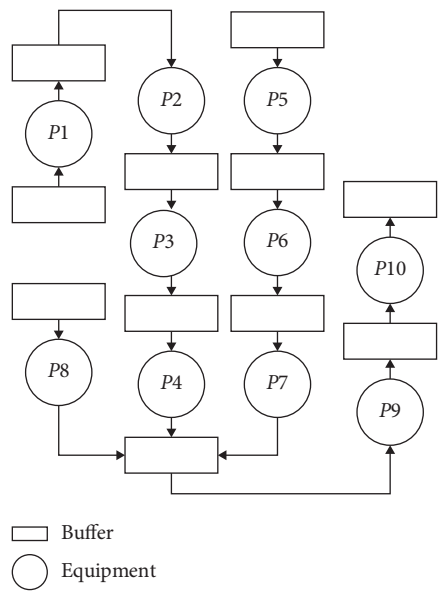


Figure 3: The layout of the workshop.

Table 6: Node state values without disturbance.

Time (d)	Node state values									
	*P*1	*P*2	*P*3	*P*4	*P*5	*P*6	*P*7	*P*8	*P*9	*P*10
1	0.875	0.400	0.875	0.625	0.538	0.763	0.429	0.542	0.365	0.625
2	0.677	0.706	0.687	0.809	0.925	0,.736	0.659	0.980	0.937	0.836
3	0.675	0.833	0.738	0.639	0.206	0.755	0.482	0.196	0.286	0.569
4	0.565	0.686	0.737	0.579	0.198	0.859	0.783	0.837	0.691	0.896
5	0.763	0.902	0.849	0.782	0.390	0.491	0.509	0.359	0.425	0.739
6	0.803	0.285	0.729	0.832	0.849	0.729	0.492	0.401	0.839	0.685
7	0.586	0.339	0.605	0.572	0.449	0.839	0.837	0.938	0.729	0.817
8	0.806	0.839	0.358	0.609	0.362	0.937	0.885	0.829	0.582	0.606

Table 7: Node state values after 3 hours of disturbance.

Time (d)	Node state values									
	*P*1	*P*2	*P*3	*P*4	*P*5	*P*6	*P*7	*P*8	*P*9	*P*10
1	0.875	0.400	0.875	0.625	0.538	0.763	0.429	0.542	0.365	0.625
2	0.677	0.706	0.687	0.809	0.925	0,.736	0.659	0.980	0.937	0.836
3	0.875	0.952	0.896	1.000	0.984	1.255	0.482	1.000	0.325	1.000
4	0.865	0.977	0.904	0.892	0.795	1.389	0.886	0.949	0.892	0.916
5	0.763	0.902	0.849	0.782	1.286	0.695	0.616	0.479	0.495	0.739
6	0.933	0.858	0.786	0.897	1.497	0.881	0.797	0.771	0.739	0.657
7	0.266	0.217	0.605	0.572	1.930	0.780	0.633	0.814	0.796	0.817
8	0.766	0.583	0.658	0.509	3.362	0.837	0.775	0.626	0.594	0.316

Table 8: Statistics of workshop network parameters.

Node	Node degree	Clustering coefficient	State fluctuation rate	Node average state value	Bottleneck node
*P*1	3	0.667	0.196	0.722	—
*P*2	4	0.333	0.175	0.672	—
*P*3	4	0.167	0.189	0.736	—
*P*4	5	0.500	0.072	0.741	—
*P*5	3	0.333	1.020	1.106	Secondary bottleneck
*P*6	6	0.667	1.108	1.528	Primary bottleneck
*P*7	2	0.167	0.207	0.532	—
*P*8	3	0.200	0.433	0.602	—
*P*9	4	0.500	0.051	0.622	—
*P*10	3	0.333	0.176	0.752	—

[27, 28] is also high; it shows that the theoretical analysis of the disturbance factors in the proposed prediction model is consistent with the simulation results, which verifies the correctness and rationality of the proposed model.

5. Conclusions

In this paper, a complex network is introduced into the production process of the production workshop. According to the characteristics of the production workshop, such as the working characteristics, resource flow, process constraints, information flow, and personnel allocation, the data are collected, and the key nodes are extracted, and then the complex network model is constructed to realize the network mathematical description of the production process of the production workshop. The disturbance factors are described mathematically, and the mechanism of the disturbance factors in the complex network system is constructed, so as to identify the bottleneck position of the production workshop under the disturbance factors.

Data Availability

The basic data used in this article are downloaded from the online public dataset weighted tardiness (http://people.brunel.ac.uk/~mastjjb/jeb/info.htm).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by a grant from the National Natural Science Foundation of China (no. 91130035).

References

[1] B. Wu, P. Liu, and X. Xu, "An evolutionary analysis of low-carbon strategies based on the government-enterprise game in the complex network context," *Journal of Cleaner Production*, vol. 141, pp. 168–179, 2017.

[2] X. Hao, H. An, H. Qi, and X. Gao, "Evolution of the exergy flow network embodied in the global fossil energy trade: based on complex network," *Applied Energy*, vol. 162, pp. 1515–1522, 2016.

[3] G. Sun and S. Bin, "Router-level internet topology evolution model based on multi-subnet composited complex network model," *Journal of Internet Technology*, vol. 18, no. 6, pp. 1275–1283, 2017.

[4] F. Ma, H. Xue, K. F. Yuen et al., "Assessing the vulnerability of logistics service supply chain based on complex network," *Sustainability*, vol. 12, no. 5, p. 1991, 2020.

[5] H. Cao, H. Zhang, and M. Liu, "Global modular production network: from system perspective," *Procedia Engineering*, vol. 23, pp. 786–791, 2011.

[6] T. Funke and T. Becker, "Complex networks of material flow in manufacturing and logistics: modeling, analysis, and prediction using stochastic block models," *Journal of Manufacturing Systems*, vol. 56, pp. 296–311, 2020.

[7] A. Giret, E. Garcia, and V. Botti, "An engineering framework for service-oriented intelligent manufacturing systems," *Computers in Industry*, vol. 81, pp. 116–127, 2016.

[8] B. Kea, B. Slsa, and C. Az, "Prediction of disc cutter life during shield tunneling with AI via the incorporation of a genetic algorithm into a GMDH-type neural network," *Engineering*, vol. 7, no. 2, pp. 238–251, 2021.

[9] C. Wei, C. Az, and D. Slsc, "An analytical method for estimating horizontal transition probability matrix of coupled Markov chain for simulating geological uncertainty," *Computers and Geotechnics*, vol. 129, p. 2021.

[10] S. L. Shen, H. M. Lyu, and A. Zhou, "Automatic control of groundwater balance to combat dewatering during construction of a metro system," *Automation in Construction*, vol. 123, no. 5, 103536.

[11] S.-S. Lin, S.-L. Shen, A. Zhou, and Y.-S. Xu, "Approach based on TOPSIS and Monte Carlo simulation methods to evaluate lake eutrophication levels," *Water Research*, vol. 187, p. 116437, 2020.

[12] K. Zhang, H.-M. Lyu, S.-L. Shen, A. Zhou, and Z.-Y. Yin, "Evolutionary hybrid neural network approach to predict shield tunneling-induced ground settlements," *Tunnelling and Underground Space Technology*, vol. 106, p. 103594, 2020.

[13] A. Ssl, B. Slsa, and Z. B. Ning, "Modelling the performance of EPB shield tunnelling using machine and deep learning algorithms," *Geoscience Frontiers*, vol. 12, no. 5, Article ID 101177, 2021.

[14] J.-Q. Wang, J. Chen, Y. Zhang, and G. Q. Huang, "Schedule-based execution bottleneck identification in a job shop," *Computers & Industrial Engineering*, vol. 98, pp. 308–322, 2016.

[15] P. Samouei, P. Fattahi, and J. Ashayeri, "Bottleneck easing-based assignment of work and product mixture determination: fuzzy assembly line balancing approach," *Applied Mathematical Modelling*, vol. 40, no. 7, pp. 4323–4340, 2016.

[16] A. Azadeh and B. M. Shoja, "A neural network meta-model for identification of optimal combination of priority dispatching rules and makespan in a deterministic job shop scheduling problem," *International Journal of Advanced Manufacturing Technology*, vol. 67, no. 8, pp. 1549–1561, 2013.

[17] Z. Rui and W. Cheng, "Bottleneck machine identification method based on constraint transformation for job shop scheduling with genetic algorithm," *Information Sciences*, vol. 188, pp. 236–252, 2012.

[18] S. Bin and G. Sun, "Optimal energy resources allocation method of wireless sensor networks for intelligent railway systems," *Sensors*, vol. 20, no. 2, p. 482, 2020.

[19] P. Brucker, E. K. Burke, and S. Groenemeyer, "A mixed integer programming model for the cyclic job-shop problem with transportation," *Discrete Applied Mathematics*, vol. 160, no. 13, pp. 1924–1935, 2012.

[20] M. Masoud, E. Kozan, and G. Kent, "A job-shop scheduling approach for optimising sugarcane rail operations," *Flexible Services and Manufacturing Journal*, vol. 23, no. 2, pp. 181–206, 2011.

[21] M. Thürer and M. Stevenson, "On the beat of the drum: improving the flow shop performance of the Drum-Buffer-Rope scheduling mechanism," *International Journal of Production Research*, vol. 56, no. 9, pp. 3294–3305, 2018.

[22] K. D. Sweeney, D. C. Sweeney, and J. F. Campbell, "The performance of priority dispatching rules in a complex job shop: a study on the Upper Mississippi River," *International Journal of Production Economics*, vol. 216, pp. 154–172, 2019.

[23] K.-S. Chin, D.-W. Tang, J.-B. Yang, S. Y. Wong, and H. Wang, "Assessing new product development project risk by Bayesian network with a systematic probability generation methodology," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9879–9890, 2009.

[24] S. Bin, G. Sun, N. Cao et al., "Collaborative filtering recommendation algorithm based on multi-relationship social network," *Computers, Materials & Continua*, vol. 60, no. 2, pp. 659–674, 2019.

[25] G. Sun, C.-C. Chen, and S. Bin, "Study of cascading failure in multisubnet composite complex networks," *Symmetry*, vol. 13, no. 3, pp. 523–537, 2021.

[26] G. Tian, S. Zhou, G. Sun, and C.-C. Chen, "A novel intelligent recommendation algorithm based on mass diffusion," *Discrete Dynamics in Nature and Society*, vol. 2020, no. 11, pp. 1–9, Article ID 4568171, 2020.

[27] W. Fang, Y. Guo, and W. Liao, "A Parallel Gated Recurrent Units (P-GRUs) network for the shifting lateness bottleneck prediction in make-to-order production system," *Computers & Industrial Engineering*, vol. 140, no. 2, pp. 1–12, 2020.

[28] A. Yazdanbakhsh, B. Thwaites, H. Esmaeilzadeh, G. Pekhimenko, O. Mutlu, and T. C. Mowry, "Mitigating the memory bottleneck with approximate load value prediction," *IEEE Design & Test*, vol. 33, no. 1, pp. 32–42, 2016.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 5560277, 18 pages
https://doi.org/10.1155/2021/5560277

Research Article

Adaptive Reinforcement Learning-Enhanced Motion/Force Control Strategy for Multirobot Systems

Phuong Nam Dao, Duy Khanh Do, and Dinh Khue Nguyen

School of Electrical Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam

Correspondence should be addressed to Phuong Nam Dao; nam.daophuong@hust.edu.vn

Received 17 January 2021; Revised 14 March 2021; Accepted 4 May 2021; Published 1 June 2021

Academic Editor: Hussein Abulkasim

This paper presents an adaptive reinforcement learning- (ARL-) based motion/force tracking control scheme consisting of the optimal motion dynamic control law and force control scheme for multimanipulator systems. Specifically, a new additional term and appropriate state vector are employed in designing the ARL technique for time-varying dynamical systems with online actor/critic algorithm to be established by minimizing the squared Bellman error. Additionally, the force control law is designed after obtaining the computation of constraint force coefficient by the Moore–Penrose pseudo-inverse matrix. The tracking effectiveness of the ARL-based optimal control is verified in the closed-loop system by theoretical analysis. Finally, simulation studies are conducted on a system of three manipulators to validate the physical realization of the proposed optimal tracking control design.

1. Introduction

In recent years, the tracking control problem for robotic systems has been remarkable, considered from not only academia but also industrial automation [1, 2]. Among different schemes, robotic systems can mainly be separated into two categories: (1) mobile robotics [3, 4] and (2) robot manipulators [5]. However, because of the growing requirements for high productivity, the networked robotic systems that are known as the collaboration of many robots can fulfill different tasks by mobile manipulators [6], tractor trailers [3], bilateral teleoperators [5], and cooperating mobile manipulators (CMMs) [7]. It is worth emphasizing that, unlike nonlinear multiagent systems [8, 9] studying the relation between agents, CMMs consider the interaction with the rigid object. In the literature of CMMs' control objectives, they can be classified into two categories, i.e., multiple mobile robot manipulators in cooperation carrying a common object with unknown parameters and disturbances [7, 10–18]; one of them tightly holds the object by the end effector, and the remaining mobile manipulators' end effector follows a trajectory on the surface of the object [19, 20]. Most of the existing control literature of networked robotics mainly focuses on the implementation of nonlinear controllers. According to Lyapunov's direct technique, authors in [5] presented the robust controller for a bilateral teleoperation under time-varying delays and without relative motion. Recently, the backstepping technique of networked robotics has achieved much attention [3, 11, 21]. In [3], a backstepping method is utilized to implement an adaptive trajectory tracking control design for a tractor trailer, where the proposed control scheme can stabilize for cascade control systems. For the purpose of easily computing the backstepping method and handling the full-state constraint, authors in [21, 22] pointed out the appropriate term α in the asymmetric barrier Lyapunov function (ABLF). This backstepping method is extended in [11] to deal with unknown dead zone nonlinearity, time-varying term using the input-driven filter term and Nussbaum function, respectively. Another consideration addressed in the existing literature is to mention the holonomic-nonholonomic constraint [12, 19, 20, 23, 24]. Motivated by this approach, the corresponding motion dynamic model and constraint force equation are established for finding appropriate controllers [12, 19, 20, 23, 24]. As a result, the robust adaptive controller was realized by estimating the derivative of the Lyapunov

function and combining the constraint force control to obtain the motion/force control scheme for multimanipulator systems [12]. Due to dynamic uncertainties and external disturbance, fuzzy and neural network (NN) method are popularly considered to estimate dynamic uncertainties for guaranteeing system stability [7, 10, 20]. However, the work in [23] utilized an estimation in the absence of fuzzy and NNs for designing a decentralized controller for cooperating manipulators. In contrast to this technique, authors in [7, 12, 25] proposed the centralized control law based on the consideration of dynamics of m manipulators. The scheme in [7] extends the neural network-based centralized coordination control to achieve the tracking of each mobile manipulator by the consideration of additional actuators using servo motors. However, these nonlinear control solutions do not take into account the challenges, such as input-state saturation, and control objectives are different from the traditional tracking control problem. Therefore, we propose motion/force control with a progressive optimality principle to improve the tracking performance.

In the implementation of the optimality principle, an important aspect is developed for robotic systems with the consideration of the robot-environment interaction [26–28]. In control design for robotic systems, and so on, there always exists an interaction between the end-effector and the environment, which may lead to various challenges of system performance. Therefore, the control scheme should maintain a solution for considering the system interactions. Needless to say, the control systems of the robot-environment interaction have been paid much attention by many researchers for several decades [28]. Due to the effect of the interaction between the robot's end effector and the unknown environment, the desired trajectory in the joint space needs to be modified by the optimality approach [26–28]. In [26], the method to generate the virtual desired trajectory can be implemented by adaptive output feedback optimal control for the unknown impedance model. A different approach for finding the virtual reference trajectory can be developed by Q-learning [27]. Furthermore, based on ARL-based admittance control, authors in [28] proposed the completed control structure with the additional term of torque estimation and adaptive controller. An extension of the robot-environment interaction can be known as the physical human-robot interaction [15, 29]. In [29], the dynamic model of an n–dimensional manipulator system is established under the task space with the consideration of the end effector of the robot and human or environment. Moreover, the dynamic model established more a desired admittance model in the task space for obtaining virtual desired trajectory under constraint region by a soft saturation function [29]. A different discussion of the human-robot interaction with the estimation of human intention was mentioned in [15]. Furthermore, the motion/force control law was proposed using the computation of constraint force without the optimality problem [15]. An approach of optimization for the dual-arm manipulator with the consideration of relative Jacobian was mentioned in [16] by the QP solver. The extension from dual-arm manipulators to multiarm robots is discussed with the classical backstepping structure and optimal control law being constructed to tackle optimal behaviors involving obstacle avoidance and energy saving [11]. Additionally, authors in [17] studied multiarm systems under time-varying communication delay by task-space synchronization with a fixed threshold. For different networked systems of bilateral teleoperation systems, the optimality principle was also considered for implementing control design [30]. However, the disadvantage of the control law in [30] is no consideration in solving the Riccati equation. Furthermore, the optimality principle in robotic systems has been almost focused on the interaction of the human-robot's end effector and the environment. This article proposes the consideration of optimal control in trajectory tracking in the motion/force control problem.

For finding an optimal control problem to minimize a given performance index in dynamical systems, it is necessary to solve the so-called Hamilton–Jacobi–Bellman (HJB) equation obtained from the dynamic principle. However, it is hard to analytically solve this HJB equation due to leading to a nonlinear partial differential equation. Among the numerical methods which have been considered to solve the HJB equation, the remarkable iterative structure has been developed to solve via online based on the reinforcement learning (RL) principle being inspired by machine learning [27, 31–39]. In order to implement the numerical algorithm for solving the HJB equation, there are two major directions, including online actor/critic [31, 32] and off-policy technique using integral reinforcement learning (IRL) [38, 39]. The first approach using online actor/critic is to develop based on the property of Hamiltonian $H(x, u^*, (\partial V^*/\partial x))$ being 0. As a result, the iterative algorithm can be considered by the Bellman error with solving the optimization problem. For designing the ARL scheme in linear dynamical systems [31] and in nonlinear systems [32, 40], the methods are realized to find the optimal control input being Kronecker product and approximating neural networks (NNs), respectively. Furthermore, this technique is extended for several situations, such as goal representation heuristic dynamic programming (GRHDP) with the multivariable tracking scheme [35] and uncertain discrete-time systems by using NNs [33]. The second approach using the off-policy technique is implemented by considering the deviation of the performance index in two different times with appropriate data collection [38, 39]. It is worth emphasizing that almost existing works discussed the ARL design in time-invariant systems [27, 33, 35, 36, 38, 39]. In contrast, for establishing the optimal control problem in robotic systems, it requires the consideration of ARL in time-varying systems because of the desired time-varying trajectory [31, 32]. Furthermore, the ARL-based optimal control with the actor/critic learning structure has just been discussed in [41] for time-invariant systems. The key idea is to establish simultaneously the learning of neural networks (NNs) in the actor and critic part [41]. The work in [32] developed the actor/critic learning structure for manipulators by combining the sliding mode control (SMC) technique in the absence of constraint force. Regarding the multimanipulator systems, the proposed control law in [12]

only implemented the robust adaptive control based on the traditional nonlinear control technique and model separation technique without considering the optimality principle.

In this paper, we investigate the ARL-based motion/force tracking control problem for a multimanipulator system in the presence of disturbance. The control goal is to obtain the unification of the optimality principle and tracking problem. The main contributions of this paper can be summarized as follows:

(1) First, it is obviously different from [26, 27, 34, 41, 42] studying optimal control for time-invariant systems; we propose ARL-based optimal control design in the situation of the time-varying nonlinear dynamical system under the influence of the time-varying desired trajectory by utilizing the online actor/critic technique for the motion dynamic model of multimanipulator systems. Furthermore, this work is still proposed for the motion/force control problem by extending more the force control term in the control scheme.

(2) Second, different from numerous existing nonlinear control approaches [10, 12, 13, 17, 20, 23–25, 43, 44], the proposed optimal control algorithm is investigated with the unification of the optimization and tracking problem to be discussed through theoretical analysis and simulation studies.

The rest of this article is organized as follows. Section 2 provides the problem statement of this paper. Section 3 describes the ARL-based motion/force control design for multimanipulator systems with tracking analysis. Simulation studies on the multirobots with 3 manipulators are presented in Section 4, and the conclusion remarks are pointed out in Section 5.

2. Preliminaries and Problem Statement

We investigate n two-link manipulators, as shown in Figure 1, which includes n robots interacted by each pair of manipulators, the unknown environment with the original Cartesian coordinate system $O(X, Y, Z)$, and the corresponding coordinate E_i of the end effector of each robot. The interaction between each pair of manipulators leads to a constraint condition as described in the following equation:

$$\Xi_{ij} = \left|\mathrm{OE}_i - \mathrm{OE}_j\right| - d_{ij} = 0\,(\forall 1 \le i < j \le n), \tag{1}$$

where OE_i is a vector with the first point being the original point and the end effector E_i under the Cartesian coordinate system and d_{ij} is the distance between end effectors of the ith robot manipulator and the jth robot manipulator. According to (1), combining all the constraints in this multimanipulator, we can obtain the constraint vector as follows:

$$\Xi(\eta) = \left[\ldots, \Xi_{ij}\left(\eta_i, \eta_j\right), \ldots\right]^T = 0, \tag{2}$$

where $\Xi(\eta) \in \mathbb{R}^{m\times 1}$, η is the joint vector of n two-link manipulator systems. It can be seen that there are two main cases in general robots. First, there exist no constraint forces due to no interaction between the robot and the rigid object. Second, the interaction leads to the existence of constraint force, and we only consider the constraint motion under the following assumption.

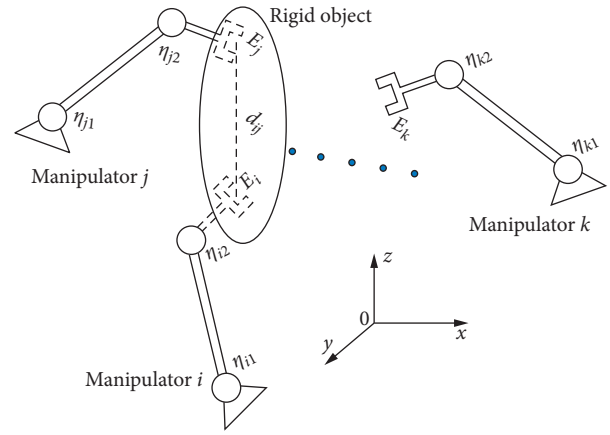


Figure 1: The cooperating manipulator system.

Assumption 1. The number of joints is more than the number of constraint conditions, i.e., $2n > m$.

In [12], the dynamic equation of n two-link robot manipulators with constraint force by the interaction between each couple of manipulators can be described as

$$M(\eta)\ddot{\eta} + C(\eta, \dot{\eta})\dot{\eta} + F(\eta, \dot{\eta}, t) = \tau - J^T(\eta)\lambda, \tag{3}$$

where $J(\eta) = (\partial\Xi/\partial\eta) \in \mathbb{R}^{m\times 2n}$ and $\lambda \in \mathbb{R}^{m\times 1}$ is the vector of Lagrange multiplier.

The control objective is to obtain not only the tracking of joints to desired trajectories but also guaranteeing the constant distance between end effectors of each pair of manipulators. Moreover, the motion/force control objective also mentions the tracking of Lagrangian coefficient λ in (3). On the contrary, due to the proposed optimality-based motion/force control design, it requires the unification of the optimal control and tracking problem.

Remark 1. It is worth emphasizing that different from classical nonlinear control approaches [7, 17, 19, 20, 23, 24, 43], the approach in this paper is developed by the optimal control technique for uncertain systems based on adaptive/approximate reinforcement learning (ARL). Moreover, regarding the mathematical model of multimanipulators, this work is the extension of Chen et al. [11] by considering the holonomic constraint force with separation of the joint vector.

3. Motion/Force Control Design with ARL

In this paper, a completed control structure in Figure 2 is investigated with 2 independent parts, including ARL-based optimal control for the motion dynamic model and the force control scheme for tracking the constraint coefficient. Moreover, the convergence and stability of the whole system

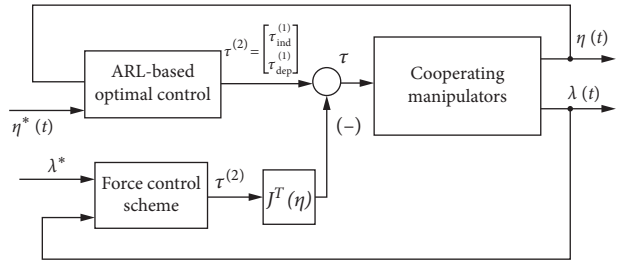


Figure 2: Control structure of the surface vessel.

(Figure 2) are discussed for determining the performance of the proposed motion/force controller.

3.1. ARL-Based Motion Controller. For the purpose of designing the motion controller in multimanipulator systems, the motion dynamic model is discussed after eliminating the constraint forces. Moreover, it is due to the description of time-varying systems in trajectory tracking control; the modification of the tracking error model needs to be implemented for developing the ARL-based optimal control design.

Because of $\Xi(\eta) \in \mathbb{R}^{m\times 1}$ with $m(m<2n)$ independent constraint conditions, there exists an algebraic map φ: $\mathbb{R}^{2n-m} \mapsto \mathbb{R}^{2n}$ satisfying

$$\eta = \varphi(\xi), \tag{4}$$

where $\xi \in \mathbb{R}^{(2n-m)\times 1}$ is the vector of independent joint variables. Taking the derivative of (4), we imply that

$$\begin{aligned} \dot{\eta} &= \Lambda\dot{\xi}, \\ \ddot{\eta} &= \Lambda\ddot{\xi} + \dot{\Lambda}\dot{\xi}, \end{aligned} \tag{5}$$

where $\Lambda = (\partial\varphi/\partial\xi) \in \mathbb{R}^{2n\times(2n-m)}$. With the purpose of eliminating the constraint force in the motion dynamic model, we use the Moore–Penrose pseudo-inverse matrix of $J(\eta)$ being $J^+ = J^T(JJ^T)^{-1}$ to achieve the matrix $Q = (I - J^+J)$ satisfying $QJ^T = 0$. Multiplying both sides of (3) with matrix Q, it leads to the following equation after utilizing $QJ^T = 0$:

$$QM\ddot{\eta} + QC\dot{\eta} + QF = Q\left(\tau - J^T\lambda\right) = Q\tau - QJ^T\lambda = Q\tau = \tau^1. \tag{6}$$

According to (5) and (6), the motion dynamic can be expressed as

$$QM\Lambda\ddot{\xi} + (QM\dot{\Lambda} + QC\Lambda)\dot{\xi} + QF = \tau^1. \tag{7}$$

Regarding the independent joint coordinates, the dynamic model can be represented after multiplying with matrix $[I_{2n-m}, 0] \in \mathbb{R}^{(2n-m)\times 2n}$ on both sides of (6) as follows:

$$\begin{cases} \dot{\xi} = \upsilon, \\ \overline{\overline{M}}\dot{\upsilon} + \overline{\overline{C}}\upsilon + \overline{\overline{F}} = \tau_{\text{ind}}^1, \end{cases} \tag{8}$$

where

$$\begin{aligned} \overline{\overline{M}} &= [I_{2n-m}, 0]QM\Lambda, \\ \overline{\overline{C}} &= [I_{2n-m}, 0](QM\dot{\Lambda} + QC\Lambda), \\ \overline{\overline{F}} &= [I_{2n-m}, 0]QF, \\ \tau_{\text{ind}}^1 &= [I_{2n-m}, 0]\tau^1. \end{aligned} \tag{9}$$

Remark 2. This paper only discusses the situation that the number of constraint conditions is smaller than the number of joints $(2n > m)$ as mentioned in Assumption 1. For the case of $(2n = m)$, it leads to no existence of independent joints. As a result, the above modifications are not implemented, and the control method is developed with different approaches [45]. Additionally, no constraint forces are considered in the special case $(m = 0)$.

The motion control design is proposed as

$$\tau_{\text{ind}}^1 = \tau_d + u, \tag{10}$$

where

$$\begin{aligned} \tau_d &= \overline{\overline{M}}\dot{\upsilon}_d + \overline{\overline{C}}\upsilon_d + \overline{\overline{F}}, \\ \upsilon_d &= \left(\dot{\xi}_d - \beta z_\xi\right), \\ z_\xi &= (\xi - \xi_d), \\ z_\upsilon &= (\upsilon - \upsilon_d). \end{aligned} \tag{11}$$

Therefore, with the assumption that $(d/dt)\xi_d$ is the function of ξ_d as $(d/dt)\xi_d = h_1(\xi_d)$, we can obtain that

$$\begin{cases} \dfrac{d}{dt}z_\upsilon = \dfrac{d}{dt}(\upsilon - \upsilon_d) = \overline{\overline{M}}^{-1}u - \overline{\overline{M}}^{-1}\overline{\overline{C}}(\upsilon - \upsilon_d), \dfrac{d}{dt}z_\xi = \upsilon - \dfrac{d}{dt}\xi_d, \dfrac{d}{dt}\xi_d = h_1(\xi_d). \end{cases} \tag{12}$$

Furthermore, υ_d is the function of (z_ξ, ξ_d) as $\upsilon_d = (\dot{\xi}_d - \beta z_\xi) = h_1(\xi_d) - \beta z_\xi \Rightarrow \upsilon = (z_\upsilon + \upsilon_d)$ is the function of $(z_\upsilon, z_\xi, \xi_d)$. Combining with (12), a continuous-time affine system has been obtained with the typical form

$$\dot{Z} = F(Z) + G(Z)u, \tag{13}$$

where the vector of state variables is $Z = [z_v, z_\xi, \xi_d]^T \in \mathbb{R}^{12}$.

Remark 3. Due to the time-varying reference in trajectory tracking, the closed system needs to be known as the time-varying property. However, the framework of using term (11) and choosing the state variable vector Z is able to obtain time-invariant systems (12) and (13). Therefore, in contrast to the existing methods in [36], the proposed control easily handles the situation of time-varying systems by the elimination of the term $-(\partial V^*/\partial t)$.

3.1.1. The Optimal Control Problem. In this section, we consider the general class of time-varying systems:

$$\frac{d}{dt}Z = F(Z, u, t), \tag{14}$$

with the associated cost function to be defined as

$$J(Z, u) = \int_0^\infty r(Z(\tau), u(\tau))dt, \tag{15}$$

where $r(Z(\tau), u(\tau)) = (Z^T Q Z + u^T \mathrm{R}u)$ under the weighting matrices $Q = Q^T \geq 0$ and $R = R^T > 0$ to be defined with appropriate dimensions. It can be seen that the optimal problem for the motion model of multimanipulators (12) and (13) is to realize an admissible control policy [31, 32] for the purpose of obtaining the minimal value. This is because that the control policy $\widehat{u}(Z)$ needs to guarantee the existence of the estimated Bellman function $\widehat{V}(Z)$ obtained from the HJB equation.

Definition 1. A control law $u(Z)$ is considered being "admissible" with respect to the performance index (15) on a compact set Ω, which is known as a continuous signal $u(Z) \in \Upsilon(\Omega)$, if the properties of the stabilization of (14) and the limitation of $J(Z, u)$ for every $Z \in \Omega$ are satisfied.

It is worth emphasizing that, in the general case of time-varying system (14), the optimal control strategy needs to be considered as a time-varying function $u^*(Z, t)$. Hence, the time-varying Bellman function with respect to arbitrary time $V^*(Z(t), t)$ can be given as

$$V^*(Z(t)) = \min_{u(Z)\in\Upsilon(\Omega)} J(Z(t), u(t)). \tag{16}$$

The HJB equation can be achieved by taking the time derivative of $V^*(X(t))$ under two different methods. Firstly, it is directly computed along system (14) as

$$\frac{d}{dt}V^*(Z(t), t) = \frac{\partial V^*}{\partial Z}\frac{dZ}{dt} + \frac{\partial V^*}{\partial t} = \frac{\partial V^*}{\partial Z}F(Z, u^*, t) + \frac{\partial V^*}{\partial t}. \tag{17}$$

Secondly, thanks to the dynamic programming principle, we also compute the derivative of $V^*(Z(t))$ with respect to time as

$$\frac{d}{dt}V^*(Z(t), t) = -r(Z(\tau), u^*(\tau)). \tag{18}$$

According to (17) and (18), the first part of the HJB problem can be given:

$$r(Z(\tau), u^*(\tau)) + \frac{\partial V^*}{\partial Z}F(Z, u^*, t) = -\frac{\partial V^*}{\partial t}. \tag{19}$$

Moreover, the second part of the HJB problem can be developed with the performance index to be formulated as

$$V^*(Z(t), t) = \min_{u(Z)\in\Upsilon(\Omega)} \int_t^\infty r(Z(\tau), u^*(\tau))d\tau, \tag{20}$$

$$V^*(Z(t), t) = \min_{u^*(Z)\in\Upsilon(\Omega)} \int_t^{t+\Delta t} r(Z, u^*)d\tau + \min_{u^*(Z)\in\Upsilon(\Omega)} \int_{t+\Delta t}^\infty r(Z, u^*)d\tau. \tag{21}$$

Because of utilizing the dynamic programming principle, the performance index is represented as

$$V^*(Z(t), t) = \min_{u(Z)\in\Upsilon(\Omega)} \int_t^{t+\Delta t} r(Z(\tau), u(\tau))d\tau + V^*(Z(t + \Delta t), t + \Delta t). \tag{22}$$

Therefore, we imply that the Bellman function can be rewritten as

$$0 = \min_{u(Z)\in\Upsilon(\Omega)} \left\{ \frac{1}{\Delta t} \int_t^{t+\Delta t} r(Z, u^*)d\tau + \frac{V^*(Z(t + \Delta t), t + \Delta t) - V^*(Z(t), t)}{\Delta t} \right\}. \tag{23}$$

As the convergence of $\Delta t \longrightarrow 0^+$, one can derive that

$$\min_{u(Z)\in\Upsilon(\Omega)}\left\{r\left(Z(t),u^*(t)\right)+\frac{\partial V^*}{\partial Z}F\left(Z,u^*,t\right)\right\}=-\frac{\partial V^*}{\partial t}. \tag{24}$$

Remark 4. It is worth noting here that the term $(\partial/\partial t)V^*(Z(t),t)$ in (19) and (24) is the disadvantage in developing the ARL algorithm, and there has been very little research considering the direct method for this challenge. Authors in [36] developed the solutions of the ARL-based time-varying optimal controller using the framework of the data-driven and Newton–Leibniz formula-based function approximation method. In this paper, we indirectly handle the time-varying tracking error model through the consideration of systems (12) and (13). Therefore, the two parts of the HJB problem can be modified as

$$r\left(Z,u^*\right)+\frac{\partial V^*}{\partial Z}\left(F(Z)+G(Z)u^*\right)=0, \tag{25}$$

$$\min_{u(X)\in\Upsilon(\Omega)}\left[r(X,u)+\frac{\partial V^*}{\partial X}\left(F(X)+G(X)u\right)\right]=0. \tag{26}$$

3.1.2. ARL-Based Control Design for Independent Joints. It is because of difficulties in directly solving HJB equation (25) to obtain corresponding optimal control algorithm. Hence, the approximate solution with a neural network (NN) is mentioned to develop the ARL design in the motion dynamic model of multimanipulators. Thanks to the smooth property of functions $V^*(Z)$ and $u^*(Z)$, they can be described over any compact domain $C\subseteq\mathbb{R}^{12}$:

$$V^*(Z)=W^T\psi(Z)+\epsilon(Z), \tag{27}$$

$$u^*(Z)=-\frac{1}{2}R^{-1}G^T(Z)\left(\left(\frac{\partial\psi}{\partial Z}\right)^T W+\left(\frac{\partial\varepsilon(Z)}{\partial Z}\right)^T\right), \tag{28}$$

where $W\in\mathbb{R}^N$ is an uncertain ideal NN weight vector, N is chosen as the number of neurons, $\psi(Z)\in\mathbb{R}^N$ denotes an activation function vector satisfying $\psi_j(0)=0$ and $(\partial\psi_j/\partial Z)|_{Z=0}=0, \forall j=1,\ldots,N$, and $\epsilon(Z)\in\mathbb{R}$ is known as the estimation error of the Bellman function $V^*(Z)$. Due to the unknown ideal NN weight vector $W\in\mathbb{R}^N$, it is necessary to obtain the corresponding adjusting mechanism $\widehat{W}_a, \widehat{W}_c$ for estimating the actor/critic NNs to achieve the optimal control design in the absence of finding the HJB equation analytically (for more details, see [41]). Moreover, it is able to choose the smooth NN activation function vector $\psi(Z)\in\mathbb{R}^N$ based on the description of multimanipulators (see Section 4). In [41], thanks to the Weierstrass approximation theorem, we can uniformly approximate not only $V^*(Z)$ but also $(\partial V^*(Z)/\partial Z)$ with $\epsilon(Z)$, $(\partial\epsilon(Z)/\partial Z)\longrightarrow 0$ as $N\longrightarrow\infty$. For a fixed number N, the approximated Bellman function of critic NN $\widehat{V}(Z)$ and the approximated optimal control policy of actor NN $\widehat{u}(X)$ are represented as

$$\widehat{V}(Z)=\widehat{W}_c^T\psi(Z), \tag{29}$$

$$\widehat{u}(Z)=-\frac{1}{2}R^{-1}G^T(Z)\left(\frac{\partial\psi}{\partial Z}\right)^T\widehat{W}_a. \tag{30}$$

It should be noted that properties (19) and (24) of Hamiltonian $H(Z,u,(\partial V/\partial Z))=r(Z(\tau),u(\tau))+(\partial V/\partial Z)(F(Z)+G(Z)u)$ with the optimal controller $u^*(Z)$ and corresponding value function $V^*(Z)$ guarantee that the adjusting mechanisms of actor NN $\widehat{W}_a$ and critic $\widehat{W}_c$ weights are simultaneously tuned to minimize the squared Bellman error δ_{hjb} and the corresponding integral, respectively. According to the property of HJB problem (25), it leads to $H^*(Z,u^*,(\partial V^*/\partial Z))=0$. After obtaining the error between approximated functions $\widehat{V}(Z),\widehat{u}(Z)$ and optimal control input $V^*(Z),u^*(Z)$, the Bellman error δ_{hjb} is represented as

$$\delta_{hjb}=\widehat{H}\left(Z,\widehat{u},\frac{\partial\widehat{V}}{\partial Z}\right)-H^*\left(Z,u^*,\frac{\partial V^*}{\partial Z}\right)=\widehat{H}\left(Z,\widehat{u},\frac{\partial\widehat{V}}{\partial Z}\right)=\widehat{W}_c^T\sigma(Z,\widehat{u})+Z^TQ_TZ+\widehat{u}^TR\widehat{u}, \tag{31}$$

where $\sigma(Z,\widehat{u})=(\partial\psi/\partial Z)(F(Z)+G(Z)\widehat{u})$ is the regression vector of the critic part.

Based on the results in [31], the adjusting mechanism of the critic NN is computed:

$$\frac{d}{dt}\widehat{W}_c=-k_c\lambda\frac{\sigma}{1+\nu\sigma^T\lambda\sigma}\delta_{hjb}, \tag{32}$$

where $\nu,k_c\in\mathbb{R}$ are constant positive coefficients and $\lambda(t)\in\mathbb{R}^{N\times N}$ is an estimated symmetric matrix being the solution of the differential equation as

$$\begin{aligned}\frac{d}{dt}\lambda&=-k_c\lambda\frac{\sigma\sigma^T}{1+\nu\sigma^T\lambda\sigma}\lambda,\\ \lambda\left(t_s^+\right)&=\lambda(0)=\varphi_0 I,\end{aligned} \tag{33}$$

where t_s^+ denotes the resetting time satisfying that $\lambda_{\min}\{\lambda(t)\}\le\varphi_1, \varphi_0>\varphi_1>0$. In [31], because of the condition that $\lambda(t)$ is positive definite and preventing the covariance wind-up problem, the covariance matrix $\lambda(t)$ is chosen as

$$\varphi_1 I\le\lambda(t)\le\varphi_0 I. \tag{34}$$

Additionally, the adjusting mechanism of the actor NN part is represented by minimizing the squared Bellman error:

$$\frac{d}{dt}\widehat{W}_a = -\frac{k_{a1}}{\sqrt{1+\sigma^T\sigma}}\frac{\partial\psi}{\partial Z}\mathrm{GR}^{-1}G^T\frac{\partial\psi^T}{\partial Z}\left(\widehat{W}_a-\widehat{W}_c\right)\delta_{hjb} - k_{a2}\left(\widehat{W}_a-\widehat{W}_c\right) \tag{35}$$

Remark 5. ARL technique has been investigated with many different approaches, such as off-policy integral reinforcement learning (IRL) [46, 47] and Q-learning [42]. However, there is a wide difference between the online actor/critic technique in this paper and IRL and Q-learning. The off-policy IRL requires the complete data collection between the two sequential sampling times in computing the optimal law [46, 47]. Meanwhile, it is obviously different from the off-policy IRL technique; the online actor/critic method in (29), (30), (35), (33) requires the initial values to implement the computation. This could be exemplified by this article design control law under HJB problems (25) and (26), and the off-policy IRL method considers the relation between the deviation of integrals at two sampling times and the collected data under the dynamic programming law. In addition, the Q-learning method mentioned the Q-function in terms of not only state variables but also control inputs, being different from the Bellman function [42]. This property can result in the difficulties for implementing the Q-learning technique of the continuous-time system. Additionally, compared with the data-driven method in [48], this work provides a neural network-based technique to avoid the Kronecker product in estimating the actor/critic term. The actor/critic-based approaches have been discussed in [49, 50] for nonlinear affine systems using residual error δ_{hjb}. However, in view of the consideration of the identifier in [49, 50], it implies the difference in the computation of residual error δ_{hjb} and training laws in actor/critic weights between the proposed method and the work in [49, 50].

3.1.3. The Control Design of Dependent Joints. After completing the tracking control design of independent joints, we can achieve the remaining dependent joints' controller based on the relation between two groups of joint variables, including independent joints and dependent joints. According to (8), it leads to the following representation:

$$\ddot{\xi} = \overline{\overline{M}}^{-1}\left(\tau^1_{\mathrm{ind}} - \overline{\overline{C}}\xi - \overline{\overline{F}}\right). \tag{36}$$

On the contrary, thanks to the relation $\tau^1 = [\tau^{1T}_{\mathrm{ind}}, \tau^{1T}_{\mathrm{dep}}]^T$, multiplying both sides of (6) with $[0, I_m] \in \mathbb{R}^{(m)\times 2n}$, we achieve the dependent joints to be computed as

$$M''\ddot{\xi} + C''\dot{\xi} + F'' = \tau^1_{\mathrm{dep}}, \tag{37}$$

where

$$\begin{aligned} M'' &= [0, I_m]QM\Lambda, \\ C'' &= [0, I_m](QM\dot{\Lambda} + QC\Lambda), \\ F'' &= [0, I_m]QF, \\ \tau^1_{\mathrm{dep}} &= [0, I_m]\tau^1. \end{aligned} \tag{38}$$

Substituting (36) into (37), we obtain the control law of dependent joint variables to be implemented as follows:

$$\tau^1_{\mathrm{dep}} = M''\left\{\overline{\overline{M}}\left(\tau^1_{\mathrm{ind}} - \overline{\overline{C}}\xi - \overline{\overline{F}}\right)\right\} + C''\dot{\xi} + F''. \tag{39}$$

It is worth noting that the problem of synchronization between multiple joints can be considered by the relation between dependent and independent joints. Therefore, this problem is able to obtain the desired trajectories of independent joints and the trajectories of remaining joints depending on these independent joints. Furthermore, it leads to the relationship between the control input of independent joints and dependent joints (39).

3.2. Constraint Force Control Design. For the purpose of designing the motion/force control scheme for n manipulators, the constraint force controller and motion control law are necessary to be established based on the computation of lambda coefficient in (3). Before demonstrating the motion/force control design, we need to compute the coefficient λ to be dependent on control input τ as follows. According to (3)–(5), we achieve the following equation:

$$J^{*T}(\xi)\lambda = \tau - \left\{M^*(\eta)\ddot{\xi} + C^*(\eta,\dot{\eta})\dot{\xi} + F^*(\eta,\dot{\eta},t)\right\}, \tag{40}$$

where

$$\begin{aligned} J^{*T}(\xi) &= J^T(\varphi(\xi)), \\ M^*(\eta) &= M(\eta)\Lambda, \\ C^*(\eta,\dot{\xi}) &= C(\eta,\dot{\eta})\Lambda + M(\eta)\dot{\Lambda}, \\ F^*(\eta,\dot{\eta},t) &= F(\eta,\dot{\eta},t). \end{aligned} \tag{41}$$

Moreover, we find the method to compute coefficient λ after eliminating the term ξ by multiplying both sides with matrix N satisfying that

$$\begin{cases} \mathrm{NM}(\eta)\Lambda = 0, \\ \mathrm{NJ}^{*T}(\xi) = I. \end{cases} \tag{42}$$

Therefore, we can obtain the following coefficient λ:

$$\lambda = N\tau - \left\{\mathrm{NC}^*(\eta,\dot{\eta})\dot{\xi} + \mathrm{NF}^*(\eta,\dot{\eta},t)\right\}, \tag{43}$$

with N to be chosen as

$$N = \left(\mathrm{AJ}^{*T}(\xi)\right)^{-1}A, \tag{44}$$

where $A = J^*(\xi)M^{-1}(\eta)$.

Based on the proposed motion/force, control design is described in the following equation:

$$\tau = \begin{bmatrix} \tau^{(1)}_{\mathrm{ind}} \\ \tau^{(1)}_{\mathrm{dep}} \end{bmatrix} - J^T(\eta)\tau^{(2)}. \tag{45}$$

This leads to the following relations that

$$Q\tau^{(1)} = Q\begin{bmatrix}\tau_{\text{ind}}^{(1)}\\ \tau_{\text{dep}}^{(1)}\end{bmatrix} = \tau^{1}, \tag{46}$$

$$\lambda = N\tau^{(1)} - \tau^{(2)} - \left\{\text{NC}^{*}(\eta,\dot{\eta})\dot{\xi} + \text{NF}^{*}(\eta,\dot{\eta},t)\right\}. \tag{47}$$

It should be noted that the constraint force coefficient λ depends on the ARL-based motion dynamic control $\tau^{(1)} = \begin{bmatrix}\tau_{\text{ind}}^{(1)}\\ \tau_{\text{dep}}^{(1)}\end{bmatrix}$ to be described in (47). Therefore, with the purpose of tracking the coefficient λ, we can propose the remaining term $\tau^{(2)}$ in (47) as given in the following:

$$\tau^{(2)} = K_f e_\lambda - \lambda_d - \text{NC}^{*}(\eta,\ddot{\eta})\dot{\xi}_d. \tag{48}$$

The tracking problem will be proved in the next section.

3.3. Convergence and Stability Analysis. With the purpose of considering the stability and convergence of the closed system under the proposed motion-force control design, we need to implement several following steps. First, the Bellman error δ_{hjb}, critic estimated weight error $\widetilde{W}_c = W - \widehat{W}_c$, and the persistence of excitation (PE) condition need to be given. Second, thanks to the optimal Bellman function and PE condition, the corresponding Lyapunov function candidate is chosen. Third, the stability of the whole system (Figure 1) and the role of constraint force control and dependent joints' motion control are discussed.

According to (27)–(31), the Bellman error function δ_{hjb} is dependent on state variable vector $Z(t)$ as

$$\begin{aligned}\delta_{\text{hjb}} &= \widehat{W}_c^{T}\sigma(Z,\widehat{u}) + Z^{T}Q_T Z + \widehat{u}^{T}R\widehat{u} - W^{T}\sigma(Z,u^{*}) - u^{*T}Ru^{*}\\ &\quad - \frac{\partial\epsilon}{\partial Z}\left(F(Z) + G(Z)u^{*}\right) - Z^{T}Q_T Z\\ &= -\widetilde{W}_c^{T}\sigma(Z,\widehat{u}) + \frac{1}{4}\widetilde{W}_a^{T}\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\psi}{\partial Z}\right)^{T}\widetilde{W}_a\\ &\quad - \frac{1}{4}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\epsilon}{\partial Z}\right)^{T} - \frac{\partial\epsilon}{\partial Z}\left(F(Z) + G(Z)u^{*}\right).\end{aligned} \tag{49}$$

Combining (49) with (32), we imply the dynamic equation of critic error $\widetilde{W}_c = W - \widehat{W}_c$ to be given as

$$\begin{aligned}\frac{d}{dt}\widetilde{W}_c &= -k_c\lambda\Pi\Pi^{T}\widetilde{W}_c + k_c\lambda\frac{\sigma(Z,\widehat{u})}{1+\nu\sigma^{T}\lambda\sigma}\left(\frac{1}{4}\widetilde{W}_a^{T}\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\psi}{\partial Z}\right)^{T}\widetilde{W}_a\right.\\ &\quad \left. -\frac{1}{4}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\epsilon}{\partial Z}\right)^{T} - \frac{\partial\epsilon}{\partial Z}\left(F(Z) + G(Z)u^{*}\right)\right),\end{aligned} \tag{50}$$

where $\widetilde{W}_a = W - \widehat{W}_a$ and $\Pi(t) = (\sigma(Z,\widehat{u})/\sqrt{1+\nu\sigma(Z,\widehat{u})^{T}\lambda\sigma(Z,\widehat{u})})$ is satisfied by

$$\|\Pi\| \le \frac{1}{\sqrt{\nu\varphi_1}}. \tag{51}$$

The nominal system is given after eliminating the term of NN errors:

$$\frac{d}{dt}\widetilde{W}_c = -k_c\lambda\Pi\Pi^{T}\widetilde{W}_c. \tag{52}$$

Due to the purpose of exponential convergence of $\widetilde{W}_c$, it is necessary to guarantee the PE condition as follows:

$$\mu_2 I \ge \int_{t_0}^{t_0+\delta}\Pi(s)\Pi(s)^{T}ds \ge \mu_1 I, \quad \forall t_0 \ge 0. \tag{53}$$

The following main theorem is proposed with the assumptions being discussed in [41] to estimate the attraction region of tracking errors.

Theorem 1. *Consider the cooperating manipulators in Figure 1 with the assumption of the bound conditions of the NN including ideal weights, activation function $\psi(.)$, its derivative to be described in [41], and PE condition of the signal vector $\Pi(t)$, and the selection of parameters is implemented with the following condition:*

$$\frac{c_3}{k_{a1}} > k_1 k_2, \tag{54}$$

where the coefficients in (54) are mentioned in (55), (56), and (35). Let us consider the control structure (Figure 2), (10) with ARL-based control scheme (46), (10), (39), and (30), the associated adjusting mechanisms (35) and (33) for the actual controller, and constraint force control vector (48); then, (1) the actor-critic weight errors $\widetilde{W}_a$ and $\widetilde{W}_c$ are UUB; (2) the tracking effectiveness of not only $\widetilde{z}_v = (v - v_d)$ but also $\widetilde{z}_\eta = \xi - \xi_d$ in the closed control system (Figure 1) is also UUB; (3) the tracking of constraint force coefficient vector λ and the remaining terms of joint variables' vector $\eta_{2n} = [\xi_{(2n-m)}, p_m]$ is also UUB.

Proof. First, the tracking effectiveness of motion tracking control is considered under the ARL-based optimal control scheme for corresponding tracking error motion models (12) and (13). Due to the satisfaction of PE condition (53), there exists a time-varying function V_c: $\mathbb{R}^N \times [0,\infty) \longrightarrow \mathbb{R}$ satisfying several following inequalities:

$$\begin{aligned}&c_1\left\|\widetilde{W}_c\right\|^2 \le V_c(\widetilde{W}_c,t) \le c_2\left\|\widetilde{W}_c\right\|^2,\\ &\frac{\partial V_c}{\partial t} + \frac{\partial V_c}{\partial\widetilde{W}_c}\left(-k_c\lambda\Upsilon\Upsilon^{T}\widetilde{W}_c\right) \le -c_3\left\|\widetilde{W}_c\right\|^2,\\ &\left\|\frac{\partial V_c}{\partial\widetilde{W}_c}\right\| \le c_4\left\|\widetilde{W}_c\right\|,\end{aligned} \tag{55}$$

where $c_1, c_2, c_3, c_4 \in \mathbb{R}$ are positive constant coefficients. The function $V_c(\widetilde{W}_c,t)$ is one part of the Lyapunov function candidate. Moreover, according to the bound conditions of the proposed NN [41], we can achieve the following results

with the purpose of estimating the time derivative of the Lyapunov function in the next steps:

$$
\begin{aligned}
i.&\left\|\widetilde{W}_a\right\| \le k_1,\\
i^2.&\left\|\frac{\partial\psi}{\partial Z}GR^{-1}G^T\left(\frac{\partial\psi}{\partial Z}\right)\right\| \le k_2,\\
i^3.&\left\|\frac{1}{4}\widetilde{W}_a^T\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\psi}{\partial Z}\right)^T\widetilde{W}_a - \frac{1}{4}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\epsilon}{\partial Z}\right)^T - \frac{\partial\epsilon}{\partial Z}\left(F(Z)+G(Z)u^*\right)\right\| \le k_3,\\
i^4.&\left\|\frac{1}{2}W^T\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\epsilon}{\partial Z}\right)^T + \frac{1}{2}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\epsilon}{\partial Z}\right)^T\right.\\
&\left.+\frac{1}{2}W^T\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\psi}{\partial Z}\right)^T\widetilde{W}_a + \frac{1}{2}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\psi}{\partial Z}\right)^T\widetilde{W}_a\right\| \le k_4.
\end{aligned}
\tag{56}
$$

Thanks to Bellman function $V^*(Z)$ (20) and time-varying $V_c(\widetilde{W}_c,t)$ (55), a Lyapunov function candidate can be chosen to determine the stability of the whole of the cascade control system (Figure 2) and the tracking of the weights of the actor NN and critic NN:

$$
V_L \triangleq \frac{1}{2}\rho z_\xi^T z_\xi + V^*(Z) + V_c\left(\widetilde{W}_c,t\right) + \frac{1}{2}\widetilde{W}_a^T\widetilde{W}_a. \tag{57}
$$

It is worth noting that the term $(1/2)\rho z_\xi^T z_\xi$ (ρ is a positive constant coefficient) is the additional term for the purpose of tracking of the whole motion control system. Because $V^*(Z)$ is a smooth function and positive definite, there exist two K-class functions γ_1 and γ_2 such that

$$
\gamma_1(\|Z\|) \le V^*(Z) \le \gamma_2(\|Z\|). \tag{58}
$$

Because of (55) and (8), we can determine the strict proper representation of Lyapunov function $V_L(Z)$:

$$
\begin{aligned}
\frac{1}{2}\rho\|z_\xi\|^2 + \alpha_1(\|Z\|) + c_1\|\widetilde{W}_c\|^2 + \frac{1}{2}\|\widetilde{W}_a\|^2 \le V_L\\
\le \frac{1}{2}\rho\|z_\xi\|^2 + \alpha_2(\|Z\|)\\
+ c_2\|\widetilde{W}_c\|^2 + \frac{1}{2}\|\widetilde{W}_a\|^2.
\end{aligned}
\tag{59}
$$

Implementing the derivative of V_L along motion system (13) under the control input $\widehat{u}(Z)$, we imply that

$$
\begin{aligned}
\frac{d}{dt}V_L &= \rho z_\xi^T\left(z_v - \beta z_\xi\right) + \frac{\partial V^*}{\partial Z}\left(F(Z)+G(Z)\widehat{u}\right)\\
&+ \frac{\partial V_c}{\partial t} + \frac{\partial V_c}{\partial\widetilde{W}_c}\left(\Theta_1+\Theta_2\right)\\
&- \widetilde{W}_a^T\dot{\widehat{W}}_a + \frac{\partial V^*}{\partial Z}G(Z)\Delta,
\end{aligned}
\tag{60}
$$

where

$$
\begin{aligned}
\Theta_1 &= -k_c\lambda\Pi\Pi^T\widetilde{W}_c\\
\Theta_2 &= k_c\lambda\frac{\sigma}{1+\nu\sigma^T\lambda\sigma}\left(\frac{1}{4}\widetilde{W}_a^T\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\psi}{\partial Z}\right)^T\right.\\
&\left.\widetilde{W}_a - \frac{1}{4}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^T\left(\frac{\partial\epsilon}{\partial Z}\right)^T - \frac{\partial\epsilon}{\partial Z}\left(F(Z)+G(Z)u^*\right)\right).
\end{aligned}
\tag{61}
$$

According to (19), the relation between the Bellman function and the optimal control law can be represented as

$$
\frac{\partial V^*}{\partial Z}F(Z) = -\frac{\partial V^*}{\partial Z}G(Z)u^*(Z) - Z^TQ_TZ - u^{*T}(Z)\mathrm{R}u^*(Z). \tag{62}
$$

According to (62) and optimal law (28), (35), and (55) in (60), we can conclude that

$$
\begin{aligned}
\frac{d}{dt}V_L \le &-\rho\beta\|z_\xi\|^2 + \rho z_\xi^T z_v - z_v^TQz_v - u^{*T}\mathrm{R}u^*\\
&+ 2u^{*T}R\left(u^*-\widehat{u}\right) - c_3\|\widetilde{W}_c\|^2 + c_4\|\widetilde{W}_c\|\|\Theta_2\| +\\
&k_{a2}\widetilde{W}_a^T\left(\widehat{W}_a-\widehat{W}_c\right) - 2u^{*T}R\Delta + \frac{k_{a1}}{\sqrt{1+\sigma^T\sigma}}\widetilde{W}_a^T\frac{\partial\psi}{\partial Z}\\
&G(Z)R^{-1}G^T(Z)\left(\frac{\partial\psi}{\partial Z}\right)^T\left(\widehat{W}_a-\widehat{W}_c\right)\delta_{\mathrm{hjb}}.
\end{aligned}
\tag{63}
$$

Using Young's inequality, one can obtain that

$$
\rho z_\xi^T z_v \le \frac{\rho}{2}\|z_\xi\|^2 + \frac{\rho}{2}z_v^Tz_v = \frac{\rho}{2}\|z_\xi\|^2 + \frac{\rho}{2}\|z_v\|^2. \tag{64}
$$

From (30), (19), and (56), the following inequality is given:

$$2u^{*T}R(u^{*}-\hat{u})=\frac{1}{2}W^{T}\frac{\partial\phi}{\partial Z}GR^{-1}G^{T}\left(\frac{\partial\epsilon}{\partial Z}\right)^{T}+\frac{1}{2}W^{T}\frac{\partial\phi}{\partial Z}GR^{-1}G^{T}\left(\frac{\partial\phi}{\partial Z}\right)^{T}\widetilde{W}_{a}+\frac{1}{2}\frac{\partial\epsilon}{\partial Z}GR^{-1}G^{T}\left(\frac{\partial\phi}{\partial Z}\right)^{T}\widetilde{W}_{a}+\frac{1}{2}\frac{\partial\epsilon}{\partial Z}GR^{-1}G^{T}\left(\frac{\partial\epsilon}{\partial Z}\right)^{T}\leq k_{4}, \tag{65}$$

and the term Θ_2 is bounded by

$$\left\|\Theta_{2}\right\|\leq\frac{k_{c}\varphi_{0}}{2\sqrt{\nu\varphi_{1}}}k_{3}. \tag{66}$$

Following (56) and (49), we have

$$\frac{k_{a1}}{\sqrt{1+\sigma^{T}\sigma}}\widetilde{W}_{a}^{T}\frac{\partial\psi}{\partial Z}G(Z)R^{-1}G^{T}(Z)\left(\frac{\partial\psi}{\partial Z}\right)^{T}\left(\widehat{W}_{a}-\widehat{W}_{c}\right)\delta_{\mathrm{hjb}}=\frac{k_{a1}}{\sqrt{1+\sigma^{T}\sigma}}\widetilde{W}_{a}^{T}\frac{\partial\psi}{\partial Z}G(Z)R^{-1}G^{T}(Z)\left(\frac{\partial\psi}{\partial Z}\right)^{T}(\widetilde{W}_{c}-\widetilde{W}_{a})\times\left(-\widetilde{W}_{c}^{T}\sigma+\frac{1}{4}\widetilde{W}_{a}^{T}\left(\frac{\partial\psi}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\psi}{\partial Z}\right)^{T}\widetilde{W}_{a}-\frac{1}{4}\left(\frac{\partial\epsilon}{\partial Z}\right)GR^{-1}G^{T}\left(\frac{\partial\epsilon}{\partial Z}\right)^{T}-\frac{\partial\epsilon}{\partial Z}(F(Z)+G(Z)u^{*})\right)\leq k_{a1}k_{1}k_{2}\left\|\widetilde{W}_{c}\right\|^{2}+k_{a1}k_{1}^{2}k_{2}\left\|\widetilde{W}_{c}\right\|+k_{a1}k_{1}k_{2}k_{3}\left\|\widetilde{W}_{c}\right\|+k_{a1}k_{1}^{2}k_{2}k_{3}. \tag{67}$$

It is obvious that

$$k_{a2}\widetilde{W}_{a}^{T}\left(\widehat{W}_{a}-\widehat{W}_{c}\right)=k_{a2}\widetilde{W}_{a}^{T}\left(\widetilde{W}_{c}-\widetilde{W}_{a}\right)\leq k_{a2}k_{1}\left\|\widetilde{W}_{c}\right\|-k_{a2}\left\|\widetilde{W}_{a}\right\|^{2}. \tag{68}$$

Moreover, it can be seen that

$$-u^{*T}\mathrm{R}u^{*}-2u^{*T}R\Delta\leq\Delta^{T}R\Delta\leq\lambda_{\max}(R)\overline{\Delta}^{2}. \tag{69}$$

Replacing (64) and (68) in (63), we can conclude that

$$\frac{d}{\mathrm{d}t}V_{L}\leq-\rho\left(\beta-\frac{1}{2}\right)\left\|z_{\xi}\right\|^{2}-z_{v}^{T}(Q-0.5\rho I)z_{v}-(c_{3}-k_{a1}k_{1}k_{2})\left\|\widetilde{W}_{c}\right\|^{2}-k_{a2}\left\|\widetilde{W}_{a}\right\|^{2}+k_{a1}k_{1}^{2}k_{2}k_{3}+k_{4}+\left(\frac{c_{4}k_{c}\varphi_{0}}{2\sqrt{\nu\varphi_{1}}}k_{3}+k_{a1}k_{1}k_{2}k_{3}+k_{a1}k_{1}^{2}k_{2}+k_{a2}k_{1}\right)\left\|\widetilde{W}_{c}\right\|+\lambda_{\max}(R)\overline{\Delta}^{2}. \tag{70}$$

Based on the estimation $\mathrm{ab}\leq\gamma a^{2}+(1/4\gamma)b^{2}$, we can obtain that

$$\frac{d}{\mathrm{d}t}V_{L}\leq-\rho\left(\beta_{\xi}-\frac{1}{2}\right)\left\|z_{\eta}\right\|^{2}-z_{v}^{T}(Q-0.5\rho I)z_{v}+\lambda_{\max}(R)\overline{\Delta}^{2}-(1-\theta)(c_{3}-k_{a1}k_{1}k_{2})\left\|\widetilde{W}_{c}\right\|^{2}-k_{a2}\left\|\widetilde{W}_{a}\right\|^{2}+k_{a1}k_{1}^{2}k_{2}k_{3}+k_{4}+\frac{1}{4\theta(c_{3}-k_{a1}k_{1}k_{2})}\left(\frac{c_{4}\eta_{c}\varphi_{0}}{2\sqrt{\nu\varphi_{1}}}k_{3}+k_{a1}k_{1}k_{2}k_{3}+k_{a2}k_{1}+k_{a1}k_{1}^{2}k_{2}\right)^{2}. \tag{71}$$

Let us select the coefficients to be satisfied such that $\beta>1/2, 0<\rho<2\lambda_{\min}(Q), 0<\theta<1$, and $c_3>k_{a1}k_1k_2$.

With the state vector $z=[z_{\eta}^{T},z_{v}^{T},\widetilde{W}_{c}^{T},\widetilde{W}_{a}^{T}]^{T}$, it is able to determine that there exist two K-class functions α_3 and α_4 satisfying

$$\alpha_{3}(\|z\|)\leq\rho\left(\beta_{\eta}-\frac{1}{2}\right)\left\|z_{\eta}\right\|^{2}+z_{v}^{T}(Q-0.5\rho I)z_{v}+(1-\theta)(c_{3}-k_{a1}k_{1}k_{2})\left\|\widetilde{W}_{c}\right\|^{2}+k_{a2}\left\|\widetilde{W}_{a}\right\|^{2}\leq\alpha_{4}(\|z\|). \tag{72}$$

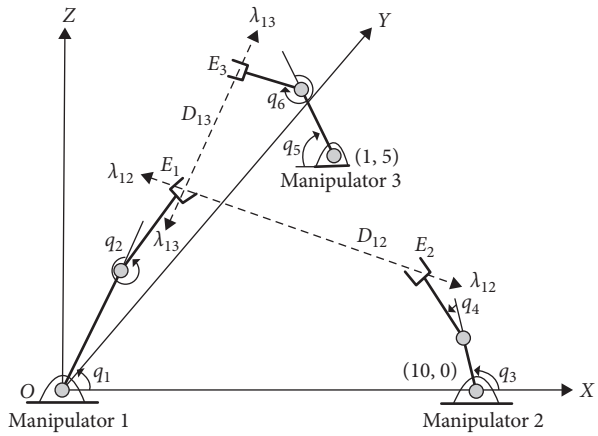


FIGURE 3: The constrained multimanipulator system [12].

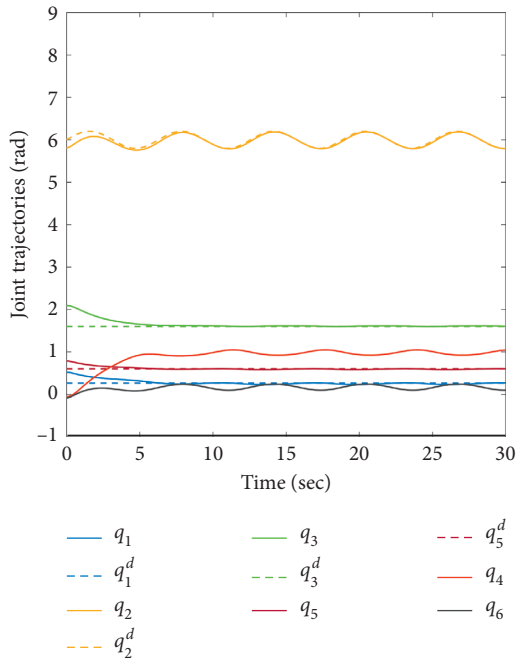


FIGURE 4: The response of 6 joint variables.

Based on (72), inequality (71) can be written as

$$\frac{d}{dt}V_L \leq -\alpha_3(\|z\|) + k_{a1}k_1^2k_2k_3 + k_4 + \lambda_{\max}(R)\overline{\Delta}^2 + \frac{1}{4\theta(c_3 - k_{a1}k_1k_2)}\left(\frac{c_4k_c\varphi_0}{2\sqrt{\nu\varphi_1}}k_3 + k_{a1}k_1k_2k_3 + k_{a1}k_1^2ka_2 + k_{a2}k_1\right)^2. \tag{73}$$

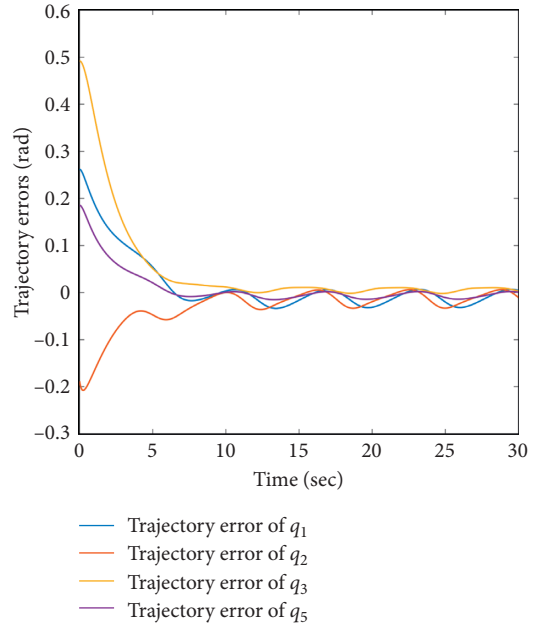


Figure 5: The tracking errors of 6 joint variables.

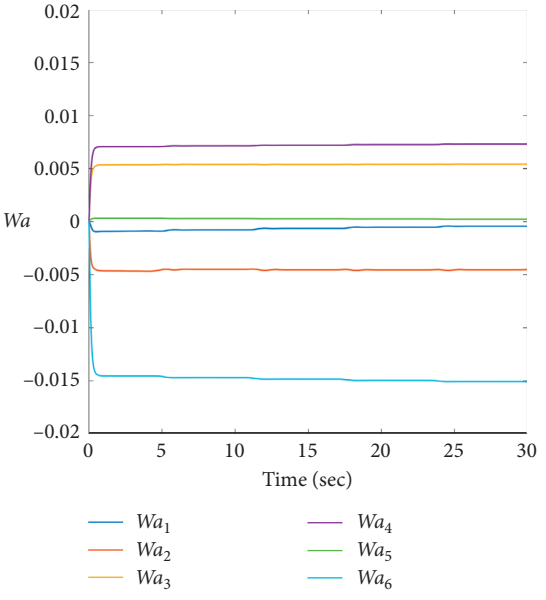


Figure 6: Weights 1–6 of the actor NN.

It is obvious that $(d/dt)V_L$ is negative if $z(t)$ lies outside the attraction region as

$$\Omega_z \triangleq \left\{ z: \|z\| \le \alpha_3^{-1}\left(\frac{1}{4\theta(c_3 - k_{a1}k_1k_2)}\left(\frac{c_4k_c\varphi_0}{2\sqrt{\nu\varphi_1}}k_3 + k_{a1}k_1k_2k_3 + k_{a1}k_1^2k_2 + k_{a2}k_1\right)^2 + k_{a1}k_1^2k_2k_3 + k_4 + \lambda_{\max}(R)\overline{\Delta}^2\right)\right\}. \tag{74}$$

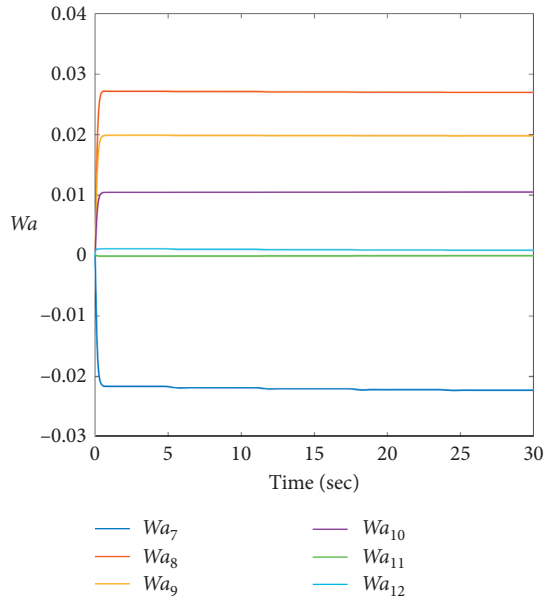


Figure 7: Weights 7–12 of the actor NN.

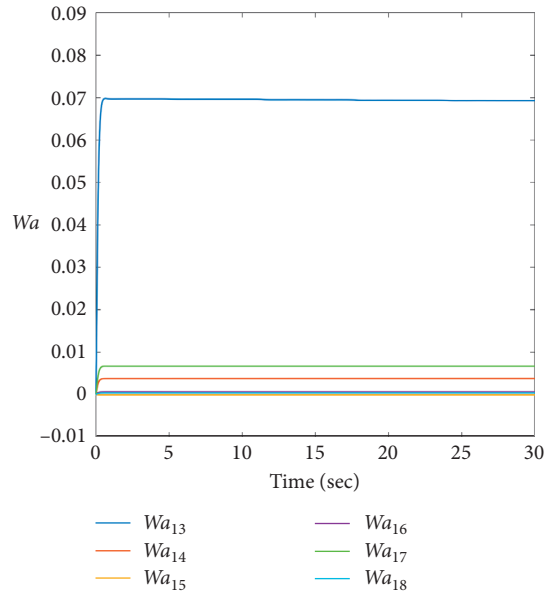


Figure 8: Weights 13–18 of the actor NN.

As a result of this, we can conclude that $\|z\|$ is UUB with attraction region (74) depending on the number of neurons in the critic NN.

Second, the tracking of constraint force is considered through the dynamic of $e_\lambda = \lambda - \lambda_d$ obtained from (47) and (48) as

$$\left(I + K_f\right)e_\lambda = N\tau^{(1)} - \mathrm{NC}^*\left(\dot{\xi} - \dot{\xi}_d\right) - \mathrm{NF}^*\left(\xi, \dot{\xi}, t\right). \quad (75)$$

According to $\dot{z}_\xi = \dot{\xi} - \dot{\xi}_d = z_v - \beta z_\xi$ and z_v, z_ξ are UUB, τ^1, N, and $F^*(\eta, \dot{\eta}, t)$ are bounded; then, e_λ is UUB with the attraction region depending on K_f and the bound of ARL-based motion dynamic control $\tau^{(1)} = \begin{bmatrix} \tau_{\mathrm{ind}}^{(1)} \\ \tau_{\mathrm{dep}}^{(1)} \end{bmatrix}$. The proposed controller is absolutely implemented for uncertain multi-manipulators with attraction region (74) depending on model error $\overline{\Delta}$. This completes the proof. □

4. Simulation Results

In this section, an example with 3 manipulators (Figure 3) is given to validate the effectiveness of the proposed control structure (Figure 2) by the *m*-file script in MATLAB software. This example has been considered in [12] by the classical robust nonlinear controller. However, it is obviously different from the existing method in [12], and the proposed motion/force control based on ARL is first structured to obtain optimality satisfaction. Similar to the model in [12], the constrained multimanipulator system is represented by dynamic equation (3) with the following constraints:

$$\begin{aligned} \phi_{12} &= \left[\left(x_{E1} - x_{E2}\right)^2 + \left(z_{E1} - z_{E2}\right)^2\right]^{(1/2)} - D_{12} = 0 \\ \phi_{13} &= \left[\left(x_{E1} - x_{E3}\right)^2 + \left(y_{E1} - y_{E3}\right)^2 + \left(z_{E1} - z_{E3}\right)^2\right]^{(1/2)} - D_{13} = 0. \end{aligned} \quad (76)$$

In this example, the matrices in (3) can be represented as in [12]:

$$\begin{aligned} M &= \begin{bmatrix} M_1 & & \\ & M_2 & \\ & & M_3 \end{bmatrix}, \\ M_1 &= \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \\ M_2 &= \begin{bmatrix} M_{33} & M_{34} \\ M_{43} & M_{44} \end{bmatrix}, \\ M_3 &= \begin{bmatrix} M_{55} & M_{56} \\ M_{65} & M_{66} \end{bmatrix}, \\ C &= \begin{bmatrix} C_1 & & \\ & C_2 & \\ & & C_3 \end{bmatrix}, \\ C_1 &= \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \\ C_2 &= \begin{bmatrix} C_{33} & C_{34} \\ C_{43} & C_{44} \end{bmatrix}, \\ C_3 &= \begin{bmatrix} C_{55} & C_{56} \\ C_{65} & C_{66} \end{bmatrix}. \end{aligned} \quad (77)$$

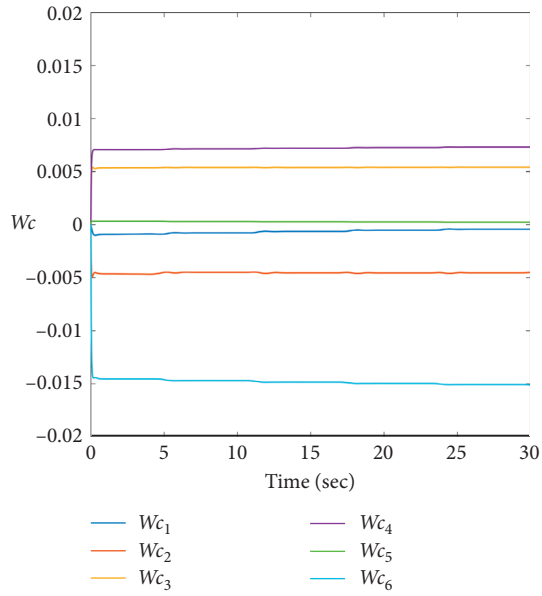


Figure 9: Weights 1–6 of the critic NN.

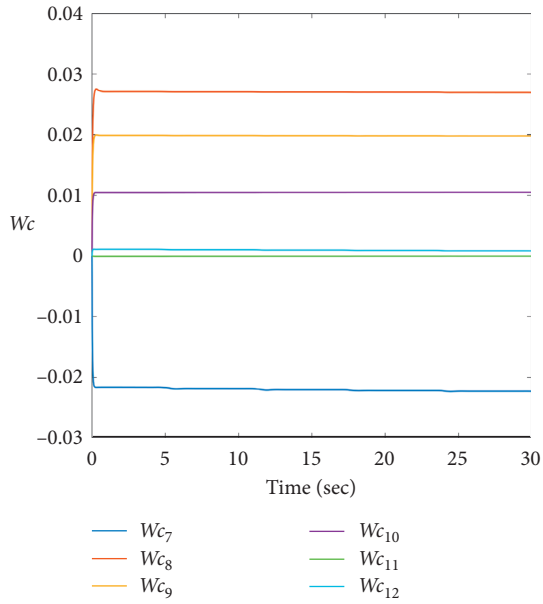


Figure 10: Weights 7–12 of the critic NN.

Moreover, to easily compare with the existing work [12], the parameters of multimanipulators are given as $D_{12} = 7\,\text{m}$, $D_{13} = 3.5\,\text{m}$, $l_1 = l_2 = 1\,\text{m}$, $m_1 = m_2 = 1\,\text{kg}$, $r_1 = r_2 = 0.9\,\text{m}$, $l_3 = l_4 = 1.5\,\text{m}$, $m_3 = m_4 = 2\,\text{kg}$, $r_3 = r_4 = 1.2\,\text{m}$, $l_5 = l_6 = 1.2\,\text{m}$, $m_5 = m_6 = 1.5\,\text{kg}$, and $r_5 = r_6 = 1\,\text{m}$. Because of two constraints (76), this multimanipulator has 4 independent joints η_1, η_2, η_3, and η_5 with the desired trajectories $\eta_{1\,d} = (\pi/12), \eta_{2\,d} = 1.91\pi + 0.2\sin(t), \eta_{3\,d} = 0.51\pi$, and $\eta_{5\,d} = 0.191\pi$, and the initial values of these independent joints are $\eta_1(0) = \pi/6, \eta_2(0) = 1.92\pi, \eta_3(0) = 2\pi/3, \eta_4(0) = -0.0565, \eta_5(0) = \pi/4$, and $\eta_6(0) = -0.0853$. According to the given multimanipulator and above parameters, we achieve the necessary models for investigating the control design, including motion dynamic model (7), (37), and constraint force (43). In light of Theorem 1, the control parameters are given as follows: $k_1 = 0.5, \eta_c = 5, \eta_{a1} = 5, \eta_{a2} = 10$, and $\nu = 0.01$. The dynamic model of 3 manipulators with the above parameters and ARL-based motion/force control scheme is established by the *m*-file script and Simulink in MATLAB software. The purpose of simulations is to verify the tracking effectiveness of not only joint variables but also constraint force coefficients. Additionally, due to the implementation of optimal control for the motion dynamic model by the radial basis function (RBF) network-based actor/critic reinforcement learning algorithm, the convergence of the adjusting mechanism of

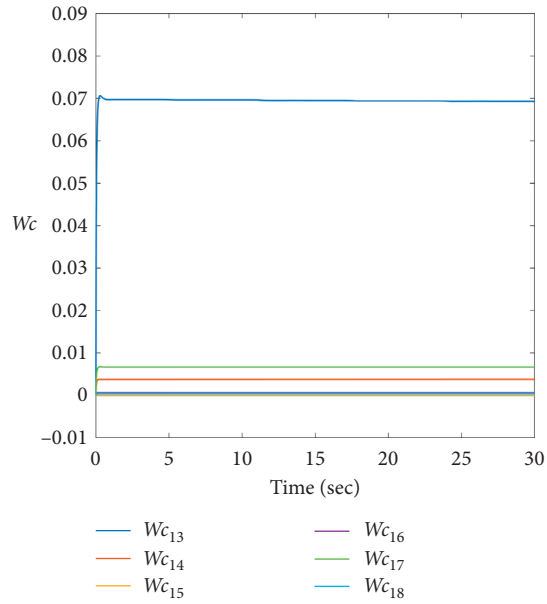


Figure 11: Weights 13–18 of the critic NN.

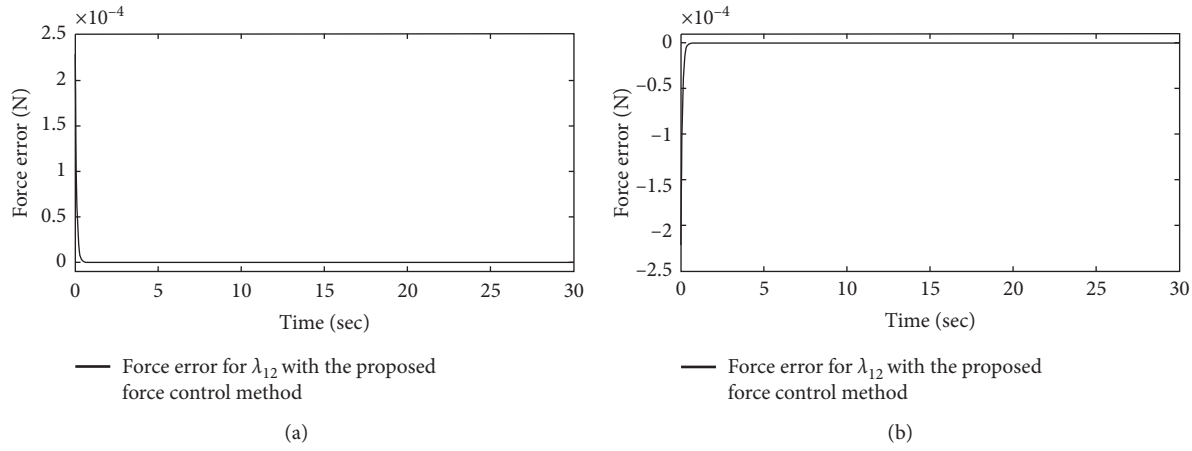


Figure 12: The response of constraint coefficients.

actor NN (35) and critic NN (33) should be verified, and the activation function is chosen depending on multi-manipulator systems:

$$\psi(Z) = \left[z_1^2, z_1z_2, z_1z_3, z_1z_4, z_2^2, z_2z_3, z_2z_4, z_3^2, z_3z_4, z_4^2, z_1^2z_9^2, z_2^2z_{10}^2, z_3^2z_{11}^2, z_4^2z_{12}^2, z_1^2z_5^2, z_2^2z_6^2, z_3^2z_7^2, z_4^2z_8^2\right], \tag{78}$$

where the vector of state variables is $Z = [z_v, z_\xi, \xi_d]^T \in \mathbb{R}^{12}$. Moreover, the adjusting mechanisms of weights in actor and critic are implemented as in (32 and 35). On the contrary, in order to complete the proposed motion/force control structure, the remaining terms need to be realized, (39) and (48). In order to illustrate the performance of the ARL scheme, we examine to obtain the actor/critic weight response and the tracking effectiveness of joint variables. Referring to Figures 4 and 5, it is seen that the response of joint variables 1–6 with tracking errors converging to zero is 10 s, which leads to that the effectiveness of closed systems is highly precise under the ARL technique. Additionally, the convergence of adjusting mechanisms of the actor NN and critic NN to the ideal weight is also obtained (0.0006, −0.0055, 0.0045, 0.0038, 0.014, −0.018, −0.009, 0.019, 0.011, 0.0055, 0.000041, 0.00051, 0.0492, 0.002, 0.000015, 0.0024, 0.0047, 0.00019) (Figures 6–11). Finally, the high-performance responses of constraint force and control inputs are shown in Figures 12 and 13. It should be noted that these simulation results are conducted for three manipulators to be described in the work [12]. However, they are obviously different with [12];

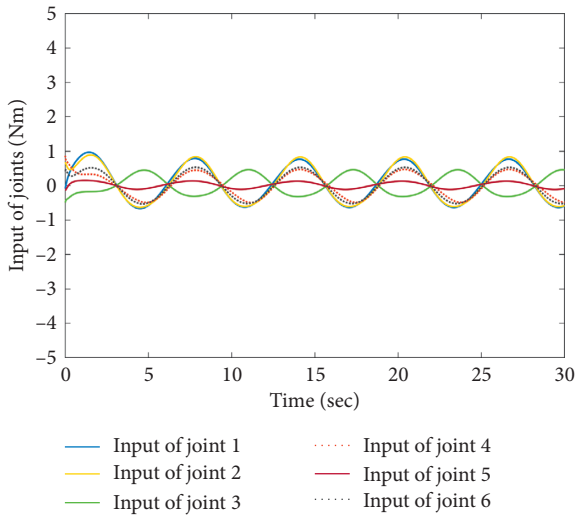


Figure 13: The control inputs of multimanipulator systems.

due to the ARL-based approach, the proposed control scheme also considers the convergence of weights in actor/critic neural networks (Figures 6–11).

5. Conclusions

This paper presented an ARL-based motion/force tracking control for multimanipulator systems, which consists of the ARL-based motion control law and nonlinear force controller. Both the minimization of the performance index and closed-loop stability are guaranteed by theoretical analysis. Moreover, the proposed control structure is able to handle the time-varying model using the additional term. The simulation studies are conducted to illustrate the efficacy of the proposed control structure. Future work is to develop completely uncertain multimanipulator systems by integrating model-free reinforcement learning.

Data Availability

This publication is supported by multiple datasets, which are available at locations cited in the reference section.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Ministry of Education and Training, Vietnam, under Grant B2020-BKA-05.

References

[1] J. Gallardo-Alvarado and J. H. Tinajero-Campos, "A parallel manipulator with planar configurable platform and three end-effectors," *Mathematical Problems in Engineering*, vol. 2019, Article ID 7972837, 12 pages, 2019.

[2] J. Shi, Y. Mao, P. Li et al., "Hybrid mutation fruit fly optimization algorithm for solving the inverse kinematics of a redundant robot manipulator," *Mathematical Problems in Engineering*, vol. 2020, Article ID 6315675, 13 pages, 2020.

[3] N. T. Binh, N. A. Tung, D. P. Nam, and N. H. Quang, "An adaptive backstepping trajectory tracking control of a tractor trailer wheeled mobile robot," *International Journal of Control, Automation and Systems*, vol. 17, no. 2, pp. 465–473, 2019.

[4] T. Nguyena, T. Hoang, M. Pham, and N. Dao, "A Gaussian wavelet network-based robust adaptive tracking controller for a wheeled mobile robot with unknown wheel slips," *International Journal of Control*, vol. 92, no. 11, pp. 2681–2692, 2019.

[5] Y.-C. Liu, P. N. Dao, and K. Y. Zhao, "On robust control of nonlinear teleoperators under dynamic uncertainties with variable time delays and without relative velocity," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1272–1280, 2020d.

[6] Z. P. Wang, T. Zhou, Y. Mao, and Q. J. Chen, "Adaptive recurrent neural network control of uncertain constrained nonholonomic mobile manipulators," *International Journal of Systems Science*, vol. 45, no. 2, pp. 133–144, 2014.

[7] M. Rani, N. Kumar, and H. P. Singh, "Motion/force control scheme for electrically driven cooperative multiple mobile manipulators," *Control Engineering Practice*, vol. 88, pp. 52–64, 2019.

[8] J. Han, X. Liu, X. Gao, and X. Wei, "Intermediate observer-based robust distributed fault estimation for nonlinear multiagent systems with directed graphs," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7426–7436, 2020.

[9] X. Liu, X. Gao, and J. Han, "Distributed fault estimation for a class of nonlinear multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 9, pp. 3382–3390, 2020a.

[10] B. Baigzadehnoe, Z. Rahmani, A. Khosravi, and B. Rezaie, "On position/force tracking control problem of cooperative robot manipulators using adaptive fuzzy backstepping approach," *ISA Transactions*, vol. 70, pp. 432–446, 2017.

[11] C. Chen, Z. Liu, Y. Zhang, and S. Xie, "Coordinated motion/force control of multiarm robot with unknown sensor nonlinearity and manipulated object's uncertainty," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, pp. 1123–1134, 2016.

[12] H. Dou and S. Wang, "Robust adaptive motion/force control for motion synchronization of multiple uncertain two-link manipulators," *Mechanism and Machine Theory*, vol. 67, pp. 77–93, 2013.

[13] L. Kong, W. He, C. Yang, Z. Li, and C. Sun, "Adaptive fuzzy control for coordinated multiple robots with constraint using impedance learning," *IEEE Transactions on Cybernetics*, vol. 49, no. 8, pp. 3052–3063, 2019.

[14] Y. Li, C. Yang, W. Yan, R. Cui, and A. Annamalai, "Admittance-based adaptive cooperative control for multiple manipulators with output constraints," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3621–3632, 2019.

[15] Z. Li, B. Huang, A. Ajoudani, C. Yang, C. Y. Su, and A. Bicchi, "Asymmetric bimanual control of dual-arm exoskeletons for human-cooperative manipulations," *IEEE Transactions on Robotics*, vol. 34, pp. 264–271, 2017b.

[16] Z. Li, W. Yuan, S. Zhao, Z. Yu, Y. Kang, and C. P. Chen, "Brain-actuated control of dual-arm robot manipulation with relative motion," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, pp. 51–62, 2017c.

[17] V. T. Ngo and Y. C. Liu, "Object transportation with force-sensorless control and event-triggered synchronization for networked uncertain manipulators," *IEEE Transactions on Industrial Electronics*, vol. 68, pp. 1–10, 2020.

[18] T. Zhao, Y. Liu, Z. Li, C. Y. Su, and Y. Feng, "Adaptive control and optimization of mobile manipulation subject to input saturation and switching constraints," *IEEE Transactions on Automation Science and Engineering*, vol. 16, pp. 1543–1555, 2018.

[19] Z. Li, P. Y. Tao, S. S. Ge, M. Adams, and W. S. Wijesoma, "Robust adaptive control of cooperating mobile manipulators with relative motion," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, pp. 103–116, 2008b.

[20] Z. Li, C. Yang, and Y. Tang, "Decentralised adaptive fuzzy control of coordinated multiple mobile manipulators interacting with non-rigid environments," *IET Control Theory & Applications*, vol. 7, no. 3, pp. 397–410, 2013.

[21] M. Deng, Z. Li, Y. Kang, C. P. Chen, and X. Chu, "A learning-based hierarchical control scheme for an exoskeleton robot in human–robot cooperative manipulation," *IEEE Transactions on Cybernetics*, vol. 50, pp. 112–125, 2018.

[22] X. Yu, W. He, H. Li, and J. Sun, "Adaptive fuzzy full-state and output-feedback control for uncertain robots with output constraint," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 99, pp. 1–14, 2020.

[23] Z. Li, S. Deng, C. Y. Su et al., "Decentralised adaptive control of cooperating robotic manipulators with disturbance observers," *IET Control Theory & Applications*, vol. 8, no. 7, pp. 515–521, 2014a.

[24] Z. Li, S. S. Ge, and A. Ming, "Adaptive robust motion/force control of holonomic-constrained nonholonomic mobile manipulators," *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 3, pp. 607–616, 2007.

[25] Z. Li, W. Chen, and J. Luo, "Adaptive compliant force-motion control of coordinated non-holonomic mobile manipulators interacting with unknown non-rigid environments," *Neurocomputing*, vol. 71, no. 7-9, pp. 1330–1344, 2008a.

[26] X. Liu, S. S. Ge, F. Zhao, and X. Mei, "Optimized impedance adaptation of robot manipulator interacting with unknown environment," *IEEE Transactions on Control Systems Technology*, vol. 29, pp. 411–419, 2020b.

[27] X. Liu, S. S. Ge, F. Zhao, and X. Mei, "Optimized interaction control for robot manipulator interacting with flexible environment," *IEEE/ASME Transactions on Mechatronics , Early Access*, vol. 99, p. 1, 2020c.

[28] C. Yang, G. Peng, Y. Li, R. Cui, L. Cheng, and Z. Li, "Neural networks enhanced adaptive admittance control of optimized robot–environment interaction," *IEEE Transactions on Cybernetics*, vol. 49, pp. 2568–2579, 2018.

[29] W. He, C. Xue, X. Yu, Z. Li, and C. Yang, "Admittance-based controller design for physical human-robot interaction in the constrained task space," *IEEE Transactions on Automation Science and Engineering*, vol. 62, 2020b.

[30] S. Ganjefar, S. Najibi, and H. Momeni, "A novel structure for the optimal control of bilateral teleoperation systems with variable time delay," *Journal of the Franklin Institute*, vol. 348, no. 7, pp. 1537–1555, 2011.

[31] P. N. Dao and Y. C. Liu, "Adaptive reinforcement learning strategy with sliding mode control for unknown and disturbed wheeled inverted pendulum," *International Journal of Control, Automation and Systems*, vol. 60, pp. 1–12, 2021.

[32] P. N. Dao, P. T. Loc, T. Q. Huy et al., "Sliding variable-based online adaptive reinforcement learning of uncertain/disturbed nonlinear mechanical systems," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 2, pp. 1–10, 2021.

[33] J. Li, J. Ding, T. Chai, F. L. Lewis, and S. Jagannathan, "Adaptive interleaved reinforcement learning: robust stability of affine nonlinear systems with unknown uncertainty," *IEEE Transactions on Neural Networks and Learning Systems , Early Access*, vol. 34, no. 1, pp. 1–11, 2021.

[34] Y. Lv, X. Ren, and J. Na, "Adaptive optimal tracking controls of unknown multi-input systems based on nonzero-sum game theory," *Journal of the Franklin Institute*, vol. 356, no. 15, pp. 8255–8277, 2019.

[35] C. Mu, Z. Ni, C. Sun, and H. He, "Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems," *IEEE Transactions on Cybernetics*, vol. 47, pp. 1460–1470, 2016.

[36] T. Sun and X. M. Sun, "An adaptive dynamic programming scheme for nonlinear optimal control with unknown dynamics and its application to turbofan engines," *IEEE Transactions on Industrial Informatics*, vol. 17, p. 14, 2020.

[37] Z. Wang, H. Li, H. Wu, and Z. Wu, "Improving maneuver strategy in air combat by alternate freeze games with a deep reinforcement learning algorithm," *Mathematical Problems in Engineering*, vol. 2020, Article ID 7180639, 17 pages, 2020.

[38] H. N. Wu and Z. Y. Liu, "Data-driven guaranteed cost control design via reinforcement learning for linear systems with parameter uncertainties," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, 2019.

[39] X. Yang, H. He, and X. Zhong, "Approximate dynamic programming for nonlinear-constrained optimizations," *IEEE Transactions on Cybernetics*, vol. 99, pp. 1–14, 2019.

[40] W. He, H. Gao, C. Zhou, C. Yang, and Z. Li, "Reinforcement learning control of a flexible two-link manipulator: an experimental investigation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 11, 2020a.

[41] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

[42] K. G. Vamvoudakis, "Q-learning for continuous-time linear systems: a model-free infinite horizon optimal control approach," *Systems & Control Letters*, vol. 100, pp. 14–20, 2017.

[43] Z. Li and S. S. Ge, *Fundamentals in Modeling and Control of Mobile Manipulators*, CRC Press, Boca Raton, FL, USA, 2013.

[44] Z. Li, C. Yang, C. Y. Su, S. Deng, F. Sun, and W. Zhang, "Decentralized fuzzy control of multiple cooperating robotic manipulators with impedance interaction," *IEEE Transactions on Fuzzy Systems*, vol. 23, pp. 1044–1056, 2014b.

[45] W. He, Y. Chen, and Z. Yin, "Adaptive neural network control of an uncertain robot with full-state constraints," *IEEE Transactions on Cybernetics*, vol. 46, pp. 620–629, 2015.

[46] Y. Zhu, D. Zhao, and X. Li, "Using reinforcement learning techniques to solve continuous-time non-linear optimal tracking problem without system dynamics," *IET Control Theory & Applications*, vol. 10, no. 12, pp. 1339–1347, 2016.

[47] J. Li, B. Kiumarsi, T. Chai, F. L. Lewis, and J. Fan, "Off-policy reinforcement learning: optimal operational control for two-time-scale industrial processes," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4547–4558, 2017a.

[48] Y. Peng, Q. Chen, and W. Sun, "Reinforcement q-learning algorithm for h tracking control of unknown discrete-time

linear systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, 2019.

[49] Y. Lv, X. Ren, and J. Na, "Online nash-optimization tracking control of multi-motor driven load system with simplified rl scheme," *ISA Transactions*, vol. 98, pp. 251–262, 2020.

[50] J. Na, Y. Lv, K. Zhang, and J. Zhao, "Adaptive identifier-critic-based optimal tracking control for nonlinear systems with experimental validation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 60, no. 1, pp. 1–14, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 5558508, 13 pages
https://doi.org/10.1155/2021/5558508

Research Article

GPU Preconditioning for Block Linear Systems Using Block Incomplete Sparse Approximate Inverses

Wenpeng Ma,[1] **Yiwen Hu**,[1] **Wu Yuan**,[2] **and Xiazhen Liu**[2]

[1]*College of Computer and Information Technology, Xinyang Normal University, Xinyang, Henan 464000, China*
[2]*Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China*

Correspondence should be addressed to Wenpeng Ma; mawp@xynu.edu.cn

Received 5 January 2021; Revised 29 March 2021; Accepted 9 May 2021; Published 26 May 2021

Academic Editor: Hussein Abulkasim

Solving sparse triangular systems is the building block for incomplete LU- (ILU-) based preconditioning, but parallel algorithms, such as the level-scheduling scheme, are sometimes limited by available parallelism extracted from the sparsity pattern. In this study, the block version of the incomplete sparse approximate inverses (ISAI) algorithm is studied, and the block-ISAI is considered for preconditioning by proposing an efficient algorithm and implementation on graphical processing unit (GPU) accelerators. Performance comparisons are carried out between the proposed algorithm and serial and parallel block triangular solvers from PETSc and cuSPARSE libraries. The experimental results show that GMRES (30) with the proposed block-ISAI preconditioning achieves accelerations $1.4\times$ –$6.9\times$ speedups over that using the cuSPARSE library on NVIDIA Tesla V100 GPU.

1. Introduction

Sparse triangular solves are essential steps for the incomplete LU- (ILU-) factorized preconditioning [1] in a Krylov subspace solver. However, they are sequentially designed and pose a performance challenge on today's parallel computers. To make full use of the state of art parallel architectures such as multicore CPUs and GPU accelerators, various works have been studied and presented. The well-known algorithm is the level-scheduling scheme [1–5] which groups the possible parallelism in levels of sets each of which can be performed in parallel. The implementations on multicore architectures and GPU accelerators are discussed in [1, 4, 5] and [2, 3], respectively. The NVIDIA's cuSPARSE [6] library offers a function interface for this algorithm which can be called by users directly in practice. However, there are two drawbacks in some of the applications. One is that the extracted parallelism depends on the sparsity pattern of the matrix; thus, the parallelism and performance is highly limited for some cases. The other is that it is usually a time-consuming job to search parallelism (known as preprocessing) before the actual computation is conducted. Liu et al. [7, 8] proposed a method in which the preprocessing step is not necessarily performed when the sparse matrix is expressed in compressed sparse column (CSC) format, and they showed performance improvement on GPU over the level scheduling method in cuSPARSE.

Compared to using exact computation for preconditioning, some inexact preconditioning ideas [9–13] are attractive in recent years because they seek a tradeoff between exactness and parallelism. Chow et al. [12, 13] proposed a fine-grained algorithm to compute ILU factorization asynchronously on Intel MIC architecture [14] and GPUs. They showed that 5 asynchronous sweeps usually make the inexact ILU factorization comparable to the exact one, but with a significant factor of speedups. Anzt et al. [9] implemented an asynchronous method on GPU for solving triangular systems using several number of Jacobi iterations. Although the inexact preconditioning step results in more number of solver iterations, it shows an advantage over the exact preconditioning in terms of the total compute time of the linear solver. Some methods are based on sparse approximate inverses (SAI) [15–19] where the matrix inverse of the triangular factors is estimated through solving least squares problems on a preset pattern. And the resulting inverses can transform the preconditioning step into sparse

matrix-vector multiplications with much more possible of concurrency computing on parallel computers. Most recently, Anzt et al. [10, 11] proposed an incomplete SAI (ISAI) algorithm in which the least square problems are replaced with solving square systems with cheaper computations and faster convergence.

Linear systems arising from block sparse matrices are widely used in scientific computing especially in multiphysics problems. Most of the numerical algorithms for block linear systems are derived from that for scalar systems. Even though they are quite similar to each other in a mathematical formula, the implementation strategies and performance tuning techniques could be very different from each other, especially on GPU accelerators. For example, the block sparse matrix-vector multiplication (BSpMV) on GPU [20] is performed in the multiplications of blocks and vectors, whereas a general SpMV is realized in scalar multiplications. Due to the principles of global memory access and the use of shared memory on the GPU, direct migration of the code from scalar case to block case may hamper the performance much [20]. Therefore, algorithms for block matrices usually require redesigned work. In multiphysics problems, the coupling feature of the physical fields results in block matrices that usually have blocks with a small size such that the inverse of a block can be explicitly expressed. Motivated by this kind of numerical applications and the ISAI preconditioning proposed by Anzt et al. [10, 11], we focus on the GPU preconditioning in a block format in this study. The main contributions of this study are the following.

(1) The GPU preconditioning framework [10, 11] is extended to block matrices with block sizes up to 5.

(2) An efficient, warp-based GPU implementation exploiting fine-grained programming model for block-ISAI preconditioning is proposed and elaborately explained.

(3) Detailed comparisons are made between the proposed algorithm and block triangular solvers from popular libraries, including PETSc [21] and cuSPARSE [6]. On block matrices selected from the SuiteSparse collection [22] and real multiphysics areas, the proposed algorithm shows an advantage over the PETSc's serial and cuSPARSE's parallel implementations of block triangular solvers in terms of total computing time for GMRES (30).

The rest of this study is organized as follows. In Section 2, some backgrounds, including sparse approximate inverse (SAI), ISAI, and block matrices, are introduced. In Section 3, the GPU preconditioning framework for block linear systems is proposed, and the strategy for GPU implementation is introduced and discussed. In Section 4, performance results and comparisons for four testing cases are shown. Concluding remarks are given in Section 5.

2. Background

2.1. Incomplete Sparse Approximate Inverses. For a given sparse matrix $A_{n\times n}$ with n rows and n columns, the SAI algorithm [1, 15–19] gives an approximation of the inverse of A by minimizing the Frobenius norm of $(\mathrm{AW} - I)$ as

$$\min_{W\in S_W} \|\mathrm{AW} - I\|_F^2 = \sum_{j=0}^{n-1} \min_{W_j\in S_W} \left\|\mathrm{AW}_j - I_j\right\|_2^2, \tag{1}$$

where W is the estimated inverse of A for a given sparsity pattern S_W, I is the identity matrix, W_j and I_j represent the j^{th} column of W and I, respectively. Solving (1) is equivalent to solve

$$\sum_{j=0}^{n-1} \min_{W_j(J)} \left\|A(R,J)W_j(J) - I_j(R)\right\|_2^2, \tag{2}$$

where J is the non-zero pattern for W_j, R represents the indices of the rows that contain non-zero values at J columns, and $A(R,J)$ is a submatrix extracted from the R^{th} rows and J^{th} columns of A. Note that $J\subseteq R$ if the diagonal values of A are non-zeros, and (2) can be viewed as the solve of a batch of least-squares problems. The SAI algorithm fits parallel computers because the problems in (2) can be solved independently of each other.

The work in [10, 11] proposed an incomplete-SAI (ISAI) by considering each problem in (2) as

$$\min_{W_j(J)} \left\|A(J,J)W_j(J) - I_j(J)\right\|_2^2. \tag{3}$$

This simplification makes $A(R,J)$ a square matrix, and the least-squares problems in (2) are changed into linear systems in (3). In this way, the ISAI can be easily applied to the matrices having special sparsity patterns. For example, in the case where incomplete LU (ILU) factorizations-based preconditioning is performed, the inverse of the lower triangular factor (L) and upper triangular factor (U) are approximately calculated with high concurrency using (3) [10, 11]. This avoids solving large sparse triangular systems which are regarded as the bottleneck in today's state-of-art parallel computers and accelerators.

2.2. Block Sparse Matrix. Block sparse matrices are widely used in scientific computing applications, especially in multiphysics problems. Figure 1 shows a 3×3 block sparse matrix with a block size of 2. It is neither a general sparse matrix nor a dense matrix, but it falls somewhere in between. It has the sparse property in global but dense features in local blocks.

There are many storage formats for block sparse matrices, such as block compressed sparse row (BCSR) and block compressed sparse column (BCSC) in PETSc [21] and cuSPARSE [6]. Different from the general CSR format for storing scalar values, BCSR format treats each block as a unit and stores all non-zero positions block by block (Figure 1). The values in each block are stored consecutively either in a row-major or column-major format. In the following parts of this study, block matrices are focused on, and column-major format is employed to store non-zero blocks in BCSR or BCSC formats.

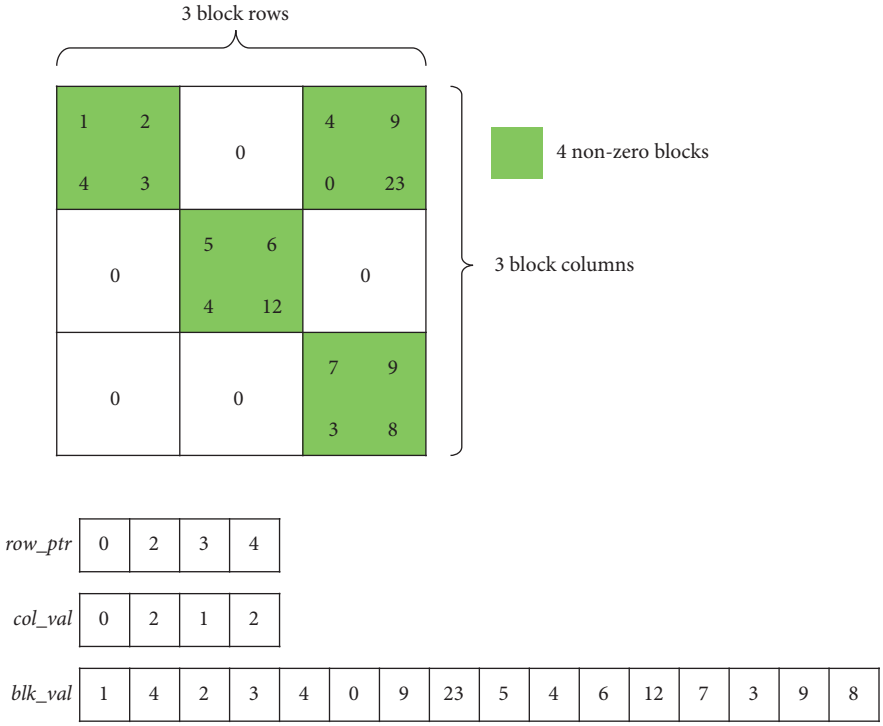


Figure 1: Example of 3×3 block matrix with four non-zero blocks, and the three arrays storing the matrix in BCSR with column-major within block.

3. Block-ISAI Preconditioning on GPUs

To solve an asymmetric block linear system $A_B x = b$, the Krylov subspace-based generalized minimal residual (GRMES) algorithm [23] with right preconditioning is employed as

$$\begin{aligned} A_B M^{-1} x_d &= b, \\ x &= M^{-1} x_d, \end{aligned} \tag{4}$$

where M is a preconditioner.

The well-known method to construct a preconditioner M is to perform a block incomplete LU factorization (BILU) on A_B [21, 24]. Then, the preconditioning step in which the inverse of M applies to a vector v, i.e., $M^{-1}v$, can be executed by

$$M^{-1}v = z \Rightarrow LUz = v, \tag{5}$$

where L and U are the block lower and upper triangular matrices factorized by A_B, respectively, and z is the preconditioned vector.

3.1. Block-ISAI on GPU. Traditionally, the blocked-based forward and backward substitutions are applied to the two triangular systems in (5). However, these are totally sequential operations. Instead of computing (5) accurately, Anzt et al. [10, 11] introduced an inexact idea in which the inverse of the lower and upper factors are estimated by highly parallel ISAI, and (5) is transformed into the problems of ISAI relaxation steps or ISAI SpMV. This method seeks a tradeoff between parallelism and exactness and shows an advantage in terms of total computing time of a linear solver.

Motivated by the scalar version of ISAI on a single GPU [10, 11] and the block version [25] on Intel's MIC (many integrated core) architecture [14], we propose a GPU-enabled block-ISAI algorithm for preconditioning in block matrices. For the convenience of our statement, we assume all the matrices in the following discussion are block matrices.

In the context of incomplete factorization preconditioners, we applied the incompleteness idea of (3) to the block lower factor L and upper factor U to have the estimations of the inverses, NL and NU, for the factors as

$$\begin{aligned} &\min_{NL_j(J_l)} \left\| L(J_l, J_l) NL_j(J_l) - I_j(J_l) \right\|_2^2, \\ &\min_{NU_j(J_u)} \left\| U(J_u, J_u) NU_j(J_u) - I_j(J_u) \right\|_2^2, \end{aligned} \tag{6}$$

where J_l and J_u are the indices of non-zero blocks at the j^{th} block column of L and U, respectively. As $L(J_l, J_l)$ and $U(J_u, J_u)$ are square block matrices, the solve of (6) is equivalent to the solve of a series of triangular systems in the block format independently

$$\begin{aligned} L(J_l, J_l) NL_j(J_l) &= I_j(J_l), \quad \text{and } U(J_u, J_u) NU_j(J_u) \\ &= I_j(J_u) \quad (j = 0, 1, 2, \ldots, n-1). \end{aligned} \tag{7}$$

The sizes of the systems in (7) are determined by the numbers of non-zero blocks in block columns, and they

usually are much smaller than the general large sparse systems. Therefore, the solve of the systems provides relatively fine-grained parallelism that fits the GPU architecture. We show the GPU-enabled block-ISAI in Algorithm 1.

To compute the preconditioned vector z, Algorithm 1 applies the preconditioner M to the input vector v on the GPU. Specifically, the GPU accepts the preconditioner M expressed in two block triangular factors L and U and a vector v as input parameters and outputs the preconditioned vector z. This can be implemented in two phases. The first phase aims to obtain the approximate inverses for L and U (denoted NL and NU, respectively) by the block-ISAI consisting of four main steps. The first step gives the guesses of the sparsity patterns for NL and NU. Following the strategy in [10, 11], we use $S(|L|^k)$ and $S(|U|^k)$ $(k \geq 1)$ as the sparsity patterns for NL and NU, respectively, where $S(|L|^k)$ and $S(|U|^k)$ represent the sparsity patterns for the multiplications of k times of absolute values of L and U, respectively. The second step extracts two block triangular systems, denoted $L(J_l, J_l)$ and $U(J_u, J_u)$, according to the non-zero patterns J_l and J_u at each column of NL and NU. The extraction process is illustrated in Figure 2, which shows how the block lower triangular system corresponding to the fourth column of NL is formed. The non-zero pattern for NL is constructed using $S(|L|^2)$ which is denser than $S(|L|)$. The non-zero pattern $J_l = \{4, 5, 7, 8\}$ for the fourth column of NL provides a row and column set (J_l, J_l) for indexing L, and then, the 4^{th}, 5^{th}, 7^{th}, and 8^{th} block rows and columns are extracted from L as a small block lower triangular matrix. In the third step, looping over all block columns of NL (NU) and storing all corresponding triangular matrices consecutively, we form two groups of small block triangular systems, LGSBTS and UGSBTS. The fourth triangular system in the LGSBTS, also shown in Figure 2, is formed by

$$L(J_l, J_l)\text{Sol}_4 = I_4(J_l), \tag{8}$$

where Sol_4 is the solution for the fourth block column of NL, and $I_4(J_l)$ is the right hand side with the first block being identity matrix. The last step is to solve the LGSBTS and UGSBTS block column by block column for the solutions of NL and NU, respectively. Generally, since the preconditioning matrix $M = \text{LU}$ remains unchanged during the solve of $A_B x = b$, the first phase needs to be performed only once.

Once the approximate solutions of NL and NU are obtained, the second phase can be accomplished by performing two block sparse matrix-vector multiplications, i.e., $y = \text{NL}v$ and $z = \text{NU}y$. These two operations must be conducted in every iteration because v changes at every iteration.

3.2. GPU Implementation. Although the block version of the ISAI looks similar to the scalar version, the computations in the two versions are entirely different. For example, Sol_4 in (8) consists of four blocks instead of four scalars, matrix-matrix multiplications is conducted instead of scalar-scalar multiplications, and the inverse of a non-zero block is computed instead of the inverse of a scalar. Therefore, to make full use of fine-grained features of a GPU, the implementations and tuning strategies for the two versions are also different. In the following discussion, an implementation of Algorithm 1 is introduced by developing several efficient GPU kernels and exploiting the cuSPARSE library [6].

For the first step in S1 of Algorithm 1, the essence of estimating the sparsity patterns of NL and NU is to conduct block matrix-matrix multiplication, which usually involves symbolic multiplication and numerical multiplication. However, since only the non-zero patterns of NL and NU are needed, only the symbolic multiplication of L^k and U^k is employed on the GPU. This is realized by calling the presetup function which estimates the non-zero pattern for the multiplication result of two general sparse matrices.

In the scalar version of the ISAI, there are two strategies introduced in [10, 11] for the GPU implementation, including the following three steps. One can separately design three kernels, each of which is responsible for a single step of Algorithm 1. In this way, the LGSBTS and UGSBTS must be formed explicitly for the solve [10]. An alternative method [11] is to merge all three steps into a single kernel in which the data $L(J_l, J_l)$ (or $U(J_u, J_u)$) extracted from L (or U) are directly used for the partial solution of one system from LGSBTS (UGSBTS). This means each system is formed temporarily, and the data for one system could be overwritten by subsequent systems. The main advantage of the first strategy is that one can explicitly express the LGSBTS (or UGSBTS) in an order that the memory pattern is perfect for coalesced memory accesses on the GPU. However, the drawback is obvious. To solve NL's j^{th} block column with $\text{nnzbl}(j)$ blocks, according to (7), it requires to explicitly storing the corresponding $\text{nnzbl}(j) \times \text{nnzbl}(j)$ lower triangular block matrix that contains $((\text{nnzbl}(j) \times (\text{nnzbl}(j) + 1))/2)$ blocks, i.e., $(s^2 \times \text{nnzbl}(j) \times (\text{nnzbl}(j) + 1)/2)$ scalar elements, where s is the block size. By adding the number of blocks up for all nbr block columns in NL targeting parallel computing, it requires approximately $s^2 \times \sum_{j=1}^{\text{nbr}} ((\text{nnzbl}(j) \times (\text{nnzbl}(j) + 1))/2)$ double precision memory spaces for storing all matrices in (7). It is estimated that explicitly storing LGSBTS for $\text{nbr} = 200,000$, $s = 3$, and $\text{nnzbl}(j) = 10$ requires over 750 MB of memory, and the memory requirement exceeds 2 GB when the block size is 5. The memory requirements will double if the LGSBTS and UGSBTS are both stored. By considering the space complexity of the first strategy and the limited memory resources on the GPU, in the block format, explicitly forming the LGSBTS and UGSBTS in the implementation proposed in this study is avoided.

A key point in S1 is the solving of the small block triangular systems (SBTSs) that fit the GPU architecture very well because they are independent of each other. However, the GPU strategy [10, 11] for the scalar version is not an optimized choice for the block version here because mapping a thread to a matrix block for numerical operations leads to significantly uncoalesced memory accesses on the GPU. Moreover, the experiments on block sparse matrix-vector multiplication in [20] show that using consecutive threads (usually 32 threads in a warp) collaboratively for the

Input:
(1) The BILU triangular factors, L and U ($\mathbf{M} = \mathbf{LU}$);
(2) Input vector, v;
Output:
(1) The preconditioned vector, z;
S1: **setup step**:
S1.1: estimate the sparsity patterns for the inverses of L and U expressed as NL and NU: $S(\mathrm{NL}) = S(|L|^k)$ and $S(\mathrm{NU}) = S(|U|^k)$;
S1.2: extract $L(J_l, J_l)$ and $U(J_u, J_u)$ from L and U, where J_l and J_u are the non-zero patterns at the j^{th} column of NL and NU, respectively, indexing J_l^{th} block rows and columns in L to form $L(J_l, J_l)$ and indexing J_u^{th} block rows and columns in U to form $U(J_u, J_u)$;
S1.3: form two groups of small block triangular systems in (7), LGSBTS: $L(J_l, J_l)\mathrm{NL}_j(J_l) = I_j(J_l)$ and UGSBTS: $U(J_u, J_u)\mathrm{NU}_j(J_u) = I_j(J_u)$, by looping overall J_l and J_u;
S1.4: solve LGSBTS and UGSBTS for the solutions of NL and NU: solve n block lower and upper triangular systems in (7) in parallel.
S2: preconditioning step:
S2.1: perform BSpMV: $y = L^{-1}v = \mathrm{NL}v$;
S2.2: perform BSpMV: $z = U^{-1}y = \mathrm{NU}y$.

Algorithm 1: Block-ISAI preconditioning on GPU: $z = \mathbf{M}^{-1}v$.

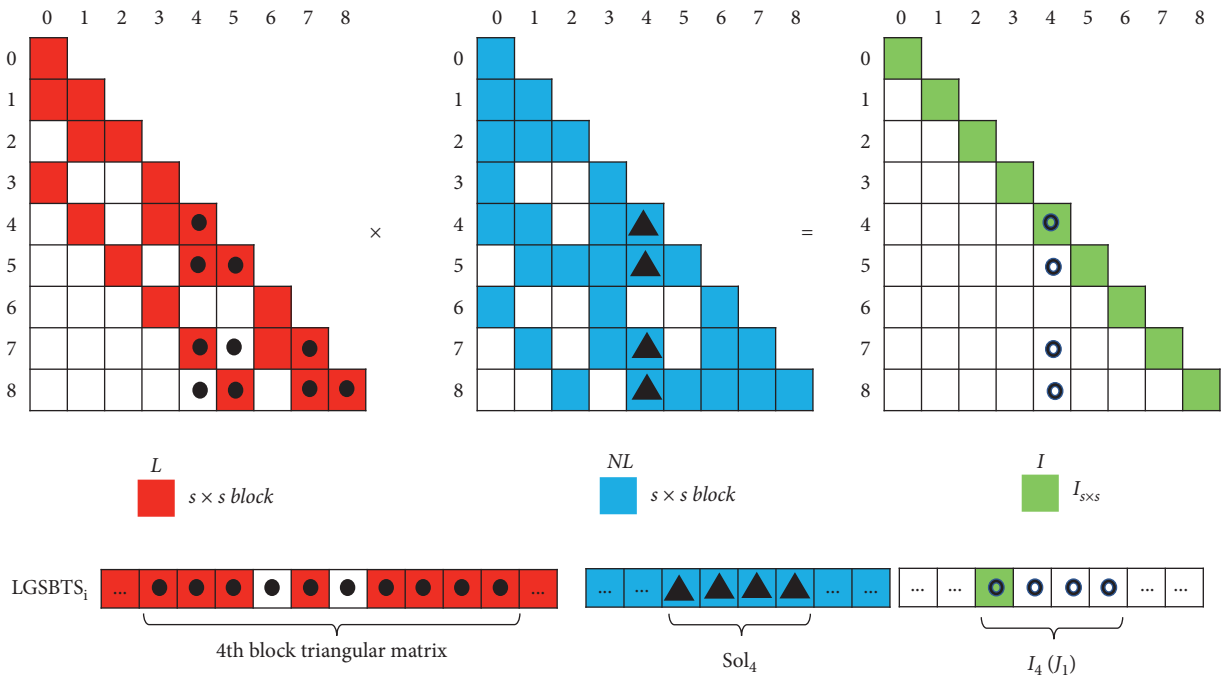


Figure 2: Illustration of extracting block triangular matrix and forming groups of triangular systems for solving NL (s is the block size).

computing of consecutive blocks is good for bandwidth utilization. Therefore, 32 threads in a warp are employed for the solve of a SBTS. The idea of the algorithm is shown in Figure 3, which shows the first warp in a thread block of size 128 for the solution of (8) with block size 4 in parallel. The small block triangular matrix (SBTM) is extracted from L in columns, so that a warp can handle a column with possible data concurrency. Specifically, when the j^{th} block column is accessed by a warp, the j^{th} block of Sol_4, denoted as $\mathrm{Sol}_4(j)$, is ready to be solved, and once $\mathrm{Sol}_4(j)$ is obtained, it can contribute to the solving of $\mathrm{Sol}_4(k)\,(k > j)$ by performing in parallel

$$I_4(k) = I_4(k) - \mathrm{SBTM}(k, j)\mathrm{Sol}_4(j), \quad (k > j). \tag{9}$$

Taking the SBTS's 0^{th} block column consisting of four blocks as an example, the 0^{th} block is used to solve $\mathrm{Sol}_4(0)$, and this is followed by accumulating $I_4(1)$ and $I_4(2)$ simultaneously. As a warp can only cover two 4×4 blocks, it restarts the computation for the rest (i.e., $I_4(3)$) in the next cycle, until it has gone through all blocks in the column.

The GPU kernel developed by the CUDA C/C++ language for solving the LGSBTS in Algorithm 2 is now described. The one-dimensional thread organization is employed. Since a warp (32 threads) is chosen and implicitly

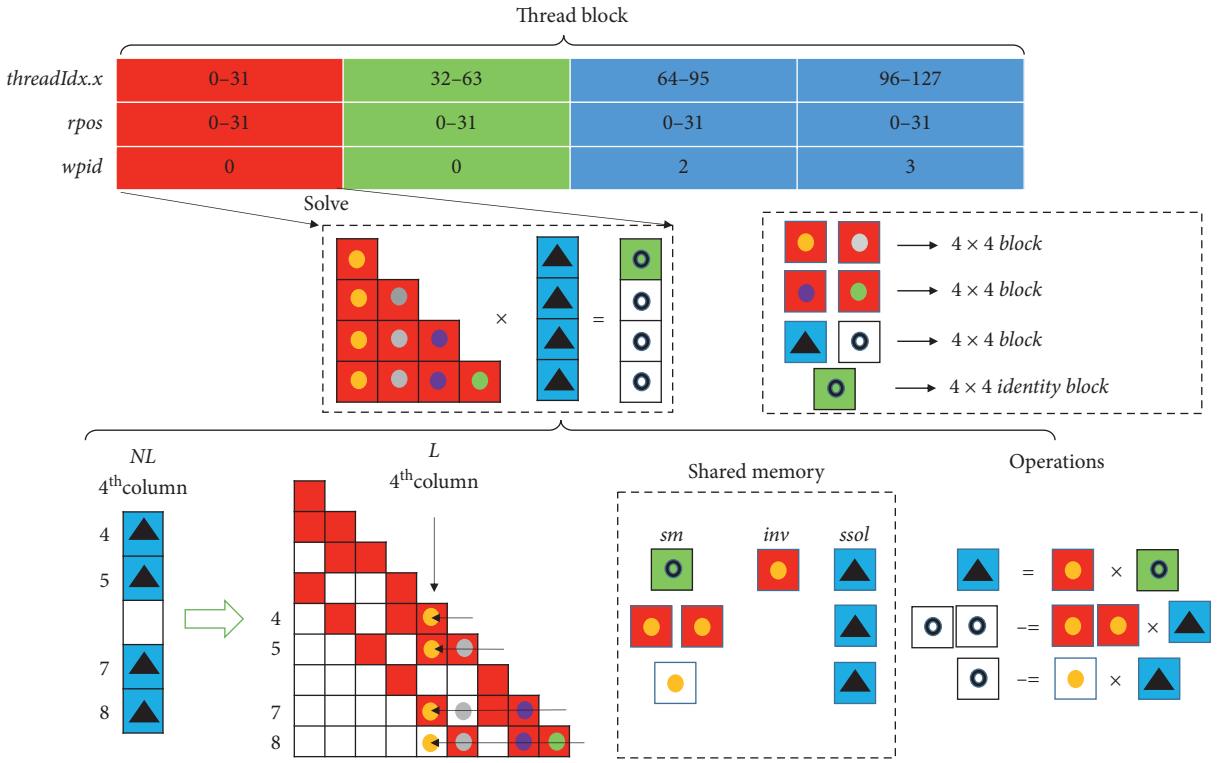


Figure 3: Illustration of warp-based formation of a SBTM and solving a SBTS.

```
(1) __shared__ double sm[warps_per_block][warp_size];
(2) __shared__ double inv[warps_per_block][bsz][bsz];
(3) __shared__ double ssol[warps_per_block][bsz][bsz];
(4) tid = threadIdx.x + blockIdx.x × blockDim.x;
(5) wpid = threadIdx.x/warp_size;
(6) gwpid = tid/warp_size;
(7) rpos = threadIdx.x%warp size;
(8) range = (warp_size/bsz^2) × bsz^2;
(9) if gwpid < nbr then
(10)   if rpos < range then
(11)     r = rpos%bsz;
(12)     c = (rpos/bsz)%bsz;
(13)     InvStIdx = NLColPtr[gwpid];
(14)     InvEdIdx = NLColPtr[gwpid + 1];
(15)     for idx = InvStIdx to InvEdIdx − 1 do
(16)       irow = NLRowVal[idx];
(17)       if rpos < bsz^2 then
(18)         tmpidx = LRowPtr[irow + 1] − 1;
(19)         ssol[wpid][r][c] = L[tmpidx ×bsz^2 + rpos];
(20)         inv[wpid][r][c] = 1.0/det ×inverse(ssol)_{r,c};
(21)         sm[wpid][rpos] = Rhs[idx ×bsz^2 + rpos];
(22)         ssol[wpid][r][c] = ∑_{t=0}^{bsz−1} inv[wpid][r][t] × sm[wpid][c × bsz + t];
(23)         NL[idx ×bsz^2 + rpos] = ssol[wpid][r][c];
(24)       end if
(25)       myidx = idx + 1 + (rpos/bsz^2);
(26)       triEdIdx = triStIdx + (n − icol) ×bsz^2;
(27)       solIdx = solStIdx + bsz^2 + rpos;
(28)       while myidx < InvEdIdx do
(29)           sm[wpid][rpos] = 0.0;
(30)           myrow = NLRowVal[myidx];
```

Algorithm 2: Continued.

```
(31)              tmpidx = LRowPtr[myrow + 1] − 1;
(32)              while tmpidx ≥ LRowPtr[myrow] && LColVal[tmpidx] == irow do
(33)                tmpidx−−;
(34)              end while
(35)              if tmpidx ≥ LRowPtr[myrow] && LColVal[tmpidx] == irow then
(36)                s[wpid][lane] = L[tmpidx ×bsz² + rpos%bsz²];
(37)              end if
(38)            Rhs[myidx ×bsz² + lane%bsz²]− = ∑_{t=0}^{bsz−1} sm[wpid][(lane/bsz²) × bsz² + t × bsz + r] × ssol[wpid][t][c];
(39)            myidx+ = warp_size/bsz²;
(40)          end while
(41)        end for
(42)      end if
(43)  end if
```

Algorithm 2: GPU kernel for solving LGSBTS.

dispatched by CUDA (compute unified device architecture), for solving of a SBTS, the total number of launched threads is $32 \times \text{nbr}$ where nbr is the number of block rows (columns) of the matrix. The total number of CUDA blocks is calculated by ceil $((32 \times \text{nbr})/\text{threads\_per\_block})$. By default, the one-dimensional block organization is employed, but when the number of blocks exceeds the maximum number of blocks along one-dimension provided by the CUDA architecture, the two-dimensional block arrangement is conducted.

To identify which warp corresponds to the desired column of NL, a global thread identifier tid is first transformed to the global warp identifier gwpid and relative warp identifier wpid in a thread block (lines 5 and 6). Then, the local index within a warp for a thread, denoted rpos, can be obtained by threadId$x.x$%warp_size. For a block size $\text{bsz}\,(\text{bsz} \le 5)$, the maximum number of complete blocks a warp can cover at a time is floor $(\text{warp\_size}/\text{bsz}^2)$, and the maximum number of active threads in a warp is estimated by $\text{range} = \text{floor}\,(\text{warp\_size}/\text{bsz}^2) \times \text{bsz}^2$. As a result, the following task of the kernel (from lines 10 to 42) is limited in range threads. As illustrated in Figure 3, the warp goes through the row indices (J_l) of the gwpid$^{\text{th}}$ block column of NL. For each row index irow, the first task is to extract the blocks at the J_l rows in the irow$^{\text{th}}$ block column of L. Since L is expressed in BCSR format, the irow$^{\text{th}}$ block column cannot be accessed directly and continuously. Moreover, a warp may not be able to cover all blocks in the column simultaneously, and the type of computation on the first (diagonal) block is different from that on the remaining blocks in the column. By considering these issues, the task of going through a column of L is divided into two subtasks.

First, the first bsz^2 threads solve a block of the solution by performing a block-block multiplication (lines 17–24). To exploit the fine-grained feature of the GPU, the computing of only one element for the result of multiplication is assigned to a thread. The thread corresponding to the r^{th} row and c^{th} column of the result is responsible for the inner product between the r^{th} row of the left-hand block and c^{th} column of the right-hand block. Because values in the two blocks are repeatedly visited, the multiplication is implemented through the shared memory, where the two operating blocks are loaded to avoid repeated access from the global memory. Note that the first block (diagonal block of a SBTM) is inversed by using the symbolic expression of the adjoint matrix of a block. For example, letting $B = \{b_{r,c}, r, c = 0, 1, 2, 3\}$ be a 4×4 block, then the r^{th} row and c^{th} column of the adjoint matrix of B, denoted as $B^*_{r,c}$, can be expressed explicitly as

$$B^*_{r,c} = (-1)^{r+c} \begin{vmatrix} b_{(c+1)\%4,(r+1)\%4} & b_{(c+1)\%4,(r+2)\%4} & b_{(c+1)\%4,(r+3)\%4} \\ b_{(c+2)\%4,(r+1)\%4} & b_{(c+2)\%4,(r+2)\%4} & b_{(c+2)\%4,(r+3)\%4} \\ b_{(c+3)\%4,(r+1)\%4} & b_{(c+3)\%4,(r+2)\%4} & b_{(c+3)\%4,(r+3)\%4} \end{vmatrix}, \tag{10}$$

and therefore, each thread of the bsz^2 threads is only computing one element of the adjoint matrix. The symbolic expressions for $\text{bsz} = 3$ and $\text{bsz} = 5$ can be derived in a similar way.

Second, the entire warp executes (9) by going over all remaining blocks in the column (lines 25–40). In the circumstance in which a warp is unable to cover all of the remaining blocks at one time, several cycles are performed until the end block is reached. In each cycle, before loading blocks into the available shared memory, every bsz^2 threads search myrow from right to left to identify whether there is a non-zero block at the irow$^{\text{th}}$ column. The block is set to a zero block if the searching fails. Once the searching is completed, up to floor $(\text{warp\_size}/\text{bsz}^2)$ blocks update the right-hand side by block-block multiplications.

The implementation for the preconditioning step comprises two block matrix-vector multiplications that can be realized by calling *cusparseDbsrmv* in the cuSPARSE library. Compared to performing triangular solves in a traditional way, the block matrix-vector multiplication offers more possible parallelism that may greatly reduce the computing overhead of the preconditioning step in each iteration.

4. Experiments

4.1. Experiment Setup. A computing node of the GPU cluster was used for all experiments. The node is configured with two Intel E5-2640 V4@2.40 GHz CPUs, with 128 GB of memory and four Nvidia Tesla V100 cards. The Tesla V100

card, with a compute capability of 7.5, is configured by 4096-bit HBM2 16 GB memory and has 80 stream multiprocessors (SMs) with 5120 FP32 cores, 2560 FP64 cores, and 640 Tensor cores in total. The peak single-precision and double-precision floating point performance is 15.7 and 7.8 TFLOPS, respectively.

The algorithms presented in this study are implemented based on the PETSc-3.14.2 (portable, extensible toolkit for scientific computation) framework [21] and CUDA 10.0. The GPU version of the GMRES algorithm with restart 30, developed for block matrices based on PETSc's data structure and cuSPARSE library, is used as the test solver.

For each test case, comparisons are made between different implementations for the preconditioning step. In the first implementation, the preconditioning step is realized using PETSc's block triangular solver on the CPU, and the results are copied to the GPU for the remaining part of computing by the GMRES. The second implementation uses parallel block triangular solves on the GPU from the cuSPARSE library. The input for the cuSPARSE's triangular solver is obtained by calling the ILU factorization routine in BCSR format. The last implementation is the block-ISAI preconditioning algorithm on the GPU as proposed in the present study. The lower and upper block triangular factors are provided by the block ILU factorization in PETSc and used as the input of Algorithm 1. For convenience, the wall clock time of computing the preconditioning step for 30 iterations is denoted as T_p, the total computing time until the GMRES converges to the relative tolerance is denoted as T_{total}, and the total number of GMRES iterations is denoted as G_{it}.

4.2. Atmosmodd Case. The first test matrix, atmosmodd, is from the SuiteSparse matrix collection [22]. It consists of 1,270,432 rows and columns and 8,814,880 non-zeros. The matrix is transformed into the BCSR format with block size 4 using cuSPARSE library [6]. As a result, the block matrix consists of 317,608 block rows and columns and 2,190,844 non-zero blocks. The relative tolerance for GMRES convergence is set to 10^{-8}.

Table 1 provides the time costs for 30 preconditioning steps and the GMRES (30) method, as well as the total number of GMRES iterations using different preconditioning methods. It is reported in [3] that the sparsity pattern for atmosmodd provides valid parallelism and make the level-scheduling algorithm on the GPU gain accelerations over serial triangular solvers on the CPU. It is observed from Table 1 that the level-scheduling scheme in block format is also more efficient than the serial block triangular solves implemented in PETSc, and approximately 12.4x speedup can be achieved. The preconditioning costs decrease significantly when block-ISAI ($|X|$) and block-ISAI ($|X|^2$) are applied, which is attributed to the high parallelism provided by the block sparse matrix-vector multiplication in the proposed algorithm. The resulting G_{it}, however, is more than that produced by exact triangular solves using PETSc and cuSPARSE. This is because we only compute an approximate inverse of the lower and upper block triangular matrices based on a guessing sparsity pattern.

Table 1: Atmosmodd case: T_p, T_{total}, and G_{it} for different preconditioning methods.

-	PETSc	cuSPARSE	Block-ISAI ($\|X\|$)	Block-ISAI ($\|X\|^2$)
T_p	1072 ms	86 ms	19.7 ms	44 ms
T_{total}	7654 ms	859 ms	594 ms	685 ms
G_{it}	201	201	295	242

The overhead consisting of three parts for setting up block-ISAI($|X|$) and block-ISAI ($|X|^2$) is shown in Figure 4. As described in Algorithm 1, NL and NU are solved in columns, and thus, the BCSR to BCSC transformations are required for NL and NU after the estimations of sparsity patterns have been made. Once the block values in NL and NU are filled by the GPU kernel, they are transformed into BCSR formats for the block matrix-vector multiplications in the preconditioning step. Therefore, four format transformations in total have to be conducted. Compared to using block-ISAI ($|X|$), using block-ISAI ($|X|^2$) results in a lower number of GMRES iterations but more computing time for the GMRES algorithm and constructing an effective preconditioner. Although the overhead of generating block-ISAI increases with making the sparsity pattern denser, it only costs 2.5% and 5.8% compared to T_{total} for block-ISAI ($|X|$) and block-ISAI ($|X|^2$), respectively.

By adding the overhead of setting up block-ISAI, the overall times for the above four methods, as well as the speedups over the baseline GPU GMRES (30) with PETSc's preconditioning are shown in Figure 5. A speedup of 12.6x is obtained by block-ISAI ($|X|$) over the baseline GMRES (30), which is the best among the listed methods. Compared to cuSPARSE, both versions of block-ISAI are faster.

4.3. af_shell3 Case. The second matrix, af shell3, is also from the SuiteSparse matrix collection [22]. The size of the matrix is $504,855 \times 504,855$ with 17,588,875 non-zeros. By setting the block size to 5, the matrix is transformed into the block format. The number of block rows and columns is 100,971 and that of non-zero blocks is 703,555. The relative tolerance for GMRES (30) is set to 10^{-8}.

In this case, $S(|X|)$, $S(|X|^2)$, and $S(|X|^3)$ are considered for the estimation the inverses of L and U. Table 2 provides the comparisons of T_p, T_{total}, and G_{it} obtained by the five methods. Note that the PETSc, cuSPARSE, and block-ISAI ($|X|^3$) methods result in less than 30 GMRES iterations, so the corresponding T_p values are reported using the actual number of iterations; otherwise, T_p for 30 iterations is reported. It is observed that the exact triangular solver from cuSPARSE is efficient in extracting possible parallelism from the sparsity pattern, and it runs approximately 2.55x faster than the PETSc's serial block triangular solver. However, the preconditioning overhead of cuSPARSE, T_p, still accounts for 88% of the total computing time. In contrast, the block-ISAI method results in much less time for T_p, but at the sacrifice of increasing the total number of GMRES iterations. In addition, G_{it} decreases when the estimated patterns for inverses become denser.

By counting the overhead of setting up block-ISAI in Figure 6, the overall time and speedups are shown in

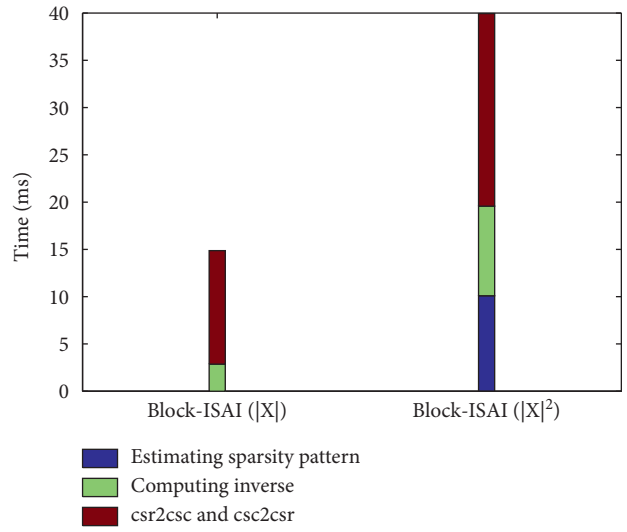


Figure 4: Atmosmodd case: breakdowns of costs for setting up block-ISAI $|X|$ and block-ISAI $|X|^2$.

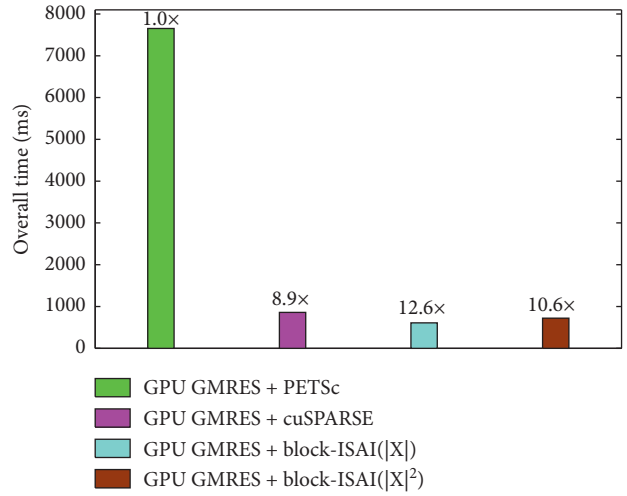


Figure 5: Atmosmodd case: comparisons of overall times between different preconditioning methods.

Table 2: af_shell3 case: T_p, T_{total}, and G_{it} for different preconditioning methods.

-	PETSc	cuSPARSE	Block-ISAI ($\|X\|$)	Block-ISAI ($\|X\|^2$)	Block-ISAI ($\|X\|^3$)
T_p	435 ms	170 ms	9.03 ms	15.1 ms	22.0 ms
T_{total}	470 ms	193 ms	40.1 ms	35.5 ms	39.7 ms
G_{it}	25	25	45	29	26

Figure 7. Although the costs of block-ISAI are comparable to T_{total}, the overall times obtained by block-ISAI methods are still much less than those in the PETSc's and cuSPARSE implementations. Figure 7 indicates that all three block-ISAI with different sparsity patterns can achieve more than 5x over PETSc's implementation. Block-ISAI with $S(|X|)$ is the most efficient in terms of the overall time and is nearly 3.9x faster than the cuSPARSE implementation.

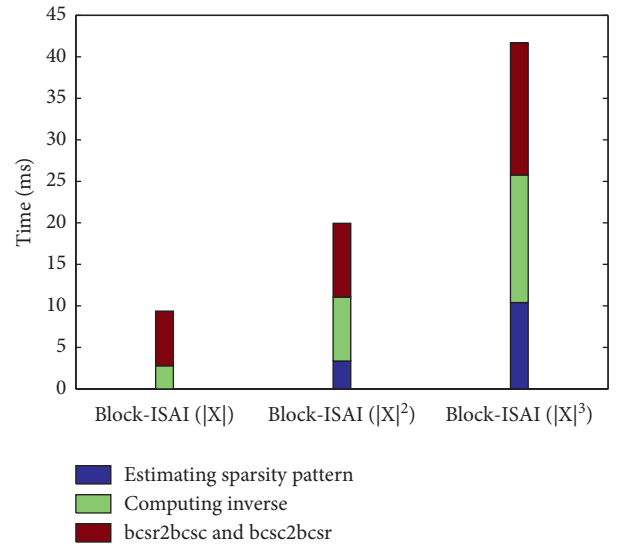


Figure 6: af shell3 case: breakdowns of costs for setting up block-ISAI ($|X|$), block-ISAI ($|X|^2$), and block-ISAI ($|X|^3$).

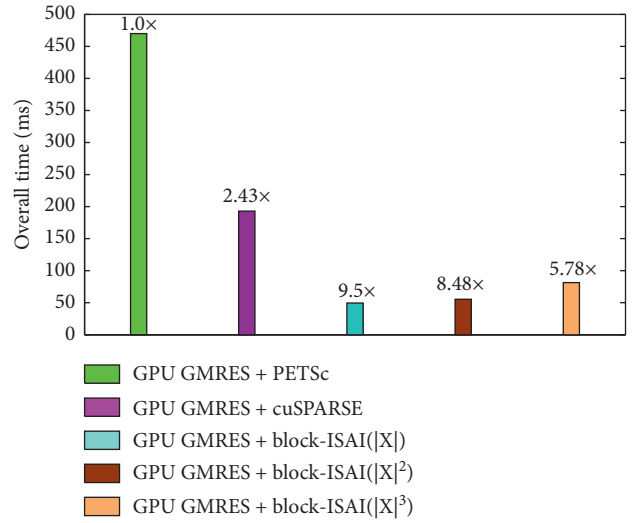


Figure 7: af_shell3 case: comparisons of overall times between different preconditioning methods.

4.4. Lid-Driven Cavity Case. In the third experiment, a linear system with the block matrix from a lid-driven cavity case is solved on a 300×300 mesh in the velocity-vorticity formulation using the five-point finite difference method. The block matrix, with a block size of 3, is extracted from the first Newton step of the nonlinear solver. The number of block rows and columns is 90,000 and that of non-zero blocks is 448,800. The relative tolerance is set to 10^{-5} for GMRES (30).

It is observed from Table 3 that cuSPARSE on the Tesla V100 fails to accelerate the preconditioning step that is composed of two block triangular systems. This is due to the strong data dependency in the sparsity pattern. Block-ISAI, in contrast, can offer a better solution in this circumstance. Eventhough the number of GMRES iterations using block-ISAI is 14%, 34%, and 75% more than that using exact

Table 3: Lid-driven cavity case: T_p, T_{total}, and G_{it} for different preconditioning methods.

-	PETSc	cuSPARSE	Block-ISAI ($\|X\|$)	Block-ISAI ($\|X\|^2$)	Block-ISAI ($\|X\|^3$)
T_p	130 ms	129 ms	3.72 ms	5.20 ms	6.88 ms
T_{total}	3150 ms	3074 ms	560 ms	468 ms	443 ms
G_{it}	623	623	1092	836	711

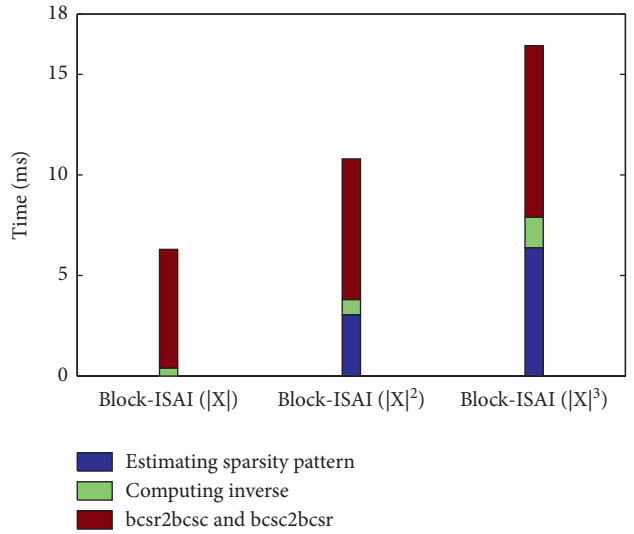


Figure 8: Lid-driven cavity case: breakdowns of costs for setting up block-ISAI ($|X|$), block-ISAI ($|X|^2$), and block-ISAI ($|X|^3$).

triangular solves, block-ISAI preconditioning is still more efficient in terms of T_{total}. In Figure 8, the breakdown costs of block-ISAI with different sparsity patterns indicates that compared to T_{total}, the overhead is only up to 4%, which keeps the cost from affecting the overall time.

Figure 9 shows the comparisons of overall times obtained by PETSc, cuSPARSE, and block-ISAI. Block-ISAI preconditioning shows an obvious advantage over the serial and parallel preconditioning from PETSc and cuSPARSE. It is further observed that both T_{total} and G_{it} obtained by block-ISAI preconditioning decreases upon increasing the density level of the sparsity pattern. GMRES (30) using block-ISAI ($|X|^3$) preconditioning is approximately 6.9x faster than that using PETSc or cuSPARSE preconditioning.

4.5. Venkat50 Case. The last matrix is from the unstructured two-dimensional Euler solver provided by Venkatakrishnan at the SuiteSparse matrix collection [22]. The fluid fields associated with mesh points are ordered continuously so that the resulting matrix has a block structure naturally, with a block size of 4, which can be viewed in Figure 10. The matrix has 15,606 block rows and columns with 107,362 non-zero blocks in total. The relative tolerance is 10^{-8} for the GMRES (30) solver.

It is given in Table 4 that the preconditioning (triangular solver) using cuSPARSE is nearly 4 times slower than the serial ones provided by PETSc. This is normal because the level scheduling-based scheme cannot always be efficient on

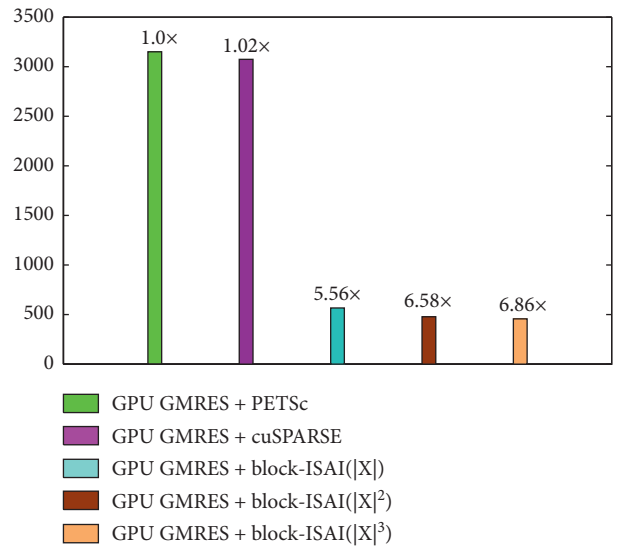


Figure 9: Lid-driven cavity case: comparisons of overall times between different preconditioning methods.

the GPU when slight parallelism is extracted from a given sparsity pattern [3]. Compared to T_{total} obtained by PETSc, T_{total} using block-ISAI ($|X|$) does not show an obvious advantage because the approximate inverses lack of accuracy and result in 6 times more GMRES iterations in G_{it}. Upon increasing the density of the sparsity pattern, the decrease of both G_{it} and T_{total} is observed, and block-ISAI ($|X|^3$) is

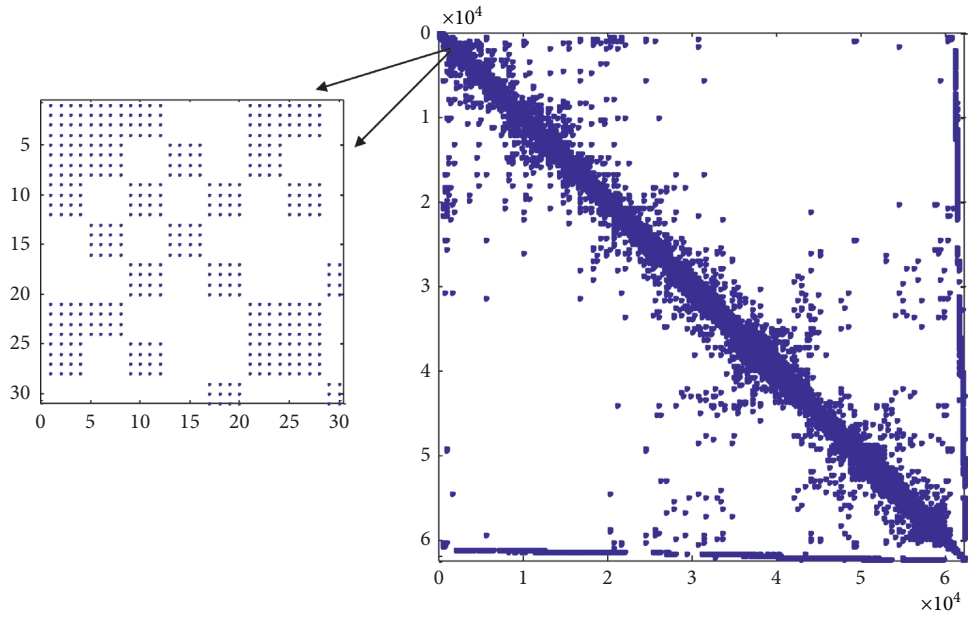


Figure 10: Sparsity pattern for venkat50 matrix with natural block structure.

Table 4: Venkat50 case: T_p, T_{total}, and G_{it} for different preconditioning methods.

-	PETSc	cuSPARSE	Block-ISAI ($\|X\|$)	Block-ISAI ($\|X\|^2$)	Block-ISAI ($\|X\|^3$)
T_p	50 ms	192 ms	1.96 ms	3.02 ms	4.25 ms
T_{total}	734 ms	2556 ms	678 ms	458 ms	369 ms
G_{it}	367	367	2208	1324	948

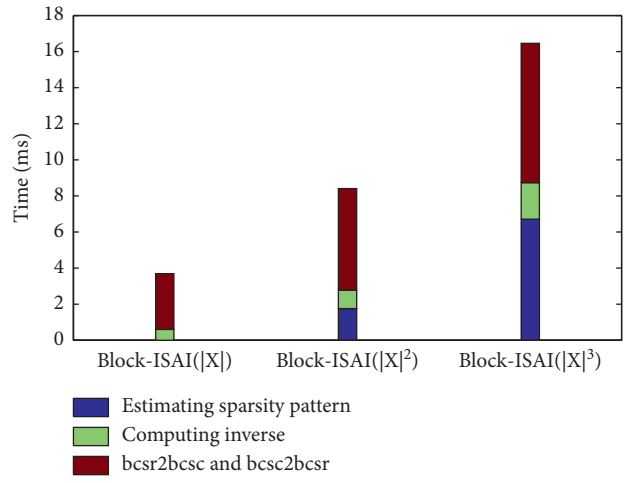


Figure 11: Venkat50 case: breakdowns of costs for setting up block-ISAI ($|X|$), block-ISAI ($|X|^2$), and block-ISAI ($|X|^3$).

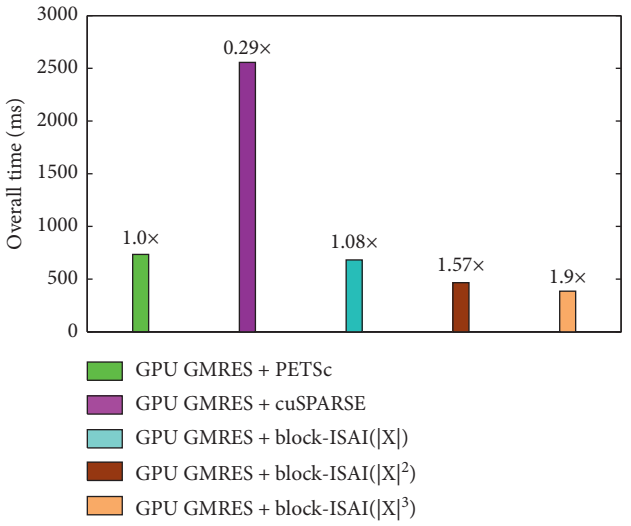


Figure 12: Venkat50 case: comparisons of overall times between different preconditioning methods.

shown to be the best preconditioner in term of T_{total}. The breakdown of costs for computing block-ISAI upon increasing the dense levels is shown in Figure 11. The overall speedups on GMRES are calculated and shown in Figure 12 by counting the overhead for computing block-ISAI, and speed improvements of approximately 1.92x and 6.69x are achieved compared to PETSc's and cuSPARSE's preconditioning, respectively.

5. Conclusions

In this study, block-ISAI preconditioning for block sparse matrices is investigated by proposing an efficient, warp-based algorithm to approximate the inverses of block triangular factors on a Tesla V100 GPU. Instead of solving triangular systems globally for preconditioning, a group of

small triangular systems with very high concurrency is solved to approximate the inverses and conduct block SpMV in the preconditioning step. The results on four cases from a matrix collection website and multiphysics areas show that the proposed GPU algorithm outperforms the serial and parallel block triangular solver-based preconditioning from the PETSc and cuSPARSE libraries in terms of the total computing time of the GMRES (30) algorithm. It is noted that all comparisons are performed under the condition that the L and U factors are provided because the target for comparison is the block triangular solves in the preconditioning step. Planned future work includes extending the algorithm to multiple GPUs that are connected by the NVLink technology using block-Jacobi or the additive Schwarz method (ASM) preconditioners, as well as the coupling of sparsity patterns and relaxation steps [10] for extreme cases.

Data Availability

The main matrix data used to support the findings of this study are available from the SuiteSparse matrix collection (https://sparse.tamu.edu/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by a grant from the National Key R&D Program of China (2019YFB1704202), the National Natural Science Foundation of China (61702438), and the Nanhu Scholar Program of XYNU and the Innovation Team Support Plan of University Science and Technology of Henan Province (19IRTSTHN014). The authors thank LetPub for its linguistic assistance during the preparation of this manuscript.

References

[1] Y. Saad, *Iterative Methods for Sparse Linear Systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2 edition, 2003.

[2] M. Naumov, *Parallel Solution of Sparse Triangular Linear Systems in the Preconditioned Iterative Methods on the GPU*, NVIDIA Technical Report NVR-2011-001, NVIDIA Corporation, Santa Clara, CF, USA, 2011.

[3] R. Li and Y. Saad, "GPU-accelerated preconditioned iterative linear solvers," *The Journal of Supercomputing*, vol. 63, no. 2, pp. 443–466, 2013.

[4] H. Kabir, J. D. Booth, G. Aupy, A. Benoit, Y. Robert, and P. Raghavan, "STS-K: a multilevel sparse triangular solution scheme for NUMA multicores," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Austin, TX, USA, November 2015.

[5] J. Park, M. Smelyanskiy, N. Sundaram, and P. Dubey, "Sparsifying synchronization for high-performance shared-memory sparse triangular solver," in *Proceedings of International Supercomputing Conference*, Springer International Publishing, Leipzig, Germany, June 2014.

[6] NVIDIA cuSPARSE, https://docs.nvidia.com/cuda/archive/10.0/cusparse/index.html. (accessed on 30 December 2020).

[7] W. Liu, A. Li, J. Hogg, I. S. Duff, and B. Vinter, "A synchronization-free algorithm for parallel sparse triangular solves," *Euro-Par 2016: Parallel Processing*, vol. 9833, pp. 617–630, 2016.

[8] W. Liu, A. Li, J. D. Hogg, I. S. Duff, and B. Vinter, "Fast synchronization-free algorithms for parallel sparse triangular solves with multiple right-hand sides," *Concurrency and Computation: Practice and Experience*, vol. e4244, pp. 1–18, 2017.

[9] H. Anzt, E. Chow, and J. Dongarra, "Iterative sparse triangular solves for preconditioning," in *Proceedings of European Conference on Parallel Processing*, Springer, Berlin, Heidelberg, August 2015.

[10] H. Anzt, T. K. Huckle, J. Bräckle, and J. Dongarra, "Incomplete sparse approximate inverses for parallel preconditioning," *Parallel Computing*, vol. 71, pp. 1–22, 2018.

[11] H. Anzt, E. Chow, T. Huckle, and J. Dongarra, "Batched generation of incomplete sparse approximate inverses on GPUs," in *Proceedings of Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems*, IEEE, Salt lake, UT, USA, November 2016.

[12] E. Chow and A. Patel, "Fine-grained parallel incomplete LU factorization," *SIAM Journal on Scientific Computing*, vol. 37, no. 2, pp. C169–C193, 2015.

[13] E. Chow, H. Anzt, and J. Dongarra, "Asynchronous iterative algorithm for computing incomplete factorizations on GPUs," *Lecture Notes in Computer Science*, vol. 9137, pp. 1–16, 2015.

[14] Intel Many Integrated Core Architecture (Intel MIC Architecture)-Advanced. https://www.intel.com/content/www/us/en/architecture-and-technology/many-integrated-core/intel-many-integrated-core-architecture.html. (accessed on 31 December 2020).

[15] L. Y. Kolotilina and A. Y. Yeremin, "Factorized sparse approximate inverse preconditionings I. Theory," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 1, pp. 45–58, 1993.

[16] A. C. N. V. Duin, "Scalable parallel preconditioning with the sparse approximate inverse of triangular matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 4, 1999.

[17] D. Bertaccini and S. Filippone, "Sparse approximate inverse preconditioners on high performance GPU platforms," *Computers & Mathematics with Applications*, vol. 71, no. 3, pp. 693–711, 2016.

[18] G. He, R. Yin, and J. Gao, "An efficient sparse approximate inverse preconditioning algorithm on GPU," *Concurrency and Computation Practice and Experience*, vol. 32, no. 3, 2019.

[19] J. Gao, Q. Chen, and G. He, "A thread-adaptive sparse approximate inverse preconditioning algorithm on multi-GPUs," *Parallel Computing*, vol. 101, no. 3, Article ID 102724, 2021.

[20] R. Eberhardt and M. Hoemmen, "Optimization of block sparse matrix-vector multiplication on shared-memory parallel architectures," in *Proceedings of 2016 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, IEEE, Chicago, IL, USA, May 2016.

[21] Balay S.; Abhyankar S.; Adams M. F.; et al. PETSc Web Page. https://www.mcs.anl.gov/petsc.(accessed on 30 December 2020).

[22] T. A. Davis and Y. Hu, "The university of Florida sparse matrix collection," *ACM Transactions on Mathematical Software*, vol. 38, no. 1, pp. 1–25, 2011.

[23] Y. Saad, M. H. Schultz, and M. H. Schultz, "GMRES: a generalized minimal residual algorithm for solving non-symmetric linear systems," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 3, pp. 856–869, 1986.

[24] B. Yang and H. Liu, "Accelerating the GMRES solver with block ILU(k) preconditioner on GPUs in reservoir simulation," *Journal of Geology and Geophysics*, vol. 4, no. 2, pp. 1–7, 2015.

[25] A. Kashi and S. Nadarajah, *Fine-grain Parallel Smoothing by Asynchronous Iterations and Incomplete Sparse Approximate Inverses for Computational Fluid Dynamics*, AIAA, Orlando, FL, USA, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6624330, 8 pages
https://doi.org/10.1155/2021/6624330

Research Article

Adaptive Configuration Method of Low-Frequency Electromechanical Sampling Information in Building Electrical System

Bai Yu and Zhixin Zhang

College of Information, Tianjin University of Commerce, Tianjin 300134, China

Correspondence should be addressed to Bai Yu; yubai2020@21cn.com

Received 3 January 2021; Revised 9 April 2021; Accepted 6 May 2021; Published 20 May 2021

Academic Editor: Hussein Abulkasim

In order to enhance the reliability of the electrical systems in low frequency, an adaptive configuration method of low-frequency electromechanical sampling information based on thyristor controlled series compensation (TCSC) is designed. The electrical system is simplified to a linear invariant system, and a stochastic subspace identification (SSI) method is established by using the singular value decomposition principle to collect low-frequency electromechanical sampling information. The reference channel technology is introduced to design the reference channel covariance matrix to judge whether low-frequency information is generated and improve the efficiency of sampling information acquisition. The architecture and working principle of the controllable series compensation device are analyzed, and the test signal method is used to evaluate the low-frequency modes and the information required by the device among the electrical system regions of buildings. The alternative input signals are selected by comparing different input signal residue ratios. The TCSC device parameters are adjusted by the compensation residue phase method, so as to realize the adaptive configuration of different low-frequency electromechanical sampling information and ensure the stable operation of the electrical system. The experimental results show that the proposed configuration scheme can effectively improve the damping ratio of the system and has an excellent effect on suppressing the continuous oscillation under a low-frequency fault.

1. Introduction

The building electrical system includes key content such as building power distribution, indoor lighting, outdoor landscape lighting, and management automation [1]. With the improvement of building quality, the structure of the building electrical system has become increasingly complex, and the interconnection between different areas has gradually become dense. Sampling in low-frequency electromechanical systems can obtain the electromechanical oscillation conditions based on the sampled information. The oscillation with a frequency range of 0.2~3 Hz belongs to low-frequency oscillation [2, 3]. During the oscillation process, the rotors of the generators participating in the oscillation will swing relatively, the power transmission of the transmission line will fail, and the continuous increase of the oscillation amplitude will affect the normal operation of the electrical system, even causing the system to disassemble.

For the problem of low-frequency electromechanical oscillation, Yaqi and Zhigang [4] proposed a predictive current control method using a traction grid-side converter model based on a continuous control set in the coordinate system. By discretizing the state equation of the AC side, the mathematical prediction model of the single traction grid-side converter was derived. The performance function was solved, that is, the difference between the predicted current value and the reference value and the change in the control voltage, to calculate the optimal control variable. The bipolar sinusoidal pulse width modulation was fused, to solve the phenomenon of traction blockage caused by low-frequency oscillation. But the method could not extract effective control parameters. Based on the discrete Fourier transform

adaptive notch filter, Shen et al. [5] reduced the harm of continuous vibration. The notch filter instead of the low-pass filter was used to eliminate the deviation of resonance frequency and oscillation frequency, which greatly simplified the suppression process and reduced the frequency of oscillation frequency detection. At the same time, an oscillation detection method was designed, where interval detection and prediction methods were used to shorten the detection time. However, the running time of the notch filter was too long to satisfy the immediacy of low-frequency electromechanical oscillation suppression in the electrical system.

In this research, a TCSC-based adaptive configuration method of low-frequency electromechanical sampling information is proposed. The linear invariant difference equation of the building electrical system is established. Under the premise of the singular value decomposition principle, a low-frequency electromechanical sampling information extraction method based on the improved SSI method is designed to determine whether the low-frequency electromechanical oscillation phenomenon exists. The TCSC device is used to complete the damping control, and the oscillation suppression mode is adaptively adjusted according to different oscillation levels.

2. Low-Frequency Electromechanical Sampling Information Extraction Based on Improved SSI Method

The building electrical system is a large-scale nonlinear system. When exploring system modes and modals, the system is generally simplified to a linear invariant system. The system measurement signals are obtained through time sequence, and the system is described as a set of linear invariant difference equations as shown in equation (1) [6], and the system also meets the convergence condition in equation (2).

$$\begin{cases} x(kT+T) = Ax(kT) + Bu(kT) + w(kT), \\ y(kT) \quad = Cx(kT) + Du(kT) + v(kT), \end{cases} \tag{1}$$

$$E\left[\begin{pmatrix} w(pT) \\ v(qT) \end{pmatrix}\left(w'(pT)v'(qT)\right)\right] = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix}\delta_{pq}, \tag{2}$$

where $x \in R^{n\times 1}$ is the system state column vector, $y \in R^{l\times 1}$ is the system output column vector, $u \in R^{m\times 1}$ is the known system input control column vector, $A \in R^{n\times n}$ is the system state matrix, $B \in R^{n\times m}$ is the system control matrix, $C \in R^{l\times n}$ is the system output matrix, $D \in R^{l\times m}$ is the system control output matrix, T represents the time-step length, k, p, and q are all time-step coding, w is the system input stochastic interference column vector, and v is the system output measurement stochastic deviation column vector.

In most states, the system control input u is an unknown quantity. By including the influence of u in the system stochastic interference and measurement stochastic deviation [7, 8], we get

$$\begin{cases} x(kT+T) = Ax(kT) + w(kT), \\ y(kT) \quad = Cx(kT) + v(kT). \end{cases} \tag{3}$$

Regarding the state matrix A of the differential model, eigenvalue and relative right eigenvector φ_i are obtained by eigen analysis. The eigenvalues of the continuous system model can be obtained by

$$s_i = \frac{\ln(e_i)}{T}. \tag{4}$$

The modal corresponding to mode s_i of each output channel in the output column vector y is

$$V_i = C\varphi_i. \tag{5}$$

The above calculation process can be summarized as follows. Under the action of interference input w and measurement deviation v of approximate white noise, the building electrical system obtains the time sequence of the output signal y in the system through the measurement value and obtains the system state matrix and the output matrix through the relevant information extraction algorithm. The system mode and modal status can be specified according to equations (4) and (5), which in this paper refers to detecting whether the low-frequency electromechanical has oscillation phenomenon.

The SSI method is a commonly used method in the field of system information collection. It is built on the principle of singular value decomposition [9] and has strong numerical stability. Through measurement, the system output time sequence can be obtained as $\{y(0), y(T), \ldots, y((N-1)T), y(NT)\}$.

The output covariance matrix of the time lag iT is

$$\begin{aligned} \Lambda_i &= E\left[(y(k+1)T)y'(kT)\right] \\ &= \lim_{N\longrightarrow\infty} \frac{1}{N}\sum_{k=0}^{N-1}\left((y(k+1)T)y'(kT)\right). \end{aligned} \tag{6}$$

The following block Toeplitz matrix is set:

$$T_{1|i} = \begin{bmatrix} \Lambda_i & \Lambda_{i-1} & \cdots & \Lambda_1 \\ \Lambda_{i+1} & \Lambda_i & \cdots & \Lambda_2 \\ \cdots & \cdots & \cdots & \cdots \\ \Lambda_{2i-1} & \Lambda_{2i-2} & \cdots & \Lambda_i \end{bmatrix}. \tag{7}$$

In the real scene, the length of the sequence is limited [10], and the covariance matrix should be approximated by

$$\Lambda_i \approx \widehat{\Lambda}_i = \frac{1}{N}\sum_{k=0}^{N-1}\left((y(k+1)T)y'(kT)\right). \tag{8}$$

Then, the calculation process of the SSI algorithm is as follows. The time sequence of system output y is obtained through measurement; the covariance matrix $\Lambda_1, \Lambda_2, \ldots, \Lambda_i$ under each time lag is calculated according to equation (8), where i satisfies $li \gg n$, and equation (7) is used to construct the Toeplitz matrix $T_{1|i}$; singular value decomposition is performed on $T_{1|i}$; then, we get

$$T_{1|i} = USV^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 S_1 V_1^T. \quad (9)$$

The extended observability matrix is obtained:

$$Q_i = U_1 S_1^{1/2} Z. \quad (10)$$

The output matrix and state matrix are obtained:

$$\begin{aligned} C &= Q_i(1{:}\ l, :), \\ A &= Q_i(1{:}\ l(i-1), :)^{\dagger} Q_i(l+1{:}\ li, :). \end{aligned} \quad (11)$$

In the SSI algorithm, in order to obtain the modal information of some busbars relative to a certain model in the building electrical system [11], it is necessary to clarify whether there is oscillation information inside the electromechanical system. With the continuous expansion of the system scale, the number of system buses will increase, and the measurement of the buses will be included in the output column vector. The reference channel technology is used to maximize operational efficiency without reducing the accuracy of sampling information extraction.

The concept of output channel originates from the field of signal processing. The concept of the output channel in this paper is the output signal measured by the system. The measurement output series obtained in modal analysis generally contains much redundant information [12, 13], and the amount of calculation can be reduced by reducing the redundant information. The reference channel is a subset $y^{\text{ref}} \in R^r$ of the output column vector, where r is the dimension of the reference channel, and the following equation is obtained:

$$y = \begin{bmatrix} y^{\text{ref}} \\ y^{\sim\text{ref}} \end{bmatrix}, \quad (12)$$

where

$$\begin{aligned} y^{\text{ref}} &= Ly, \\ L &= \begin{bmatrix} I_r & 0 \end{bmatrix}, \end{aligned} \quad (13)$$

where $y^{\sim\text{ref}}$ is the output vector except the reference channel in the initial vector [14], and the reference channel covariance matrix is recorded as

$$\begin{aligned} \Lambda_i^{\text{ref}} &= E\left[(y(k+i)T)\left(y^{\text{ref}}(kT) \right)' \right] \\ &= \lim_{N \longrightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left((y(k+1)T) y^{\text{ref}}(kT)' \right), \end{aligned} \quad (14)$$

where Λ_i^{ref} is the output covariance matrix with the time lag of the reference channel being iT.

Λ_i^{ref} is introduced into equation (7) to form a brand-new matrix $T_{1|i}^{\text{ref}}$, and the selection rule for setting iteration parameter i is $ri \gg n$. After introducing the reference channel, the value of Λ_i^{ref} is calculated, and singular value decomposition is performed. According to equation (5), the current low-frequency motor oscillation condition of all output channels can be calculated.

3. Adaptive Configuration of Low-Frequency Electromechanical Sampling Information Based on TCSC

Through the above content, it is possible to clarify whether there is oscillation information inside the low-frequency electromechanical system. If oscillation information is generated, TVSC can be used to adaptively configure different degrees of oscillation problems to ensure the stability of the building electrical system.

The TCSC device can adjust the reactance value of the series compensation device in a large area [15], which can enhance the power transmission quality of the high-intensity power transmission system, effectively control the power flow of the system, and reduce the voltage deviation. The topology of the device adopts a low-level control design, which can well reduce the generation of low-frequency electromechanical oscillation information.

Damping is the key to damping the low-frequency oscillation information of the device. The control process is presented in Figure 1.

The electrical system input has two parts of voltage and line transmission power. The damping output is formed through processes such as deviation comparison, dead zone control, and phase shifting rectification. Z_{TCSC0} represents the command impedance issued by the personal computer, ΔZ_1 is the damping impedance of power loop, ΔZ_2 represents the damping impedance of voltage loop, and Z_{TCSC} is the output impedance of final stage. The performance and parameters of each link are defined as follows.

V is the positive sequence voltage monitoring value of reference point, V_0 is the voltage reference value of reference point, that is, the steady value before oscillation occurs, P_{L0} is the transmission power reference value of the transmission channel where the device is located, and P_L is the power transmission monitoring value of the transmission channel. The deviation comparison has the characteristics of first-order inertia [16], its time constant presents the characteristics of the real measurement device, and the transfer function of the deviation comparison link is $K_m/(1 + T_m S)$.

The dead zone is to prevent the device from oscillating under a small stochastic interference and plays a certain protective role. DC blocking reset is a high-pass filter that can prevent the generation of DC components in the TCSC device's control input signal after low-frequency oscillations, thereby reducing its negative impact on the oscillation suppression effect. Phase shifting is the core content of damping control, allowing the capacitive reactance of the device to obtain the phase difference between the control input signal and the control input signal with the change of the control input signal [17]. If the appropriate phase difference is selected, the change in the capacitive reactance of TCSC will fit the change of the phase difference, which helps adaptively obtain an additional matrix close to the phase of the speed deviation and deal with the low-frequency oscillation of the building electrical system, with the transfer function being $(1 + T_1 S)/(1 + T_2 S)$. The transient stability control

link is a control process that improves the transient stability of the electrical system. After the oscillation is eliminated, the equivalent reactance of the device is immediately adjusted to a fixed value for forced compensation to reduce the electrical distance between the transmitting and receiving ends. In the protection limit link, since the capacitive reactance of the TCSC can only be adjusted in a fixed range, the adjustment range depends on the performance of the capacitor bank, the overvoltage protection level, and the current through the TCSC. The protection limit link can take reasonable constraints on the command reactance obtained by the transient stability control and damping control [18], thereby guaranteeing the normal operation of the protection device.

In order to meet the diversified operation mode of the building electrical system, the input signal of the TCSC device should be selected reasonably to improve the robustness of the adaptive configuration. When the device is not connected, the test signal is used to obtain the open-loop transfer function from the anti-group reference signal to the input signal:

$$G(s) = \frac{\Delta B(s)}{\Delta X_0(s)} = \sum_{i=1}^{n} \frac{R_i}{s - \lambda_i}, \tag{15}$$

where λ_i is the pole of $G(s)$ and R_i represents the residue relative to λ_i.

The transfer function clarifies the controllable and observable conditions in the electrical system. Therefore, the movement of the pole λ_i in $G(s)$ in the signal B is observable. Assuming λ_i is the pole of the low-frequency oscillation information mode, its relative residue R_k can be used as an observable indicator. Calculating $|R_k|$ can clarify the change of observable level with the change of different operating modes, and the residue ratio is used as the indicator to weigh the observability of the alternative input signal of the TCSC device [19]. The residue is expressed as

$$\rho_k = \frac{|R_{k,BI}|}{|R_{k,B0}|}, \tag{16}$$

where k represents the low-frequency electromechanical regional oscillation information code and $|R_{k,BI}|$ and $|R_{k,B0}|$ represent the digital modulus of the first regional oscillation information in the adaptive configuration operation mode and the digital modulus of the first regional oscillation information in the standard operation mode, respectively. In different operation modes, the less the residue ratio changes, the better the robustness of the input signal to the system operation.

TCSC device parameters can be adjusted using the compensation residue phase algorithm [20]. The device is expressed in the form of a transfer function:

$$\mathrm{KH}(s) = \frac{K}{1 + sT_1} \cdot \frac{sT_w}{1 + sT_w} \cdot \left[\frac{1 + sT_2}{1 + sT_3}\right]^m. \tag{17}$$

The adaptive configuration parameters of the phase shifting link are solved by

$$\begin{cases} \phi = 180^\circ - \arg R_k, \\ \alpha = \dfrac{T_3}{T_2} = \dfrac{[1 - \sin(\phi/m)]}{[1 + \sin(\phi/m)]}, \\ T_2 = \dfrac{1}{(\omega_k \sqrt{\alpha})}, \\ T_3 = \alpha T_2. \end{cases} \tag{18}$$

Now, the whole process of the adaptive configuration of low-frequency electromechanical sampling information is completed, which greatly reduces the probability of low-frequency electromechanical failures and provides a solid foundation for the stable operation of the building electrical system.

4. Simulation Experiment

The power system analysis integrated program, i.e., Power System Analysis Software Package (PSASP), was used to study the damping characteristics of the building electrical system without a TCSC device. The operation mode was that A transmitted 1000 MW active power to B, and the transmission result had 4 dominant low-frequency oscillation modes, i.e., 3 inter-regional oscillations and 1 intra-regional oscillation, indicating the unstable state of the system.

In order to prove that the proposed method could effectively improve the damping of the building electrical system, a TCSC device was installed in the system, and a 6% step interference signal was added to the excitation voltage amplitude of the system. Adaptive configuration optimization was conducted on the system's low-frequency electromechanical information and its circuit. The simulation results are shown in Figures 2 and 3. Table 1 shows the dataset showing comparison of low-frequency speed information of building electrical system. Table 2 shows the dataset showing comparison of low-frequency power information of link circuit.

As can be seen from Figures 2 and 3, the installed TCSC devices could effectively suppress the generation of low-frequency information. This was because the proposed method provided sufficient damping for each low-frequency mode and alleviated the adverse effects caused by the oscillation of the low-frequency electromechanical system.

In order to verify the ability of the TCSC-based adaptive configuration method on suppressing low-frequency information in the state of large interference, three short-circuit faults were set near the system's link circuit, and the line was disconnected after 0.2 s; then, the output power of the system's line is shown in Figure 4.

It can be seen from Figure 4 that the system could maintain stable operation in the operation mode of a single link circuit before and after using the adaptive configuration device, but the TCSC device showed a better optimization effect. The advantage of the TCSC device was that it used the improved SSI method to determine whether the system had low-frequency information. As a result, the overall status of

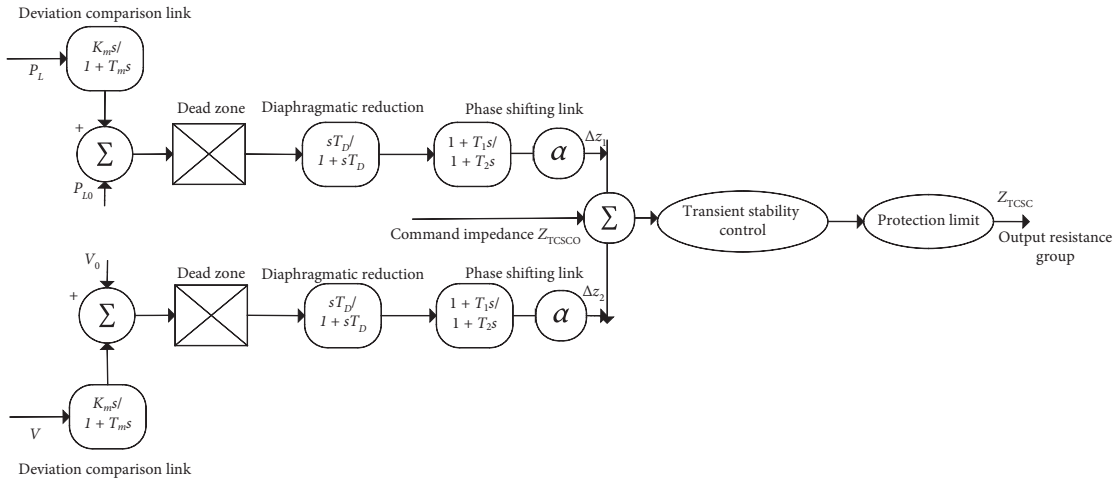


Figure 1: Schematic diagram of the damping control process.

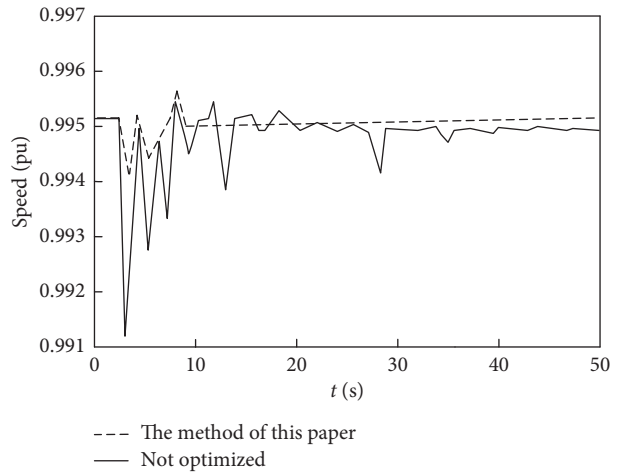


Figure 2: Comparison of low-frequency speed information of building electrical system.

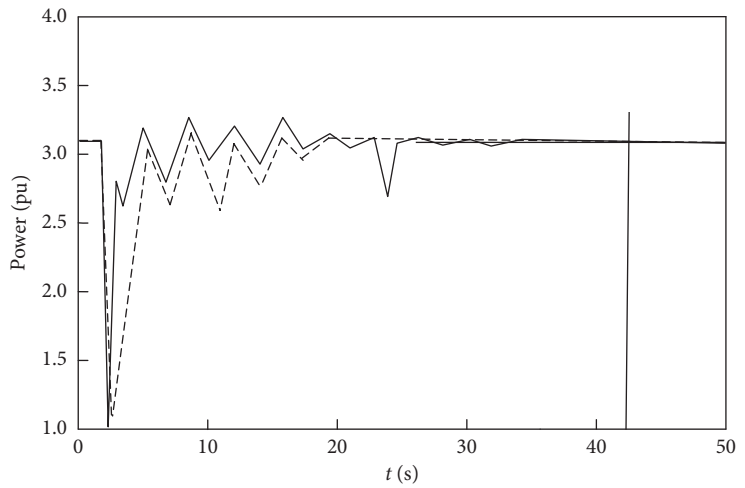


Figure 3: Comparison of low-frequency power information of link circuit.

Table 1: Dataset showing comparison of low-frequency speed information of building electrical system.

S.no.	Time (s)	Not optimized Speed (pu)	Proposed method Speed (pu)
1	0	0.9951	0.9951
2	0.5	0.9951	0.9951
3	1	0.9951	0.9951
4	3	0.9912	0.9940
5	3	0.9942	0.9940
6	4	0.9951	0.9942
7	5	0.9935	0.9945
8	6	0.993	0.9946
9	7	0.9948	0.9947
10	8	0.9935	0.9949
11	9	0.9955	0.9951
12	10	0.994	0.9952
13	11	0.9955	0.9952
14	13	0.994	0.9952
15	15	0.99508	0.9952
16	16	0.9951	0.9952
17	17	0.9949	0.9952
18	19	0.9951	0.9952
19	20	0.9949	0.9952
20	22	0.995	0.9952
21	25	0.9949	0.9952
22	29	0.9942	0.9952
23	30	0.9949	0.9952
24	35	0.9948	0.9952
25	40	0.9949	0.9952
26	50	0.9949	0.9952

Table 2: Dataset showing comparison of low-frequency power information of link circuit.

S.no.	Time (s)	Not optimized Power (pu)	Proposed method Power (pu)
1	0	3.1	3.1
2	0.5	3.1	3.1
3	1	3.1	3.1
4	2	1.1	1.1
5	3	2.6	1.2
6	4	2.8	2.2
7	5	3.1	3.0
8	6	2.9	2.6
9	7	2.6	2.4
10	8	2.7	2.4
11	9	2.9	2.4
12	10	3.2	3.1
13	11	3.0	2.8
14	13	3.1	3.0
15	15	3.0	2.8
16	16	3.2	3.0
17	17	3.3	3.1
18	19	3.0	2.9
19	20	3.1	3.1
20	22	3.1	3.0
21	25	2.55	3.1
22	29	3.0	3.1
23	30	3.0	3.1
24	35	3.1	3.1
25	40	3.1	3.1
26	50	3.1	3.1

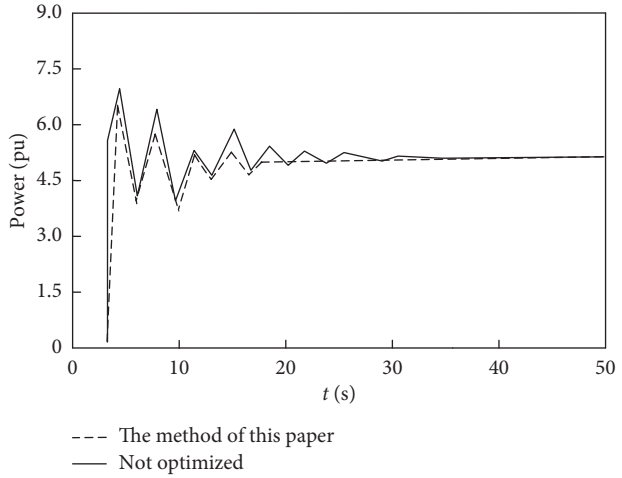


Figure 4: Comparison of low-frequency information of the line's power.

the system could be accurately analyzed, and different methods were adopted to optimize strategies according to the amount of low-frequency information and the degree of oscillation of the system, thereby obtaining a better operation effect.

5. Conclusion

In order to ensure the stability of the low-frequency electromechanical operation of the building electrical system, a TCSC-based adaptive configuration method of low-frequency electromechanical sampling information is designed in this research. The proposed method has simple calculation, effectively improves the damping of the building electrical system, and has a good effect on suppressing the low-frequency electromechanical oscillations of various degrees.

However, when measuring the output time sequence of the system, the stages of the system application were ignored, which might lead to poor accuracy of the time sequence value. This problem should be solved in further research.

Nomenclature

T:	The time-step length
k, p, and q:	Time-step coding
w:	The system input stochastic interference column vector
v:	The system output stochastic deviation column vector
u:	The system control input
A:	The state matrix
e_i:	Eigenvalue
φ_i:	Relative right eigenvector
s_i:	Continuous system model
w:	Interference input of white noise
v:	Measurement deviation of white noise
y:	Output signal
$\Lambda_1, \Lambda_2, \ldots, \Lambda_i$:	The covariance matrix
$T_{1\|i}$:	Toeplitz matrix
Q_i:	The extended observability matrix
C:	The output matrix
r:	The dimension of the reference channel
Λ_i^{ref}:	The output covariance matrix
Z_{TCSC0}:	The command impedance
ΔZ_1:	The damping impedance of power loop
ΔZ_2:	The damping impedance of voltage loop
Z_{TCSC}:	The output impedance of final stage
V:	The positive sequence voltage
V_0:	The voltage reference value
P_{L0}:	The transmission power reference value
P_L:	The power transmission monitoring value
$K_m/(1+T_mS)$:	The transfer function of the deviation comparison link
$G(s)$:	The open-loop transfer function
λ_i:	The pole of $G(s)$
R_i:	The residue relative to λ_i
λ_i:	The pole of the low-frequency oscillation information mode
R_k:	Relative residue
$KH(s)$:	Device transfer function
$\|R_{k,BI}\|$:	The digital modulus in the adaptive mode
$\|R_{k,B0}\|$:	The digital modulus in the standard mode
ρ_k:	The residue
$\varnothing, \alpha, T_2, T_3$:	The adaptive configuration parameters of the phase shifting link
V_i:	The modal corresponding to mode s_i.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Additional Points

Highlights. (1) The main aim is to enhance the reliability of the electrical systems in low frequency. (2) Adaptive electromechanical sampling information based on TCSC is designed. (3) The architecture and working principle of the controllable series compensation device are analyzed. (4) The test signal method is used to evaluate the low-frequency modes. (5). The information is required by the device among the electrical system regions of buildings. (6) The alternative input signals are selected by comparing different input signal residue ratios. (7) The experimental results show that the scheme can effectively improve the damping ratio of the system. (8) The results show that the scheme suppresses the continuous oscillation under a low-frequency fault.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Natural Science Foundation of Tianjing (no. 19JCTPJC53400).

References

[1] M. Abdollahi, J. I. Candela, J. Rocabert, M. A. Elsaharty, and P. Rodriguez, "Novel analytical method for dynamic design of renewable SSG SPC unit to mitigate low-frequency electromechanical oscillations," *IEEE Transactions on Power Electronics*, vol. 35, no. 7, pp. 7532–7544, 2020.

[2] M. Yang, Y. Wang, C. Cai, Z. Liu, H. Zhu, and S. Zhou, "Monocular vision-based low-frequency vibration calibration method with correction of the guideway bending in a long-stroke shaker," *Optics Express*, vol. 27, no. 11, p. 15968, 2019.

[3] W. Yang, P. Norrlund, J. Bladh, J. Yang, and U. Lundin, "Hydraulic damping mechanism of low frequency oscillations in power systems: quantitative analysis using a nonlinear model of hydropower plants," *Applied Energy*, vol. 212, pp. 1138–1152, 2018.

[4] W. Yaqi and L. Zhigang, "Suppression research regarding low-frequency oscillation in the vehicle-grid coupling system using model-based predictive current control," *Energies*, vol. 11, no. 7, p. 1803, 2018.

[5] H. Shen, J. Xu, and Y. Kang, "A fast adaptive resonance suppression scheme for high-damping system using R-SDFT detection," *International journal of Electronics*, vol. 106, no. 7–9, pp. 976–991, 2019.

[6] H. Shokouhandeh and M. Jazaeri, "An enhanced and auto-tuned power system stabilizer based on optimized interval type-2 fuzzy PID scheme," *International Transactions on Electrical Energy Systems*, vol. 28, no. 1, pp. 1–19, 2018.

[7] U. Agrawal, J. Follum, J. W. Pierre, and D. Duan, "Electromechanical mode estimation in the presence of periodic forced oscillations," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1579–1588, 2019.

[8] S. G. Bortnikov, V. S. Aliev, and I. A. Badmaeva, "VO2 film temperature dynamics at low-frequency current self-oscillations," *Journal of Applied Physics*, vol. 123, no. 7, pp. 75701.1–75701.6, 2018.

[9] V. J. Boucher, A. C. Gilbert, and B. Jemel, "The role of low-frequency neural oscillations in speech processing: revisiting delta entrainment," *Journal of Cognitive Neuroscience*, vol. 31, no. 8, pp. 1–11, 2019.

[10] P. Liu, S. Liu, H. Yang, X. Kang, and S. Pan, "K-band optoelectronic oscillator based on a double-Brillouin-frequency shifter," *Optical Engineering*, vol. 58, no. 10, p. 1, 2019.

[11] S. Qiang, W. Xinpeng, and W. Yixuan, "Oscillation suppression in the sense mode of a high-Q MEMS gyroscope using a simplified closed-loop control method," *Sensors*, vol. 18, no. 8, Article ID 2443, 2018.

[12] A. H. Naghshbandy and A. Faraji, "Coordinated design of PSS and unified power flow controller using the combination of CWT and Prony methods with the help of SPEA II multi-objective optimisation algorithm," *IET Generation, Transmission & Distribution*, vol. 13, no. 21, pp. 4900–4909, 2019.

[13] F. Sun, J. Zou, and H. B. Chan, "In-situ control of electrical properties of nanoelectromechanical resonators by electromigration for self-sustained oscillations," *Applied Physics Letters*, vol. 115, no. 20, Article ID 203101, 2019.

[14] P. Wang, Z. He, K. R. Thompson, and J. Sheng, "Modulation of near-inertial oscillations by low-frequency current variations on the inner scotian shelf," *Journal of Physical Oceanography*, vol. 49, no. 2, pp. 329–352, 2019.

[15] A. M. S. Fransco and E. Echer, "Wavelet analysis of low frequency plasma oscillations in the magnetosheath of Mars," *Advances in Space Research*, vol. 65, no. 9, pp. 2090–2098, 2020, https://www.sciencedirect.com/science/article/abs/pii/S027311771930657X.

[16] S. Ahmed, X. Zou, N. Jaber, M. I. Younis, and H. Fariborzi, "A low power micro-electromechanical resonator-based digital to analog converter," *Journal of Microelectromechanical Systems*, vol. 29, no. 3, pp. 320–328, 2020.

[17] Y. Dong, X. Zhang, and R. Yu, "Stability control and system protection scheme for power grid with hydropower pool and multi-DC weak sending-end," *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, vol. 42, no. 22, pp. 19–25, 2018.

[18] A. Galetzka, Z. Bontinck, and U. Romer, "A multilevel Monte Carlo method for high-dimensional uncertainty quantification of low-frequency electromagnetic devices," *IEEE Transactions on Magnetics*, vol. 55, no. 8, pp. 1–12, 2019.

[19] R. Wu, P. C. E. Roberts, C. Soutis, and C. Diver, "Downrange manoeuvre and oscillation suppression of a self-regulating centrifugally deployed flexible heat shield using a controlled reaction wheel," *Acta Astronautica*, vol. 161, no. 2, pp. 415–424, 2019.

[20] S. W. Hwang, J.-H. Bak, J. Yoon, and J. H. Park, "Oscillation reduction and frequency analysis of under-constrained cable-driven parallel robot with three cables," *Robotica*, vol. 38, no. 3, pp. 375–395, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 8883614, 7 pages
https://doi.org/10.1155/2021/8883614

Research Article

A Hardware-Efficient Elliptic Curve Cryptographic Architecture over GF (p)

Chao Cui, Yun Zhao, Yong Xiao, Weibin Lin, and Di Xu

Electric Power Research Institute of CSG, Guangzhou, Guangdong 510663, China

Correspondence should be addressed to Chao Cui; ncepucc@126.com

Received 10 September 2020; Revised 15 January 2021; Accepted 5 May 2021; Published 19 May 2021

Academic Editor: Hussein Abulkasim

This paper proposes a hardware-efficient elliptic curve cryptography (ECC) architecture over GF(p), which uses adders to achieve scalar multiplication (SM) through hardware-reuse method. In terms of algorithm, the improvement of the interleaved modular multiplication (IMM) algorithm and the binary modular inverse (BMI) algorithm needs two adders. In addition to the adder, the data register is another optimize target. The design compiler is synthesized on 0.13 μm CMOS ASIC platform. The time range of performing scalar multiplication over 160, 192, 224, and 256 field orders under 150 MHz frequency is 1.99–3.17 ms. Moreover, the gate area required for different field orders in this design is in the range of 35.65k–59.14k, with 50%–91% hardware resource less than other processors.

1. Introduction

Due to the rapid development of technology, Internet of Things- (IoT-) related devices have become popular. Most importantly, the safety must be guaranteed. In addition to the IoT devices, the safety of road networks also needs to be paid great attention [1]. Miller [2] and Koblitz [3] put forward the concept of elliptic curve cryptography (ECC), which is a kind of asymmetrical cryptosystem put forward by Miller [2] and Koblitz [3] in 1986, which has higher security than other methods like RSA encryption algorithm. Several international organizations have adopted ECC, including NIST [4], ANSI [5] and IEEE [6].

For ECC, there have been a large number of hardware architectures [7–17]. Among them, there are two methods for the realization of modular multiplication (MM), namely, the multiplier and the adder. The multiplier-based architecture includes the design based on specific prime field and the design based on Montgomery multiplication algorithm [7]. The adder-based architecture includes the design based on interleaved multiplication algorithm [9]. The processor [13] uses a design with Montgomery MM algorithm and r-bit $*$ r-bit multiplier. The processors [8, 14] use a design with n-bit $*$ n-bit multiplier. MM includes multiplication and fast reduction operation over a specific prime field. It should be noted that the multiplier-based architecture requires a lot of hardware.

In ECC, modular inversion (MI) is also a kind of cumbersome operation. Among them, binary modular inversion algorithms are usually used in hardware-efficient architectures. The MM and MI units of processor [11] are based on the adder, and the two units are independent in adder. Processors [18, 19] adopt a radix-4 booth encoding IMM algorithm. Processor [20] implements MM through a radix-2 MM algorithm and avoids MI through projective coordinates.

Traditional cryptographic algorithm software has the disadvantages of high power consumption and time delay, which can be solved by hardware implementation. This article attempts to provide security assurance with low

power consumption for IoT devices through hardware implementation. The following are the main contributions of this article.

(1) A hardware-efficient architecture based on add units is proposed to achieve as little hardware consumption as possible

(2) Through the modification of IMM algorithm and BMI algorithm with 2 full-word adders and four data registers, MM and MI can be realized

(3) Registers are optimized to minimize hardware consumption, in which four full-word register units for MM, MS, MA, and MI and eight full-word register units for SM operation

The structure of this article is divided into four parts. First, the Mathematical Background section elaborates on EC operation and SM operation. Second, the Scalar Multiplication Architecture section introduces the hardware-efficient architecture over GF(p). Third, the Implementation and Result section shows the results and then conducts comparative analysis. Fourth, the Conclusion section is a summary.

2. Mathematical Background

2.1. Elliptic Curve over GF(p). An introduction on EC over GF(p) is conducted. When the p value of nonsupersingular elliptic curve E on GF(p) is greater than 3, the following formula can be used:

$$y^2 = x^3 + ax + b, \tag{1}$$

where a, b, x, and y are elements of GF(p), and $4a^3 + 27b^2 \neq 0 \bmod p$. See [21, 22] for more information on elliptic curve cryptographic primitives.

Formulas (2) and (3) show the point adding (PA) operation and point doubling (PD) operation. With elliptic curve point $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, the computing formula for PA is $P_3(x_3, y_3) = P_1(x_1, y_1) + P_2(x_2, y_2)$, and the computing formula for PD is $P_3(x_3, y_3) = 2P_1(x_1, y_1)$.

$$\begin{cases} x_3 = \lambda^2 - x_1 + x_2 \\ y_3 = \lambda(x_1 - x_3) - y_1 \end{cases}, \tag{2}$$

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, & \text{if } P_1 \neq P_2, \\ \dfrac{3x_1^2 + a}{2y_1}, & \text{if } P_1 = P_2. \end{cases} \tag{3}$$

2.2. Elliptic Curve Scalar Multiplication. In ECC, SM is the basic operation. As for PM operation, integer k and point P on the elliptic curve are input and then performed as a sequence of PA and PD operations given in Algorithm 1. In Step 1, the point Q is initialized as a point at infinity. Step 2 is to iterate $n-1$ times, where each iteration has the PD operation. $k_i = 1$ indicates that there is a PA operation.

3. Scalar Multiplication Architecture

This part describes the bottom-up algorithm optimization on GF(p), which achieves maximum reuse by adder unit. The SM operation is implemented by using two full-word adder units. The optimization of MM and MI operations is conducive to the reduction of power consumption and the improvement of SM operation's performance.

3.1. Modular Addition/Subtraction. MA and MS operations are implemented based on Algorithm 2. In ASIC, the addition or subtraction operations can be implemented using nearly equal hardware, namely, adder units. Since MA and MS operations require a clock cycle, there is a need for 2 full-word adders. In addition, here, the adder unit is the minimum unit, and $C0_n$ and $C1_n$ are the most significant bits (MSB).

3.2. Modular Multiplication. MM is an indispensable operation in SM operation architecture. In this study, the interleaved modular multiplication algorithm is selected. The standard interleaved modulo multiplication in [16] (Algorithm 2) has certain shortcomings. Since steps 5, 6, and 7 carry out addition operations with carry propagation and steps 6 and 7 check all lengths of the operands, there is a large latency. In response to this problem, the improved algorithm in [16] (Algorithm 3) performs addition operations with carry-save adders in the loop. Moreover, the modified algorithm in [16] (Algorithm 4) reduces the area and time by lookup-table method. In [10], a new interleaved modular multiplication algorithm is proposed, which uses only two adder units. The specific steps are shown in Algorithm 5 as follows. In step 1, the variable R is initialized to zero. In step 2.1, the $R * 2$ can be realized by shifting operation. In step 2.2, the $X_i * \mathrm{Y}$ can be implemented by a multiplexer. Step 2.3 and step 2.4 require an adder unit, respectively. Therefore, if each iteration is completed within one clock cycle, then a total of two adder units are required. After the iteration of step 2, the result is limited to $[0, 2p-1]$. Therefore, it is necessary to go to step 3 to limit R to $[0, p-1]$.

3.3. Modular Inversion. In addition to the MM operation, the modular inversion (MI) operation also plays an extremely important role in the SM operation architecture. In MI operation, the same two adder units are reused to reduce hardware consumption. This paper adopts the binary modular inversion algorithm proposed in [10]. Algorithm 3 can calculate MM and MI operations in the same clock cycle. If the input $a = 1$, it is an MI operation, and $y = 1/x \bmod p$. In step 1, the variables u, v, r, s are initialized. In step 2 and step 3, the /2 operations can be realized by right shifting one bit. With a positive or negative odd r, $R/2 \bmod p = (r + p) \gg 1$, that is, it can be computed by adding r to p and then shifting right. The same is true in other situations. The above operations require one adder unit. In step 2 or step 3, the comparison between u and v in step 4 is calculated in advance, which requires two adder

```
Input: an integer k and a point P on elliptic curve
Output: kP
(1) Q = ∞;
(2) for (i = n − 1; i ≥ 0; i + + ) {
    (2.1) Q = 2Q;
    (2.2) if k_i = 1, {Q = Q + P; }
          }
(3) return Q
```

Algorithm 1: Elliptic curve scalar multiplication.

```
Input: p, A, B ∈ [0, p − 1]
Output: R = (A + B) mod p
(1) C0 = A + B
(2) C1 = C0 − p
(3) if C1n = 1 {R = C0}
(4) else {R = C1}
(5) return R
Input: p, A, B ∈ [0, p − 1]
Output: R = (A − B) mod p
(1) C0 = A − B
(2) C1=C0 + p
(3) if C0 n = 1 {R = C1}
(4) else {R = C0}
(5) return R
```

Algorithm 2: Modular addition and subtraction in GF(p).

```
Input: p, x, a ∈ [1, p − 1]
Output: y, satisfying xy = a mod p
Step 1: u = p; v = x; r = 0; s = a;
Step 2: if (u is even) {
   u = u/2;
   if (r is odd) r = (r + p)/2;
   else if (r is negative) r = (r + 2p)/2;
   else r = r/2;
   }
Step 3: if (v is even) {
   v = v/2;
   if (s is odd) s = (s + p)/2;
   else if (s is negative) s = (s + 2p)/2;
   elses = s/2;
   }
Step 4: if (u and v are odd) {
   if(u > v) r = r − s; u = u − v;
   else s = s − r; v = v − u;
   }
Step 5: if (u = 1) {
      if (r < 0) {return r = r + p; }
      else { return r.}
   }
   else if (v = 1) {
      if (s < 0) {return s = s + p; }
      else {return s.}
   }
   else {go to step2.}
```

Algorithm 3: Binary modular inversion (IBMI) algorithm.

Input: P1($x1$, $y1$), P2($x2$, $y2$),
Output: P3($x3$, $x3$) = $P1$ + P2
(1) $t2 = y1 - y2$
(2) $t1 = x1 - x2$
(3) $t1 = t2/t1$
(4) $t2 = t1 * t1$
(5) $t2 = t2 - x1$
(6) $t2 = t2 - x2$
(7) $t2 = x2 - t2$, $x1 = t2$
(8) $t2 = t1 * t2$
(9) $y1 = t2 - y2$
(10) return $x3 = x1$,$y3 = y1$
Input: P1($x1$, $y1$) = P2($x2$, $y2$)
Output: P3($x3$, $x3$) = $P1$ + P2
(1) $t2 = x1 * x1$
(2) $t1 = t2 + t2$
(3) $t1 = t2 + t1$
(4) $t2 = t1 + a$
(5) $t1 = y1 + y1$
(6) $t1 = t2/t1$
(7) $t2 = t1 * t1$
(8) $t2 = t2 - x1$
(9) $t2 = t2 - x2$
(10) $t2 = x2 - t2$, $x1 = t2$,
(11) $t2 = t1 + t2$
(12) $y1 = t2 - y2$
(13) return $x3 = x1$, $y3 = y1$

Algorithm 4: Point addition and point doubling.

Input: $p, X, Y \in [0, p-1]$
Output: $R = X * Y \bmod p$
(1) $R = 0$;
(2) for{($i = n - 1$; i ≥ 0; i + +)
(2.1) $R = R * 2$;
(2.2) $I = Xi * Y$;
(2.3) $R = R + I$;
(2.4) $R = R - R[n + 1: n] * p$;
}
(3) if $R \geq p$ {$R = R - p$; }
(4) return R

Algorithm 5: Interleaved modular multiplication algorithm.

units. In step 4, $(r - s, u - v)$ or $(s - r, v - u)$ is calculated, which requires two adder units. Step 5 requires a total of two adder units, one of which is used to determine whether r or s is less than 0, and the other is used for $r + p$ or $s + p$. Therefore, if each step is completed in one clock cycle, two adder units are required.

3.4. Point Addition and Point Doubling. Algorithm 4 provides PA and PD operations. Since modular operations (MA, MS, MM, and MI) share the same two adder units, only one modular operation is computed at a time. A total of eight registers are required, of which six are used for PA and PD operations of t_1, t_2, x_1, x_2, y_1, and y_2, and two for integer k and prime p.

3.5. Scalar Multiplier Architecture. In this part, Figure 1 shows the scalar multiplication architecture of SM on GF(p), which achieves the modular operations of MM, MS, MA, and MI as well as the point operations of SM, PA, and PD. Among them, point controller block is the main state

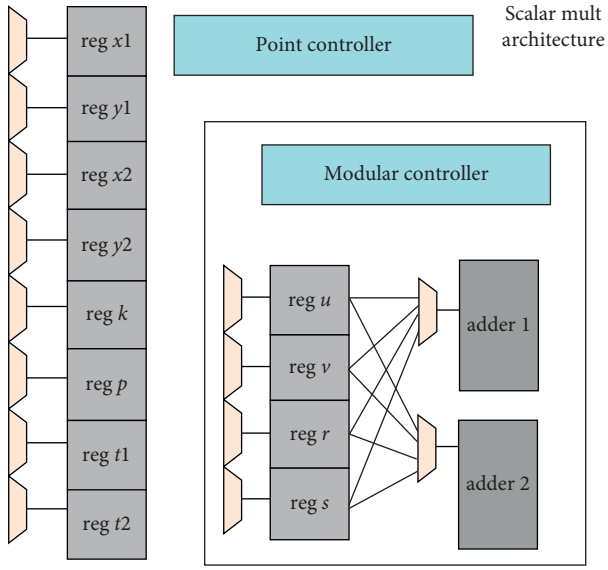


Figure 1: Scalar multiplication architecture.

machine that realizes the point operation, and modular controller block is the state machine that realizes the modular operation.

4. Implementation and Result

The ECC architecture described in this part is designed using Verilog-HDL language and adopt Design Compiler to synthesize it using SMIC 130-nm CMOS standard cell library. In addition, the experimental circuit area is evaluated by the 2-way NAND gate.

The source of the experimental simulation parameters is the FIPS 186-2 standard [8]. Figure 2 lists the main parameters for one 256 bit elliptic curve on the prime field GF(p) and the other bit elliptic curve can be found in the FIPS 186-2 standard. The coordinates of base point G on elliptic curve are Gx and Gy.

Parameter	Value
a	FFFFFFFE FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF 00000000 FFFFFFFF FFFFFFFC
b	5AC635D8 AA3A93E7 B3EBBD55 769886BC 651D06B0 CC53130F6 3BCE3C3E 27D2604B
p	FFFFFFFF 000000001 00000000 00000000 00000000 FFFFFFFF FFFFFFFF FFFFFFFF
Gx	6B17D1F2 E12C4247 F8BCE6E5 63A440F2 77037D81 2DEB33A0 F4A13945 D898C296
Gy	4FE342E2 FE1A7F9B 8EE7EB4A 7C0F9E16 2BCE3357 6B315ECE CBB64068 37BF51F5

Figure 2: Curve parameter of our experiment.

There is a need for a total of two adders and twelve data registers in the proposed architecture. According to Table 1, the required registers and adders consumed 42% of the hardware. Among them, the twelve registers are used for data storage. With the increase of field order, the adder's resource consumption percentage increases from 13.72% to 15.54%.

In Table 2, the results of the implementation and comparison of the proposed architecture are shown. By testing 100 times, the SM operation requires an average of 186, 268, 364, and 475 clock cycles on 160, 192, 224, and 256 prime fields, respectively. The proposed architecture takes 1.24, 1.78, 2.42, and 3.16 ms with the 35.65 k, 43.25 k, 49.41 k, and 59.14 k gate area for one SM operations over 160, 192, 224, and 256 prime fields, respectively.

The authors of [10, 11] use IMM algorithm and BIA to realize the inversion and multiplier units. Among them, the processor in [10] and the processor we proposed use the same method, that is, use the same unit to implement MM and MI operations. But in contrast, the proposed design has higher performance on the prime fields of 160/192/224/256 bits, which is 1.28~1.29 times faster than that of [10]. Under 160 bit prime field, the processor in [10] takes 35.43k gate area and 1.60 ms to perform an SM operation. In the area-time product (AT) parameter, the AT value of the processor we designed is relatively low, indicating that there is a better balance between hardware consumption and performance.

The processor in [11] uses two adder-based inversion units and two adder-based multiplier units, and our processor uses one combined unit. In contrast, the proposed processor has the advantage of low hardware consumption. In addition, our design saves 64.81%, 64.87%, 65.66%, and 64.69% area over the 160/192/224/256 bit prime fields than the design in [11]. Taking the 160 bit prime field as an example, the processor in [11] takes 101.3 k gate area and

Table 1: Hardware consumption of register and adder.

Field order	Total area	Area			Percent		
		Register	Adder	Register/adder	Register	Adder	Register/adder
160	35.65	9.95	4.89	14.84	27.91	13.72	41.63
192	43.25	11.83	6.19	18.02	27.35	14.31	41.66
224	49.41	13.69	7.02	20.71	27.71	14.21	41.91
256	59.14	15.87	9.19	25.06	26.83	15.54	42.37

Table 2: ECC hardware performance comparison.

Design	Technology	Field order	Area (kgate)	Frequency (MHz)	Clock cycles (k)	SM (ms)	At
This work	0.13 μm CMOS	160	35.65	150	186	1.24	44
		192	43.25	150	268	1.78	77
		224	49.41	150	364	2.42	120
		256	59.14	150	475	3.16	187
10	0.13 μm CMOS	160	35.43	150	239	1.60	57
		192	43.37	150	342	2.28	99
		224	50.38	150	468	3.12	157
		256	57.05	150	610	4.07	232
11	0.13 μm CMOS	160	101.3	150	129.3	0.87	88
		192	123.1	138	–	1.36	167
		224	143.9	130	–	1.95	281
		256	167.5	110	–	3.01	504
13	0.13 μm CMOS	160	117.5	137.7	153	1.21	142
		192	118.02	137.7	184	1.44	170
		224	120.26	137.7	297	2.34	281
		256	120.26	137.7	340	2.68	322
14	0.13 μm CMOS	256	659	163.7	3.3	0.02	13
15	0.13 μm CMOS	256	122	556	562	1.01	123

0.87 ms. Although it has higher performance, the design we propose chooses a lower AT value in order to balance hardware consumption and performance. In summary, the proposed processor has the advantages of low hardware consumption and high hardware efficiency.

The processor in [13] uses a word-based Montgomery multiplier and dynamic redundant binary converter, which can improve the performance of SM. Compared with the design in [13], our design can save 69.66%, 63.35%, 58.93%, and 50.84% area over the 160/192/224/256 bit prime fields.

The processor in [14] causes large power consumption, which is not suitable for IoT devices. More specifically, a full-size 256 bit × 256 bit multiplier requires a large hardware consumption, namely, 659 k gate. In contrast, the proposed design can save 91.03% of the area. The processor in [15] uses a systolic arithmetic unit in high frequency of 556 MHz. Based on the 256 bit prime fields, our design can save 51.52% of the area.

Compared with the abovementioned processors in [10, 11, 13, 14] and [15], our proposed processor has the least hardware consumption.

5. Conclusion

By constructing a bottom-up optimization for all operations of algorithm-level scalar multiplication on the basis of two full-word adders, a hardware-efficient elliptic curve processor over GF(p) is proposed. Through the improvement of IMM algorithm and BMI algorithm, they become suitable for two adder units. Moreover, the registers are also optimized. A total of 12 full-word register units are used to store data. Synthesized on 0.13 μm ASIC platform, the processor's hardware consumption can be controlled within the range of 35.65 k~59.14 k, which is far lower than most processors.

Data Availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Science and Technology Planning Project of Guangdong Province of China (nos. 2015B010128013, 2015B090911001, and 2015B090912001).

References

[1] J. C. Hung and Y. H. Lee, "Intelligent travel information platform based on location base services to predict user travel behavior from user-generated GPS traces," *International Journal of Computers and Applications*, vol. 39, no. 3, pp. 155–168.

[2] V. S. Miller, "Use of elliptic curves in cryptography," in *Proceedings of the Advances in Cryptology. CRYPTO 1985*, pp. 417–426, Santa Barbara, CA, USA, January 1986.

[3] N. Koblitz, "Elliptic curve cryptosystems," *Mathematics of Computation*, vol. 48, no. 177, p. 203, 1987.

[4] National Institute of Standards and Technology, *Digital Signature Standard*, FIPS Publication, Gaithersburg, MD, USA, 2000.

[5] ANSI X9.63, *Public Key Cryptography for the Financial Services Industry: Elliptic Curve Key Agreement and Key Transport Protocols*, American National Standards Institute, New Yark, NY, USA, 2000.

[6] Institute of Electrical and Electronic Engineers, NY, *P1363 Standard Specifications for Public Key Cryptography*, Institute of Electrical and Electronic Engineers, New Yark, NY, USA, 2000.

[7] P. L. Montgomery, "Modular multiplication without trial division," *Mathematics of Computation*, vol. 44, no. 170, p. 519, 1985.

[8] T. Güneysu and C. Paar, *Ultra High Performance ECC over NIST Primes on Commercial FPGAs*, Springer, Berlin, Germany, 2008.

[9] M. A. Nassar and L. A. A. El-Sayed, "Efficient interleaved modular multiplication based on sign detection," in *Proceeding of the 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, pp. 1–5, Marrakech, Morocco, November 2015.

[10] X. Hu, X. Zheng, S. Zhang, S. Cai, and X. Xiong, "A low hardware consumption elliptic curve cryptographic architecture over GF(p) in embedded application," *Electronics*, vol. 7, no. 7, p. 104, 2018.

[11] S. Ghosh, M. Alam, D. R. Chowdhury, and I. S. Gupta, "Parallel crypto-devices for GF(p) elliptic curve multiplication resistant against side channel attacks," *Computers & Electrical Engineering*, vol. 35, no. 2, pp. 329–338, 2009.

[12] K. Javeed and X. Wang, "FPGA based high speed SPA resistant elliptic curve scalar multiplier architecture," *International Journal of Reconfigurable Computing*, vol. 2016, Article ID 6371403, 10 pages, 2016.

[13] A. Satoh and K. Takano, "A scalable dual-field elliptic curve cryptographic processor," *IEEE Transactions on Computers*, vol. 52, no. 4, pp. 449–460, 2003.

[14] Z. Zhao and G. Bai, "Ultra high-speed SM2 ASIC implementation," in *Proceedings of the 2014 IEEE 13th International Conference*, pp. 182–188, Singapore, June 2014.

[15] G. Chen, G. Bai, and H. Chen, "A High-Performance elliptic curve cryptographic processor for general curves over GF(p) based on a systolic arithmetic unit arithmetic unit," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 5, pp. 412–416, 2007.

[16] D. N. Amanor, C. Paar, J. Pelzl, and V. Bunimov, "Efficient hardware architectures for modular multiplication," in *Proceedings of the International Conference on Field Programmable Logic and Applications, 2005*, Tampere, Finland, February, 2005.

[17] S. Ghosh, M. Alam, I. S. Gupta, and D. R. Chowdhury, "A Robust GF(p) parallel arithmetic unit for public key cryptography," in *Proceedings of the EUROMICRO DSD 2007*, pp. 109–115, Lubeak, Germany, August 2007.

[18] K. Javeed and X. Wang, "Low latency flexible FPGA implementation of point multiplication on elliptic curves over GF(p)," *International Journal of Circuit Theory and Applications*, vol. 45, no. 2, pp. 214–228, 2017.

[19] K. Javeed, X. Wang, and M. Scott, "High performance hardware support for elliptic curve cryptography over general prime field," *Microprocess & Microsystem*, vol. 51, pp. 331–342, 2016.

[20] M. M. Islam, M. S. Hossain, M. K. Hasan, M. Shahjalal, and Y. M. Jang, "FPGA implementation of high-speed area-efficient processor for elliptic curve point multiplication over prime field," *IEEE Access*, vol. 7, Article ID 178811, 2019.

[21] Standard Specifications for Public-key Cryptography, "IEEE standard P1363," 2000, http://grouper.ieee.org/groups/1363.

[22] D. Hankerson, A. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*, Springer, New York, NY, USA, 2004.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 5592132, 10 pages
https://doi.org/10.1155/2021/5592132

Research Article

Binary Bitwise Artificial Bee Colony as Feature Selection Optimization Approach within Taguchi's T-Method

Nolia Harudin,[1] **Faizir Ramlie**,[2] **Wan Zuki Azman Wan Muhamad**,[3] **M. N. Muhtazaruddin**,[2] **Khairur Rijal Jamaludin**,[2] **Mohd Yazid Abu**,[4] **and Zulkifli Marlah Marlan**[2]

[1]*Department of Mechanical Engineering, Universiti Tenaga Nasional, Kajang 43000, Selangor, Malaysia*
[2]*Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, Kuala Lumpur 54100, Malaysia*
[3]*Institute of Engineering Mathematics, Universiti Malaysia Perlis, Kampus Pauh Putra, Arau 02600, Perlis, Malaysia*
[4]*Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, Pekan 26600, Malaysia*

Correspondence should be addressed to Wan Zuki Azman Wan Muhamad; wanzuki@unimap.edu.my

Received 5 February 2021; Accepted 22 April 2021; Published 7 May 2021

Academic Editor: Gengxin Sun

Taguchi's T-Method is one of the Mahalanobis Taguchi System- (MTS-) ruled prediction techniques that has been established specifically but not limited to small, multivariate sample data. The prediction model's complexity aspect can be further enhanced by removing features that do not provide valuable information on the overall prediction. In order to accomplish this, a matrix called orthogonal array (OA) is used within the existing Taguchi's T-Method. However, OA's fixed-scheme matrix and its drawback in coping with the high-dimensionality factor led to a suboptimal solution. On the contrary, the usage of SNR (dB) as its objective function was a reliable measure. The application of Binary Bitwise Artificial Bee Colony (BitABC) has been adopted as the novel search engine that helps cater to OA's limitation within Taguchi's T-Method. The generalization aspect using bootstrap was a fundamental addition incorporated in this research to control the effect of overfitting in the analysis. The adoption of BitABC has been tested on eight (8) case studies, including large and small sample datasets. The result shows improved predictive accuracy ranging between 13.99% and 32.86% depending on cases. This study proved that incorporating BitABC techniques into Taguchi's T-Method methodology effectively improved its prediction accuracy.

1. Introduction

Taguchi's T-Method, which was explicitly developed for predictive analysis, is one of the Mahalanobis Taguchi System's (MTS) variants that has been increasingly used by researchers and industrial practitioners in Japan and other countries. Taguchi's T-Method was proposed for multivariate estimation to predict the integrated estimated output value. In the 1980s, Dr. Genichi Taguchi developed the Mahalanobis Taguchi System (MTS) as a pattern recognition technique that blends Mahalanobis Distance (MD) theory and Taguchi Robust Engineering concept to systematically and effectively classify and predict data in a multidimensional environment [1–6]. MTS establishes a multivariate measurement scale that recognizes a normal or healthy observation from an abnormal or an unhealthy observation and integrates it with the concept of signal-to-noise ratio (SNR) and orthogonal array (OA). Beginning with the introduction of the MT-Method as a classification technique that has so far gained much attention among scholars [7–14], Taguchi's T-Method has been established since then, which has utilized the same integration principles. The unit-space concept, the duplicate signal-to-noise ratio (SNR) adaptation as a weighting factor, zero-proportional theory, and OA as the feature selection optimization are the main elements that have been adopted in reinforcing Taguchi's T-Method robustness.

One of Taguchi's T-Method significant advantages is its ability to predict even with limited sample data. In multiple regression analyses, a limitation exists in which the sample size has to be higher than the number of variables. On the contrary, the said limitation does not apply to Taguchi's T-Method. Additionally, Taguchi's T-Method has no direct influence from multicollinearity since individual regression has been considered [2, 15, 16]. Based on the number of papers published in the literature, Taguchi's T-Method studies' progress is moving towards optimizing parameters and optimizing feature selection rather than just application purposes since the year 2012 [17–20]. The increasing pattern has indirectly triggered that there are indeed a variety of enhanced approaches towards parameter and feature selection optimization available out there that can be further explored and incorporated into Taguchi's T-Method as a hybridization or integration element.

1.1. Taguchi's T-Method for the Feature Selection Optimization Problem. In MTS, the orthogonal array (OA) is a feature selection search mechanism that has been established between a series of MTS, including Taguchi's T-Method, which share standard procedures but vary in their objective function determination. The OA element within MTS has been debated and is believed to be insufficient as it offers a suboptimal solution [21, 22]. Most OA's concerns are based on its restriction in having appropriate combinations of features to be assessed and evaluated in the search for optimality, as it relies on a fixed scheme [20, 23]. The authors of [24] argued that the fixed combination in OA is not optimal since the results may vary significantly if the column-to-column information is rearranged [6]. In [25], the authors agreed with the authors of [24] after proving 1000 random variables to the column assignment. Issues in OA have been highlighted as well by [26, 27], especially the fact that the OA design has a limitation in handling the higher-order interaction between variables, which might lead to an inconsistency in the identification of the significant variables [24, 25, 27–29]. Therefore, developing a hybrid methodology for better accuracy is a preferred solution to these concerns that drove this research's primary motivation.

Until recently, the OA element in the MTS classification approaches has been continuously improved by numerous machine learning algorithms. However, enhancing the OA element within Taguchi's T-Method as a prediction tool is still at an initial stage. In [30], the authors applied a stepwise forward and backward selection procedure for this purpose which showed an increase in accuracy in many cases conducted [30]. The author of [31] suggested a Binary Artificial Bee Colony (BABC) algorithm, and the findings revealed that T-Method + BABC worked better than T-Method + OA in a particular case study conducted [31]. The most recent reported study by [32] has specifically addressed OA's downside and suggested Binary Particle Swarm Optimization (BPSO), which indicates an increase in accuracy for specific case studies [32]. The published literature on OA improvement in Taguchi's T-Method is found not utilizing the generalization aspect thoroughly and focused on a somewhat limited case study. The previous research by [31, 32] was further expanded in this study by proposing the other variant of binary ABC called Binary Bitwise ABC algorithms with proper generalization aspect been amended into it, which is the application of bootstrap cross-validation.

2. Methodology

2.1. Taguchi's T-Method. Regression analysis aims to construct a mathematical model that describes and explains the relationship between variables for prediction or a study of causal relationships [33]. Taguchi's T-Method, which is driven by similar purposes, was built to forecast the unknown value of the output variable concerning the established value of the input variables by statistically evaluating the relevant correlation and functional relationship between those variables through a specific developed linear regression model to compute the integrated estimate output value. The integrated estimate output model in Taguchi's T-Method consists of some additional elements that differentiated it from standard linear regression: (1) zero-point proportional term, (2) inverse regression model, (3) unit-space concept, and (4) weightage SNR. All these elements have been embedded into the existing Taguchi's T-method model described by [34] to generate the specified integrated estimated model, as shown in equation (1). Taguchi's T-Method as well utilizes the ordinary least squares approach to calculate the proportional coefficient, β which is a common approach in linear regression. Equations (2)–(7) govern the inclusion of dynamic SNR as a weightage factor for each feature within the model [35]:

$$\text{Integrated estimate value } \widehat{M}_{ii} = \frac{\eta_1 \check{M}_{ii,1} + \eta_2 \check{M}_{ii,2} + \ldots + \eta_D \widehat{M}_{ii,D}}{\eta_1 + \eta_2 + \ldots + \eta_D}, \text{ where } ii \text{ (number of sample)} = 1, 2, \ldots, l \text{ and } \widehat{\mathrm{M}} = \frac{Z_{ii,j}}{\beta_j}, \text{ while } Z \text{ is the normalized input variables,} \tag{1}$$

$$\text{Effective divider, } r = M_1^2 + M_2^2 + \ldots + M_l^2, \tag{2}$$

$$\text{Total variation, } S_{Tj} = Z_{1j}^2 + Z_{2j}^2 + \ldots + Z_{lj}^2 \text{ for } j \text{ (number of features)} = 1, 2, \ldots, D, \tag{3}$$

$$\text{variation of proportional term, } S_{\beta j} = \frac{\left(M_1 Z_{1j} + M_2 Z_{2j} + \ldots + M_l Z_{lj}\right)^2}{r}, \tag{4}$$

$$\text{Error variation, } S_{ej} = S_{Tj} - S_{\beta j}, \tag{5}$$

$$\text{Error variance, } V_{ej} = \frac{S_{ej}}{l-1}, \tag{6}$$

$$\text{SNR, } \eta = \begin{cases} \dfrac{1/r\left(\text{Variation of proportional term, } S_{\beta j} - \text{Error variance, } V_{ej}\right)}{\text{Error variance, } V_{ej}}, & \left(\text{when } S_{\beta j} > V_{ej}\right), \\ 0, & \left(\text{when } S_{\beta j} \le V_{ej}\right). \end{cases} \tag{7}$$

It is seen that the higher SNR of an item will contribute to a greater degree of contribution to overall model estimation. The integrated estimate SNR (dB) is computed based on the result obtained using equation (1). The integrated estimate SNR, η (dB), is a performance measure to evaluate the input variable's relative importance towards the output variable. To further increase the model accuracy, optimization concerning the selection of features is considered a value-added approach within Taguchi's T-Method. Equations (8)–(13) are used for calculating the SNR (dB) for feature selection optimization, which as shown below. The evaluation of the relative importance of features is conducted using the two-level orthogonal array (OA). OA with a predetermined combination of "use" and "not use" of features allows for comparison of integrated estimate SNR (dB) under the setting. Table 1 shows the example of L_{12} orthogonal array with Level 1 in the array indicates that the variable will be used, while Level 2 indicates that the variable will not be used during the simulation study. Evaluation of relative importance of features is performed by computing the new integrated estimate SNR (dB) when the features are not used in computation and observed the increment or deterioration of the value. A higher integrated estimate SNR (dB) value is preferred, and a combination of input variables that yields optimal integrated estimate SNR (dB) is selected as an optimal combination:

$$\text{Linear equation, } L = M_1 \widehat{M}_1 + M_2 \widehat{M}_2 + \ldots + M_l \widehat{M}_l, \tag{8}$$

$$\text{Variation of proportional term, } S_\beta = \frac{L^2}{r}, \tag{9}$$

$$\text{Total variation, } S_T = \widehat{M}_1^2 + \widehat{M}_2^2 + \ldots + \widehat{M}_l^2, \tag{10}$$

$$\text{Error variation, } S_e = S_T - S_\beta, \tag{11}$$

$$\text{Error variance, } V_e = \frac{S_e}{l-1}, \tag{12}$$

$$\text{Integrated estimate } SN \text{ ratio, } \eta = 10 \log\left(\frac{1/r\left(S_\beta - V_e\right)}{V_e}\right), (\text{d}B), \tag{13}$$

2.2. Binary Bitwise Artificial Bee Colony (BitABC) into Taguchi's T-Method for Feature Selection Optimization. This research's binary approach is similar to the orthogonal array (OA) concept in existing Taguchi's T-Method. The Binary ABC was explicitly developed for the feature selection optimization process by changing the information of each identified food source update to the discrete-binary data type to be "1" or "0." The primary food source *(Xi)* is randomly initialized by following the identified bee's population size (NP/2 = *N*) and the total number of features (*D*) using discrete-binary data (1 or 0). The primary objective function, which is to maximize the SNR (dB) value, is then computed. The best SNR (dB) are selected as Global_max and its binary combination as Global_para. The employed bees will continue searching for a better food source, which will make a little change based on their nearby information memory and create a new source. The objective function, SNR (dB) value, is computed then and been compared to primary sources. The higher SNR (dB) value will be memorized, while the lower will be forgotten. If the previous SNR (dB) value is

Table 1: L_{12} orthogonal array combination.

No.	Control factors											Integrated estimate SNR (dB)
	A	B	C	D	E	F	G	H	I	J	K	
1	1	1	1	1	1	1	1	1	1	1	1	SNR1
2	1	1	1	1	1	2	2	2	2	2	2	SNR2
3	1	1	2	2	2	1	1	1	2	2	2	SNR3
4	1	2	1	2	2	1	2	2	1	1	2	SNR4
5	1	2	2	1	2	2	1	2	1	2	1	SNR5
6	1	2	2	2	1	2	2	1	2	1	1	SNR6
7	2	1	2	2	1	1	2	2	1	2	1	SNR7
8	2	1	2	1	2	2	2	1	1	1	2	SNR8
9	2	1	1	2	2	2	1	2	2	1	1	SNR9
10	2	2	2	1	1	1	1	2	2	1	2	SNR10
11	2	2	1	2	1	2	1	1	1	2	2	SNR11
12	2	2	1	1	2	1	2	1	2	2	1	SNR12

(Note: 1 denotes "Item will be used," 2 denotes "Item will not be used").

higher than the existing candidates, the value will remain. This decision process is called greedy selection. The employed bees will then share the information on the new position to onlooker bees once they return to their hive in the dance area. The onlooker bees will then evaluate the new position and choose to emphasize the food source's information, relying on the probability rate calculated. The onlooker bees will modify the position if the criteria are fulfilled, and SNR (dB) amount will be recalculated and updated following greedy selection criteria. The employed bee that cannot improve their position up to the defined limit will be abandoned and become a scout bee. A scout bee will randomly search for a new food source near the surrounding area of its hive. The cycle is repeated until it reaches the maximum number of cycles. The Global_max and Global_para at the maximum cycle are updated accordingly.

The method used by the bees (employed and onlooker) to search for the new food source which having more nectar amount within its neighborhood are following the approach introduced by Jia, Duan, and Khan [36] called binary Bitwise ABC (BitABC). Bitwise operators often transform an image into a binary number and represent a series of 0s and 1s on the computer [36]. However, only the logic operator results are adopted in the study conducted by Jia, Duan, and Khan [36], as it has similar characteristics with the binary space (0 and 1). The bitwise operator (∧, &, and |) to describe the trajectory of the food source within this study is illustrated in Table 2 and equations (14) and (15).

$$v_{ij} = X_{ij} \wedge \varnothing_{ij} \& \left(X_{ij} | X_{kj}\right), \quad i = \text{number of bees, while } j = \text{number of features}, \tag{14}$$

$$\text{where, } \varnothing_{ij} = \begin{cases} 1 & \text{rand}\,(0,1) < (r = 0.5), \\ 0 & \text{rand}\,(0,1) \geq (r = 0.5). \end{cases} \tag{15}$$

2.3. Data Preparation and Selection. The optimum features are selected based on the total number of use items ("1") produced by each feature across the run's total number. The combination of use item ("1") at each run represents the combination features contributed to the most optimum SNR (dB) value across the maximum cycle iteration. In demonstrating the proposed algorithm's stability and consistency, 70% of the training dataset from 20 different independent runs were set, and features that appear to be selected more than 10 times (more than 50%) are selected as the optimum features. The optimum features will be used to validate the remaining 30% validation dataset. The 70% training dataset follows the bootstrap cross-validation analysis during the training phase, which segregates the training and test set into 63.2% and 36.8%, respectively, with 1000 bootstrap cycles. The risk of overfitting is being considered and monitored accordingly within this study.

For better comparative purposes, despite the current Taguchi's T-method, the outcome of Bitwise ABC's optimum features has also been compared to another metaheuristic algorithm variant called Probability Binary Particle Swarm Optimization (PBPSO) [32] as well as the existing Taguchi's T-Method with full features and Taguchi's T-Method with optimal features provided by OA analysis [35]. Several simulations were performed on eight real-world datasets on prediction and regression with multivariate cases in assessing the suggested algorithm. Six out of eight datasets were obtained from the University of California at Irvine (UCI) Machine Learning Benchmark Repository [37]. The other two datasets were taken from the actual case study.

Both the BitABC and PBPSO are being set by the parameter configuration listed in Table 3. The optimization of all the algorithms within this study was constructed using Matlab R2018a application software. The programming algorithm compiled on 64 bits Sony VAIO VPCCA notebook with Intel i5 (2.3 GHz) 4 Gigabytes RAM capability and 212 GB data storage. The pseudocode of the proposed BitABC algorithm into Taguchi's T-Method is shown in Figure 1.

2.4. Performance Measure. Prediction is an iteration method involving model creation before performance evaluation, then proceeds to repeat the cycle until a satisfying solution is encountered. Throughout this study, two performance criteria are used to evaluate the developed algorithm's performances: the prediction accuracy and convergence rate of training, testing, and validation dataset.

In machine learning, especially on the regression analysis, the standard prediction error performance measures are computed using the mean absolute error (MAE), root

Table 2: Bitwise logical index sample operation.

Input		Bitwise operator and description		
		And a & b	Or a | b	Exclusive or a^b
A	B		Results	
1	1	1	1	0
0	1	0	1	1
1	0	0	1	1
0	0	0	0	0

Table 3: Parameter setting for all BitABC and PBPSO algorithm.

Parameter	BABC	BPSO
Population size, NP	50	30
Acceleration coefficient, C1	NA	2
Acceleration coefficient, C2	NA	2
Inertia weight, $w_{\min}$	NA	0.4
Inertia weight, $w_{\max}$	NA	0.9
Maximum number of cycle	500	100
Limit	100	NA
$V_{\max}$	NA	5
Number of run	20	
Training dataset	70%	
Validation dataset	30%	

Note. NA: not applicable.

mean squared error (RMSE), mean absolute percentage error (MAPE), and several others. In practice, the regression prediction model accuracy must be estimated over the training and validation sets and are independent of one another. In this study, after the optimum features have been identified, the integrated estimate value, $\widehat{M}_{\text{predicted}}$ will be calculated as indicated by Equation (1). MAE formula was applied for the prediction model accuracy as shown in Equation (16). The MAPE measure has also been applied in this study to provide the final increment percentages of the optimal approach toward existing Taguchi's T-Method that uses full features, as shown by equation (17):

$$\text{MAE} = \frac{1}{\text{no.of sample} \sum_1^n \left| M_{\text{actual}} - \widehat{M}_{\text{predicted}} \right|}, \tag{16}$$

$$\text{MAPE} = 100 \left[\frac{1}{\text{no.of sample} \sum_1^n \left| M_{\text{actual}} - \widehat{M}_{\text{predicted}} \right| / M_{\text{actual}}} \right]. \tag{17}$$

3. Results and Discussion

The feature selection analysis findings are addressed according to the respective case studies presented in this research using the defined integrated estimate model shown by equation (1) previously. Despite focusing on the MAE results and its SD value, the discussion is also guided with several other performance measures such as the convergence plot of the SNR (dB) value as the objective function and also MAE for the training and testing phase. Table 4 and Figure 2 illustrate the example of the performance analysis for the heating load case study. Researchers often use this dataset to interact with several other techniques that rely on regression analysis [38, 39]. Similar procedures were applied to the remaining seven datasets applied within this study. The explanation of the heating load case study will provide a general idea of how the other case studies are analyzed in terms of their MAE trend for both training and testing, as well as the SNR (dB) convergence plot. The validation phase is summarizing the overall case studies considered within this research.

Table 4 reveals that F2, F3, F6, F7, and F8 are the dominant features for both *T* Method-BitABC and *T* Method-PBPSO. The T-Method with OA shows conflicting results, with F1 identified as one of the dominant features instead of F3 and F8 as other methods.

In providing a more explicit description of how each outcome reflects the overall prediction analysis, the effects of the SNR (dB) and MAE for the training and testing are illustrated by the convergence plot shown in Figures 1(a) and 1(b). The result reveals that the *T* Method-BitABC is the most optimum approach with the highest SNR (dB) value compared to the *T* Method-PBPSO, *T* Method, and *T* Method-OA. The trend aligned with MAE's trend for the training and testing phase, with *T* Method-BitABC performing better prediction accuracy with lower MAE value than *T* Method-PBPSO.

As seen in Table 5, the validation phase results indicate the result of the trained model performance towards the validation dataset with the case studies having more than 30 sample data (large dataset), while Table 6 summarized for the case study having less than 30 sample data (small dataset). Table 5 indicates that the result of *T* Method-BitABC and *T* Method-PBPSO reflect the same MAE performance. This is possible due to similar optimal features' selection results gained from the training and testing phase. The improvement percentages range from 13.99% to 32.86% across three different case studies (Abalone, Heating, and Cooling). Body fat and Concrete Compressive Strength cases show that Taguchi's T-Method maintains the best compared to others, while *T* Method-OA is the best for the Auto MPG case study, which contributes to 45.71% improvement compared to Taguchi's T-Method. The trend for the small sample case studies is a little bit of contrast. The result for both *T* Method-BitABC and *T* Method-PBPSO seems to differ from each other. *T* Method-BitABC provides better performance for the JD dataset with 9.07% improvement compared to Taguchi's T-Method. *T* Method-OA provides the best result for the Chiller dataset with 9.54% improvement compared to Taguchi's T-Method.

The analysis results shared explicitly represent how well the *T* Method-BitABC approach is well reflected in several case studies. A few findings could be further investigated, which implicitly represent the analysis results identified. The findings shall be summarized as follows:

The adoption of BitABC into existing Taguchi's T-Method replacing the OA is found not suitable for the body fat case study. Body fat is a case study with a normal distribution trend and has a stable output

Enumerate unit-space, η , β, r, l as fix value in evaluating the objective function (snr_est) and...
...evaluating the fitness values.
Input: *dataTraining (Z (ii,j) , M(ii,j)), N, D, limit,lb,ub,maxcycle and BS*
Output: a set of optimal solutions obtained by the algorithm

1 **For** bootstrap_runcycle = 1 : BS
2 Randomly **initialize *N x D*** food sources, $\boldsymbol{X_i} = \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & 1 & 0 & \dots & 1 \end{bmatrix}$ following uniform distribution [0,1]
3 Evaluate the functions of the initialized food source, and compute their fitness values
4 $snr\_esti\ (maximize) = \boldsymbol{f(X_i)} = 10log_{10}\left[\frac{1}{r}\frac{S_\beta - V_e}{V_e}\right]$ *% using objective function pseudocode*
5 $Fit\_F_i = \begin{cases} \frac{1}{1+f(X_i)} \\ 1 + abs\,\boldsymbol{f(X_i)} \end{cases}$
6 Set the initial Globe_max and Globe_para value *% Max(snr_est) & the combination Xi at max index*
7 ***While*** *Stopping Criteria (**maxcycle**) not satisfied* ***do***

%% Employed Bee phase
8 **for** *i = 1 : N*
r is set to *0.5*, while *k* is the neighbouring location in the range *1 to N…*
9 *... update the food source trajectory using bitwise operator sequentially...*
...following uniform distribution [0,1] .

jrand=fix(rand(D))+1;* and *k=fix(rand*N)+1;*
if jrand >= D+1 *while (k==i)*
10 *jrand=D;* *k=fix(rand*(N))+1;*
elseif jrand <=0 *end*
jrand=1;
end
11 $v_{ijrand} = x_{ijrand} \wedge \varnothing_{ijrand} \& (x_{ijrand} | x_{kjrand})$ where, $\varnothing_{ijrand} = \begin{cases} 1 & rand(0,1) < r \\ 0 & rand(0,1) \geq r \end{cases}$
12 **end**
13 Calculate the SNR function and fitness as in step 4 and 5
14 Greedy selection is applied between current and candidate solution based on maximum…
...function and minimum fitness value as in step 4 & 5
15 Keep the best solution between current and candidate solution, update *Xi* position
16 Calculate the probability , $NomFit_i = Fit\_F_i / sum\ (Fit\_F_i)$

%% Onlooker Bee Phase
17 A random value between 0 to 1 is generated for an onlooker bee to compare with...
… the calculated probability, *Nomfit* value of a food source.
18 **if** *rand < Nomfit (i)*
19 This food source (*i*) is selected by onlooker bee and step 8 to 12 is followed…
… for new food source
20 **end**
21 Greedy selection between current and candidate *snr_est* value based on maximum ...
… function and minimum fitness value as in step 4 & 5.
22 Keep the best solution between current and candidate solution, update ***Xi*** position
23 Update the Globe_max value *(the maximum SNR(dB))* and ...
...Global_para *(the best among Xi position with maximum SNR (dB))*

%% Scout Bee Phase
24 If the counter value of a food source is the maximum among those of food sources and
… exceeds limit,
25 *[Max_bus ind] = max(trial);*
26 **if** *trial(ind) > limit*
27 *trial(ind) = 0;*
28 *X(ind,:) = randi([0,1],1,D);* *% a new food source for (ind) is created by a scout*
Evaluate the functions of the new food source, and compute the fitness values.
29 **end** *% bee using binary random values 0 or 1.*
cycle = cycle + 1;
31 **End While**
32 **End**
33 $optimum\_para = \sum_{1}^{BS} Global\_para \geq 50\%$ BS

Note: Mean Absolute Error (MAE) is calculated at each cycle to see the accuracy of Global_para in improving the model accuracy while optimum_para is tested using validation data.

Figure 1: The pseudocode of the proposed BitABC algorithm into Taguchi's T-Method.

Table 4: Times of the appearance of each feature over 20 runs for heating load.

	Heating load							
Methods	F1	F2	F3	F4	F5	F6	F7	F8
T Method-PBPSO	1	19	20	0	0	11	20	10
T Method-BitABC	0	20	20	0	0	12	20	15
T Method-OA	√	√	x	x	x	√	√	x

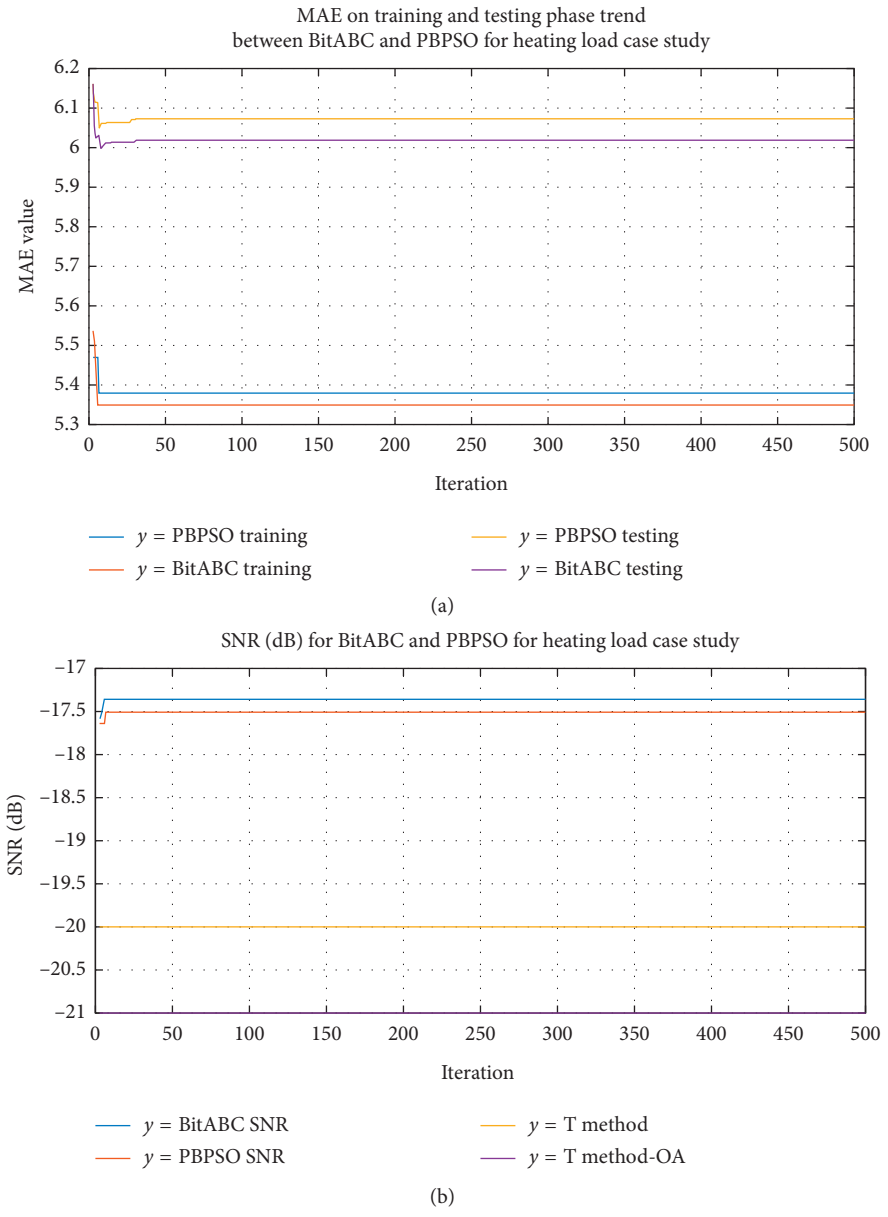


Figure 2: (a) MAE for the training and testing phase between Bitwise ABC and PBPSO for the heating load case study; (b) SNR (dB) trend between Bitwise ABC and PBPSO for the heating load case study.

performance than other cases [40]. The adoption of feature selection optimization does not provide a better trend on this type of data since the combination features are already appropriate for the model.

The Concrete Compressive Strength dataset shows how the quality of the data within each analysis affects the analysis result. By considering randomness and variation effect within datasets, it is possible to have slightly different trend results. From the result in Table 5, the slightly different trend between *T* Method, *T* Method-BitABC, and *T* Method-PBPSO shows that the proposed algorithm should provide a better deal since just

Table 5: The prediction accuracy (MAE) compilation across all proposed enhanced methods and existing Taguchi's T-Method on validation datasets for large sample data.

Dataset	# sample training	# sample testing	# features	Measure	T-Method [28]	*T* Method-BitABC and *T* Method-PBPSO [26]	T-Method + OA [28]	% MAE improvement of best result vs. T-Method
Body fat	176	76	14	MAE	0.3868	1.5663	1.651	NA
				SD	0.3250	1.0743	1.056	
				# optimum features	14	10	9	
Abalone	2924	1253	8	MAE	4.2377	3.6447	3.757	13.99%
				SD	3.3917	2.5686	2.791	
				# optimum features	8	3	4	
Heating	538	230	8	MAE	8.6033	5.7762	6.317	32.86%
				SD	8.5813	3.5191	3.831	
				# optimum features	8	5	4	
Cooling	538	230	8	MAE	8.1106	5.9515	6.312	26.62%
				SD	4.5612	4.0778	4.217	
				# optimum features	8	5	2	
Concrete Compressive Strength	721	309	7	MAE	11.4115	11.8847	12.382	NA
				SD	11.3301	12.3216	11.677	
				# optimum features	7	6	6	
Auto MPG	274	118	7	MAE	6.0035	3.5716	3.259	45.71%
				SD	3.0157	2.6609	2.717	
				# optimum features	7	2	5	

Table 6: The prediction accuracy (MAE) compilation across all proposed enhanced methods and existing Taguchi's T-Method on validation datasets for the small sample.

Dataset	# sample training	# sample testing	# Features	Measure	T-Method [28]	*T* Method-PBPSO [26]	*T* Method-BitABC	T-Method + OA [28]	% MAE improvement of best result vs. T-Method
Chiller	10	4	44	MAE	4.7925	5.3121	5.6858	4.3354	9.54%
				SD	6.2396	3.6373	4.3237	4.0135	
				# optimum features	44	12	14	16	
JD power	14	6	44	MAE	0.7866	0.9127	0.7153	0.7299	9.07%
				SD	0.7235	0.4276	0.5361	0.7431	
				# optimum features	44	24	27	23	

relying on 6 features instead of 7 total number of features. A similar situation occurs to the Auto MPG and Chiller case studies with *T* Method-BitABC, and *T* Method-PBPSO requiring fewer features compared to the *T* Method-OA with minimum MAE differences.

In this study, Taguchi's T-Method proved capable of computing a prediction analysis involving sample data much lower than the number of features than multiple linear regression that cannot compute the analysis within a similar state. This served as one of the main advantages of Taguchi's T-Method.

Adopting the BitABC replacing the OA within Taguchi's T-Method for small sample data with many features seems feasible, even though risk towards model accuracy still exists, requiring further monitoring. A considerable number of features are able to be reduced by implementing this approach. However, overfitting might be one of the risks to deal with for this small sample datasets' cases.

The adoption of BitABC seems not to differ from PBPSO for the large sample data within this study but varies for the small sample dataset. The better

exploration and exploitation search mechanism within the ABC algorithm might be the main reason for this trend since small sample data are susceptible to variation. The bootstrap, adopted as the cross-validation element, helps in reducing the risk of overfitting across training, testing, and validation dataset.

4. Conclusion

The adoption of BitABC into Taguchi's T-Method replacing the OA is shown feasible in this study. The result analysis shows that 4 out of 8 case studies reflect that BitABC adoption provides better performance than existing Taguchi's T-Method. The other case studies vary with minimal MAE differences and provide fewer significant features to be considered. Even though the trend result for both BitABC and PBPSO is similar for the large dataset, the small data samples reflected that BitABC provides much better prediction results. It was apparent that the merging of the BitABC into the current Taguchi's T-Method optimization technique to increase the SNR (dB) and predict the accuracy of the predicted integrated model was indeed practical. Further development studies should also focus on improving parameter estimates' robustness to ensure an established integrated estimated output model is reliable, especially for small sample data analysis.

Data Availability

Data are available within the repository of the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported under the Collaborative Research Grant (CRG) scheme between Universiti Teknologi Malaysia (Q. K130000.2456.08G27) and Universiti Tenaga Nasional (20200106CRGJ). This work also was funded by the Ministry of Higher Education, Malaysia under Fundamental Research Grant Scheme (FRGS/1/2019/TK08/UTM/02/4).

References

[1] G. Taguchi, S. Chowdhury, and Y. Wu, *The Mahalanobis-Taguchi System*, McGraw-Hill, New York, NY, USA, 1st edition, 2001.

[2] S. Teshima, Y. Hasegawa, and K. Tatebayashi, "Pattern recognition and the MT system," in *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, pp. 1–13, Momentum Press, New York, NY, USA, 1st edition, 2012.

[3] F. Ramlie, W. Z. A. Wan Muhamad, K. R. Jamaludin, E. Cudney, R. Dollah, and R. Dollah, "A significant feature selection in the mahalanobis taguchi system using modified-bees algorithm," *International Journal of Engineering Research and Technology*, vol. 13, no. 1, pp. 117–136, 2020.

[4] W. Z. A. W. Muhamad, K. R. Jamaludin, F. Ramlie, N. Harudin, and N. N. Jaafar, "Criteria selection for an MBA programme based on the mahalanobis Taguchi system and the Kanri Distance Calculator," in *Proceedings of the 2017 IEEE 15th Student Conference on Research and Development (SCOReD)*, pp. 220–223, Kuala Lumpur Malaysia, December 2017.

[5] S. K. M. Saad, M. H. M. Razali, M. Y. Abu et al., "Optimizing the MFlex monitoring system using Mahalanobis-Taguchi system," *IOP Conference Series: Materials Science and Engineering*, vol. 1092, no. 1, Article ID 012009, 2021.

[6] N. N. N. M. Kamil, S. N. A. M. Zaini, and M. Y. Abu, "Feasibility study on the implementation of Mahalanobis-Taguchi system and time driven activity-based costing in electronic industry," *International Journal of Industrial Management*, vol. 10, pp. 160–172, 2021.

[7] H. Sakeran, N. A. Abu Osman, and M. S. Abdul Majid, "Gait classification using Mahalanobis-Taguchi system for health monitoring systems following anterior cruciate ligament reconstruction," *Applied Sciences*, vol. 9, no. 16, 16 pages, Article ID 3306, 2019.

[8] T. Asakura, W. Yashima, K. Suzuki, and M. Shimotou, "Anomaly detection in a logistic operating system using the mahalanobis-taguchi method," *Applied Sciences*, vol. 10, no. 12, pp. 4376–12, 2020.

[9] A. M. Yazid, J. K. Rijal, M. S. Awaluddin, and E. Sari, "Pattern recognition on remanufacturing automotive component as support decision making using Mahalanobis-Taguchi system," *Procedia CIRP*, vol. 26, pp. 258–263, 2015.

[10] N. N. Nik Mohd Kamil and M. Y. Abu, "Integration of Mahalanobis-Taguchi System and activity based costing for remanufacturing decision," *Journal of Modern Manufacturing Systems and Technology*, vol. 1, no. 1, pp. 39–51, 2018.

[11] J. Zhan, W. Chen, L. Cheng, Q. Wang, F. Han, and Y. Cui, "Diagnosis of asthma based on routine blood biomarkers using machine learning," *Computational Intelligence and Neuroscience*, vol. 20208 pages, 2020.

[12] X. Xiao, D. Fu, Y. Shi, and J. Wen, "Optimized mahalanobis-taguchi system for high-dimensional small sample data classification," *Computational Intelligence and Neuroscience*, vol. 202015 pages, 2020.

[13] S. Chen, Y. Liu, L. Wei, and B. Guan, "PS-FW: a hybrid algorithm based on particle swarm and fireworks for global optimization," *Computational Intelligence and Neuroscience*, vol. 201827 pages, 2018.

[14] M. El-Banna, "Modified mahalanobis taguchi system for imbalance data classification," *Computational Intelligence and Neuroscience*, vol. 201715 pages, 2017.

[15] K. Nishino and A. Suzuki, "Taguchi's T-method using median-median line for small sample with outliers," *IEEJ Transactions on Industry Applications*, vol. 138, no. 7, pp. 598–604, 2018.

[16] S. Negishi, Y. Morimoto, S. Takayama, and A. Ishigame, "Daily peak load forecasting by Taguchi's T method," *Electrical Engineering in Japan*, vol. 201, no. 1, pp. 57–65, 2017.

[17] N. Harudin, K. R. Jamaludin, M. Nabil Muhtazaruddin, F. Ramlie, and W. Z. A. W. Muhamad, "A feasibility study in adapting Shamos Bickel and Hodges Lehman estimator into T-Method for normalization," *IOP Conference Series: Materials Science and Engineering*, vol. 319, Article ID 012033, 2018.

[18] Z. M. Marlan, K. R. Jamaludin, F. Ramlie, N. Harudin, and N. N. Jaafar, "Determination of optimal unit space data for taguchi ' s T-method based on homogeneity of output," *Open International Journal of Informatics*, vol. 7, pp. 167–179, 2019, special issue.

[19] N. Harudin, "Increasing T-method accuracy through application of Robust M-estimator," *International Journal of Engineering & Technology*, vol. 7, no. 3, pp. 44–48, 2018.

[20] W. Z. A. W. Muhamad, F. Ramlie, and K. R. Jamaludin, "Mahalanobis-taguchi system for pattern recognition: a brief review," *Far East Journal of Mathematical Sciences (FJMS)*, vol. 102, no. 12, pp. 3021–3052, 2017.

[21] A. Pal and J. Maiti, "Development of a hybrid methodology for dimensionality reduction in Mahalanobis-Taguchi system using Mahalanobis distance and binary particle swarm optimization," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1286–1293, 2010.

[22] W. H. Woodall, R. Koudelik, K.-L. Tsui, S. B. Kim, Z. G. Stoumbos, and C. P. Carvounis, "Response," *Technometrics*, vol. 45, no. 1, pp. 29-30, 2003.

[23] W. Z. A. W. Muhamad, K. R. Jamaludin, S. A. Saad, Z. R. Yahya, and S. A. Zakaria, "Random binary search algorithm based feature selection in Mahalanobis Taguchi system for breast cancer diagnosis," *AIP Conference Proceedings*, vol. 1974, no. 1, Article ID 020027, 2018.

[24] B. Abraham and A. M. Variyath, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 22–24, 2003.

[25] C. R. Foster, R. Jugulum, and D. D. Frey, "Evaluating an adaptive one-factor-at-a-time search procedure within the mahalanobis-taguchi system," *International Journal of Industrial and Systems Engineering*, vol. 4, no. 6, pp. 600–614, 2009.

[26] D. M. Hawkins, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 25–29, 2003.

[27] K. Tsui, T. Sukchotrat, and V. C. P. Chen, "A comparison study and discussion of the mahalanobis-taguchi system seoung bum kim," *International Journal of Industrial and Systems Engineering*, vol. 4, no. 6, pp. 631–644, 2009.

[28] D. M. Hawkins, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 25–29, 2003.

[29] H. Kawada and Y. Nagata, "An application of a generalized inverse regression estimator to Taguchi's T-Method," *Total Quality Science*, vol. 1, no. 1, pp. 12–21, 2015.

[30] H. Kawada and Y. Nagata, "Studies on the item selection in taguchi's T-method," *Journal of the Japanese Society for Quality Control*, vol. 45, no. 2, pp. 179–193, 2015.

[31] N. Harudin, M. N. Muhtazaruddin, F. Ramlie, and W. M. Wan Zuki Azman, "Artificial bee colony for features selection optimization in increasing T-method accuracy," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 885–891, 2018.

[32] N. Harudin, K. R. Jamaludin, F. Ramlie, M. N. Muhtazaruddin, C. M. Che Razali, and W. Z. A. Wan Muhamad, "Binary particle swarm optimization for variables selection optimization in Taguchi's T-Method," *Matematika*, vol. 36, no. 1, pp. 69–84, 2020.

[33] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*, Wiley Series in Probability and Statistics, Hoboken, NJ, USA, 2nd edition, 2003.

[34] S. Teshima, Y. Hasegawa, and K. Tatebayashi, *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, Momentum Press, New York, NY, USA, 2012.

[35] S. Teshima, Y. Hasegawa, and K. Tatebayashi, "T-Method application procedures and Key points," in *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, pp. 87–104, Momentum Press, New York, NY, USA, 1st edition, 2012.

[36] D. Jia, X. Duan, and M. K. Khan, "Binary Artificial Bee Colony optimization using bitwise operation," *Computers & Industrial Engineering*, vol. 76, pp. 360–365, 2014.

[37] M. Lichman, *UCI Machine Learning Repository*, University of California, School of Information and Computer Science, Irvine, CA, USA, 2013, http://archive.ics.uci.edu/ml.

[38] C. M. C. Razali, S. F. M. Hussein, N. Harudin, and S. S. Abdullah, "Estimation of building energy efficiency performance using radial basis function neural network," *International Journal of Engineering & Technology*, vol. 7, no. 4.35, pp. 755–759, 2018.

[39] A. Tsanas and A. Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools," *Energy and Buildings*, vol. 49, pp. 560–567, 2012.

[40] N. Harudin, "An overview of taguchi' S T-method as A prediction tool for multivariate analysis," *Open International Journal of Informatics*, vol. 7, no. 1, 2019.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 5548102, 11 pages
https://doi.org/10.1155/2021/5548102

Research Article

General Minimum Lower-Order Confounding Designs with Multi-Block Variables

Yuna Zhao

School of Mathematics and Statistics, Shandong Normal University, Jinan 250358, China

Correspondence should be addressed to Yuna Zhao; yunazhao0504@163.com

Received 7 February 2021; Revised 31 March 2021; Accepted 8 April 2021; Published 21 April 2021

Academic Editor: Gengxin Sun

Blocking the inhomogeneous units of experiments into groups is an efficient way to reduce the influence of systematic sources on the estimations of treatment effects. In practice, there are two types of blocking problems. One considers only a single block variable and the other considers multi-block variables. The present paper considers the blocking problem of multi-block variables. Theoretical results and systematical construction methods of optimal blocked 2^{n-m} designs with $(N/4)+1 \le n \le 5N/16$ are developed under the prevalent general minimum lower-order confounding (GMC) criterion, where $N = 2^{n-m}$.

1. Introduction

The regular 2^{n-m} factorial experiment has played an important role in engineering, manufacturing industry, agriculture and medicine, and so on. It allows efficient and economic experimentation to estimate treatment effects. When the size of the experimental units is large, the inhomogeneity will cause unwanted variance to the estimations of treatment effects. To reduce such bad influence, a crucial way is to partition the experimental units into blocks.

There are two kinds of blocking problems as pointed out in [1]. One is called the single block variable problem which considers only a single block variable, and the other is called the multi-block variable problem which considers two or more block variables. In the last decades, choosing optimal blocked 2^{n-m} designs with a single block variable has been well investigated; for example, the authors in [2–9] studied the blocked 2^{n-m} designs under different minimum aberration criteria; Chen et al. and Zhao et al. [10, 11] explored the blocked 2^{n-m} designs under the clear effects criterion; Zhang and Mukerjee, Zhao et al., and Zhao and Zhao [12–14] explored the constructions of the blocked 2^{n-m} designs under the general minimum lower-order confounding (GMC) criterion proposed in [15]; and Zhao et al. [16–18] gave construction methods of the blocked 2^{n-m} designs under another GMC criterion proposed in [19].

Compared to the large body of work on the blocking problem of single block variable, the studies on choosing optimal blocked 2^{n-m} designs with multi-block variables are relatively rare. However, it has been recognized that the blocking problem of multi-block variables can arise quite naturally in many practical situations. For example, in the agricultural context, Bisgaard [1] pointed that when designs are laid out in rectangular schemes, both row and column inhomogeneity effects probably exist in the soil. Another example of multi-block variables is from [20]. Considering the comparison of two gasoline additives by testing them on two cars with two drivers over two days, the "cars," "drivers," and "days" are three block variables which should be taken into account when performing experiments.

Under the clear effects criterion, Zhao and Zhao [21] proposed an algorithm for finding optimal blocked 2^{n-m} designs with multi-block variables. Under the minimum aberration criterion, Zhao and Zhao [22] developed some rules for constructing optimal blocked 2^{n-m} designs with multi-block variables. Zhang et al. [23] extended the idea of GMC criterion to the case of multi-block variable problem and developed the blocked GMC criterion, called

the B^2-GMC criterion. Inheriting the advantage of the GMC criterion, the B^2-GMC designs are particularly preferable when some prior information on the importance ordering of treatment effects is present. By computer search, Zhang et al. [23] tabulated some B^2-GMC deigns with small n and N, where $N = 2^{n-m}$. When n or N is large, computer search becomes computationally expensive. Zhao et al. [24] and Zhao and Zhao [25] completed the constructions of B^2-GMC designs with $5N/16 + 1 \le n \le N - 1$. This paper aims at providing theories and systematical construction methods of the B^2-GMC designs with $N/4 + 1 \le n \le 5N/16$.

The rest of the paper is organized as follows. Section 2 reviews doubling theory and B^2-GMC criterion. Section 3 provides theoretical results and construction methods of B^2-GMC designs. Section 4 gives concluding remarks. Some useful lemmas are deferred to Appendix.

2. Preliminaries

2.1. Doubling Theory. Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_t)$ be a matrix with entries 1 or −1. Denote $\mathbf{J}_0 = (1, 1)'$ and $\mathbf{J}_1 = (1, -1)'$, where $'$ denotes transpose. Define

$$D(\mathbf{X}) = \left((\mathbf{J}_0, \mathbf{J}_1) \otimes \mathbf{x}_1, \ldots, (\mathbf{J}_0, \mathbf{J}_1) \otimes \mathbf{x}_t\right), \tag{1}$$

as a *double* of $\mathbf{X}$, where $\otimes$ is the Kronecker product. Let $D^q(\mathbf{X})$ denote the design obtained by repeatedly doubling $\mathbf{X}q$ times, i.e., $D^q(\mathbf{X}) = D(D^{q-1}(\mathbf{X}))$. When $\mathbf{X} = 1$, we write $D^q(1) = (\mathbf{I}_{2^q}, \mathbf{1}_{2^q}, \mathbf{2}_{2^q}, \mathbf{1}_{2^q}\mathbf{2}_{2^q}, \ldots, \mathbf{1}_{2^q}\mathbf{2}_{2^q}\mathbf{3}_{2^q}, \ldots, \mathbf{q}_{2^q})$, where the subscript 2^q means the dimension of a column, $\mathbf{I}_{2^q}$ is a column of 1's,

$$\begin{aligned}
&\mathbf{1}'_{2^q} = (1, \ldots, 1, -1, \ldots, -1)\\
&\mathbf{2}'_{2^q} = (1, \ldots, 1, -1, \ldots, -1, \ldots, 1, \ldots, 1, -1, \ldots, -1)\\
&\quad\vdots\\
&\mathbf{q}'_{2^q} = (1, -1, 1, -1, \ldots, 1, -1, 1, -1),
\end{aligned} \tag{2}$$

are q independent columns, and the other columns are the component-wise products of some of these q independent columns. For example, $\mathbf{1}_{2^q}\mathbf{2}_{2^q}$ is the component-wise product of the columns $\mathbf{1}_{2^q}$ and $\mathbf{2}_{2^q}$. Write $D^q(1) = (\mathbf{I}_{2^q}, D^q(\cdot))$; then, $D^q(\cdot)$ is just the regular two-level saturated design with columns arranged in Yates order. As some subdesigns in $D^q(\cdot)$, denote $\mathbf{H}_0^q = \varnothing$, $\mathbf{H}_1^q = \mathbf{1}_{2^q}, \mathbf{H}_r^q = (\mathbf{H}_{r-1}^q, \mathbf{r}_{2^q}, \mathbf{r}_{2^q}\mathbf{H}_{r-1}^q)$ for $r = 2, 3, \ldots, q$, where $\varnothing$ denotes the empty set, the superscript q of $\mathbf{H}_r^q$ refers to that $\mathbf{H}_r^q$ is a subdesign of $D^q(\cdot)$, and $\mathbf{r}_{2^q}\mathbf{H}_{r-1}^q = (\mathbf{r}_{2^q}\mathbf{d}_{2^q}: \mathbf{d}_{2^q} \in \mathbf{H}_{r-1}^q)$, i.e., $\mathbf{H}_r^q$ consists of the first $2^r - 1$ columns of $D^q(\cdot)$. Especially, $\mathbf{H}_q^q = D^q(\cdot)$.

2.2. B^2-GMC Criterion. Before introducing the B^2-GMC criterion, we first review some principles in the multi-block variable problem. Let $b_1, b_2, \ldots, b_s$ denote the s block variables which cause the inhomogeneity of the experimental units. Suppose that the block variable b_j partitions the N (= 2^{n-m}) experimental units into 2^{l_j} blocks; then, l_j independent columns are needed to carry out this blocking plan. Denote S_j as the set of the l_j independent columns related to the block variable b_j. The block columns should follow the following rules:

(i) The l_j block columns in S_j $(j = 1, 2, \ldots, s)$ are independent of each other.

(ii) A block column from S_j is not necessarily independent of the block columns from S_i with $j \ne i$.

In this paper, we focus on the case where each block variable is at two levels, i.e., $l_j = 1$.

The effect hierarchy principle for blocked designs with multi-block variables is as follows (see [23]):

(i) The lower-order treatment factorial effects are more likely to be important than the higher-order ones, and the treatment factorial effects of the same order are equally likely to be important.

(ii) The lower-order block factorial effects are more likely to be important than the higher-order ones, and the block factorial effects of the same order are equally likely to be important.

(iii) All the interactions between treatment factors and block factors are negligible.

Since each variable or factor is assigned to one column of the design matrix when an experiment is carried out, we do not differentiate the variable, factor, and column. Based on the effect hierarchy principle and weak assumption that the effects involving three or more factors are usually not important and negligible, Zhang et al. [23] proposed the B^2-GMC criterion which pays attention to only the confounding among main treatment effects and the two-factor interactions of treatment factors (2fi's for short). For the same reason, a common assumption in blocking problem is that only the main effects of block variables and the interactions of any two block factors are potentially significant, and if a treatment effect is confounded with a potentially significant block effect, the treatment effect cannot be estimated. Thus, the confounding between the main effects of treatment factors and any potentially significant block effect is not allowed.

Denote $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ as a $2^{n-m}: 2^s$ design, where $\mathbf{D}_t$ consists of n treatment factors corresponding to a regular 2^{n-m} design and $\mathbf{D}_b$ consists of s block factors each of which can partition the 2^{n-m} runs into 2 blocks. Denote ${}^{\#}_1C_2^{(p)}(\mathbf{D})$ as the number of main treatment effects which are aliased with p2fi's but not with any potentially significant block effects, where $p = 0, 1, 2, \ldots, P$ and $P = n(n-1)/2$. Similarly, ${}^{\#}_2C_2^{(p)}(\mathbf{D})$ denotes the number of 2fi's which are aliased with the other p2fi's but not with any potentially significant block effects, where $p = 0, 1, \ldots, P$. Denote

$$\begin{aligned}
{}^{\#}_1C_2(\mathbf{D}) &= \left({}^{\#}_1C_2^{(0)}(\mathbf{D}), {}^{\#}_1C_2^{(1)}(\mathbf{D}), \ldots, {}^{\#}_1C_2^{(p)}(\mathbf{D})\right),\\
{}^{\#}_2C_2(\mathbf{D}) &= \left({}^{\#}_2C_2^{(0)}(\mathbf{D}), {}^{\#}_2C_2^{(1)}(\mathbf{D}), \ldots, {}^{\#}_2C_2^{(p)}(\mathbf{D})\right),
\end{aligned} \tag{3}$$

$$ {}^{\#}C(\mathbf{D}) = \left({}^{\#}_1C_2(\mathbf{D}), {}^{\#}_2C_2(\mathbf{D})\right). \tag{4}$$

A 2^{n-m}: 2^s blocked design $\mathbf{D} = (\mathbf{D}_t\colon \mathbf{D}_b)$ is called a B^2-GMC design if it sequentially maximizes (4). Let ${}^{\#}_{1}C_2^{(p)}(\mathbf{D}_t)$ be the number of main effects which are aliased with p 2fi's of $\mathbf{D}_t$ and ${}^{\#}_{2}C_2^{(p)}(\mathbf{D}_t)$ be the number of 2fi's which are aliased with the other p 2fi's of $\mathbf{D}_t$. Let

$$
\begin{aligned}
{}^{\#}_{1}C_2(\mathbf{D}_t) &= \left({}^{\#}_{1}C_2^{(0)}(\mathbf{D}_t), {}^{\#}_{1}C_2^{(1)}(\mathbf{D}_t), \ldots, {}^{\#}_{1}C_2^{(p)}(\mathbf{D}_t)\right),\\
{}^{\#}_{2}C_2(\mathbf{D}_t) &= \left({}^{\#}_{2}C_2^{(0)}(\mathbf{D}_t), {}^{\#}_{2}C_2^{(1)}(\mathbf{D}_t), \ldots, {}^{\#}_{2}C_2^{(p)}(\mathbf{D}_t)\right),\\
{}^{\#}C(\mathbf{D}_t) &= \left({}^{\#}_{1}C_2(\mathbf{D}_t), {}^{\#}_{2}C_2(\mathbf{D}_t)\right).
\end{aligned}
\tag{5}
$$

A 2^{n-m} design $\mathbf{D}_t$ is called a GMC design if $\mathbf{D}_t$ sequentially maximizes (5).

Let $q = n - m$. Constructing a B^2-GMC design is to choose $\mathbf{D}_t$ and $\mathbf{D}_b$ from $D^q(\cdot)$ such that (4) is sequentially maximized. In the following, without causing confusions, we omit the subscript of a column and the superscript of a design when they are taken from $D^q(\cdot)$. For example, we use $\mathbf{a}$, $\mathbf{H}_r$, and $\mathbf{H}_q$ instead of $\mathbf{a}_{2^q}$, $\mathbf{H}_r^q$, and $\mathbf{H}_q^q$, respectively. Denote

$$
U(\mathbf{D}_b) = \left\{\boldsymbol{\gamma} \in \mathbf{H}_q\colon\ \boldsymbol{\gamma} \in \mathbf{D}_b \text{ or } \boldsymbol{\gamma} = \mathbf{ab} \text{ with } \mathbf{a}, \mathbf{b} \in \mathbf{D}_b\right\}, \tag{6}
$$

and then $U(\mathbf{D}_b)$ consists of all the potentially significant block effects. As previously stated, the confounding between main treatment effects and potentially significant block effects is not allowed. This requires $\mathbf{D}_t \cap U(\mathbf{D}_b) = \varnothing$, and thus ${}^{\#}_{1}C_2^{(p)}(\mathbf{D}) = {}^{\#}_{1}C_2^{(p)}(\mathbf{D}_t)$ for $p = 0, 1, 2, \ldots, P$.

For $\mathbf{D}_t \subset \mathbf{H}_q$ and $\boldsymbol{\gamma} \in \mathbf{H}_q$, define

$$
B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = \#\{(\mathbf{d}_1, \mathbf{d}_2)\colon \mathbf{d}_1, \mathbf{d}_2 \in \mathbf{D}_t, \mathbf{d}_1\mathbf{d}_2 = \boldsymbol{\gamma}\}, \tag{7}
$$

where # denotes the cardinality of a set and $\mathbf{d}_1\mathbf{d}_2$ stands for the two-factor interaction of $\mathbf{d}_1$ and $\mathbf{d}_2$. Thus, $B_2(\mathbf{D}_t, \boldsymbol{\gamma})$ is the number of 2fi's of $\mathbf{D}_t$ appearing in the alias set that contains $\boldsymbol{\gamma}$.

Isomorphism introduced by Tang and Wu [26] is a useful concept which helps narrow down the search of the optimal blocked designs here. An isomorphism ϕ is a one-to-one mapping from $\mathbf{H}_q$ to $\mathbf{H}_q$ such that $\phi(\mathbf{xy}) = \phi(\mathbf{x})\phi(\mathbf{y})$ for every $\mathbf{x} \neq \mathbf{y} \in \mathbf{H}_q$. The 2^{n-m} designs $\mathbf{D}_t$ and $\mathbf{D}_t^*$ are isomorphic if there exists an isomorphism ϕ that maps $\mathbf{D}_t$ onto $\mathbf{D}_t^*$. The 2^{n-m}: 2^s designs $\mathbf{D} = (\mathbf{D}_t\colon \mathbf{D}_b)$ and $\mathbf{D}^* = (\mathbf{D}_t^*\colon \mathbf{D}_b^*)$ are isomorphic if there exists an isomorphism ϕ that maps $\mathbf{D}_t$ onto $\mathbf{D}_t^*$ and $\mathbf{D}_b$ onto $\mathbf{D}_b^*$.

3. Constructions of B^2-GMC Designs

3.1. B^2-GMC 2^{n-m}: 2^s Designs with $n = (N/4) + 1$. A design is of MaxC2 (see [27]) if it has resolution IV and maximum number of clear 2fi's, where a *resolution R* design has no c-factor interaction confounded with any other interaction involving less than $R - c$ factors (see [28]), and a 2fi is called clear if it is not aliased with any main treatment effect and other 2fi's. Cheng and Zhang [29] showed that a 2^{n-m} design with $n = N/4 + 1$ is a MaxC2 design if and only if it is a GMC design. They also pointed out that, up to isomorphism, the GMC 2^{n-m} design with $n = N/4 + 1$ can be uniquely expressed as $\mathbf{S}_{N/4+1} = (\mathbf{q} - \mathbf{1}, \mathbf{q}, \mathbf{q}(\mathbf{q} - \mathbf{1})\mathbf{H}_{q-2})$. It is easy to obtain that

$$
B_2(\mathbf{S}_{N/4+1}, \boldsymbol{\gamma}) = \begin{cases} 1, & \text{for } \boldsymbol{\gamma} \in \left\{(\mathbf{q} - \mathbf{1})\mathbf{H}_{q-2}, \mathbf{q}\mathbf{H}_{q-2}, \mathbf{q}(\mathbf{q} - \mathbf{1})\right\},\\ \dfrac{N}{8} - 1, & \text{for } \boldsymbol{\gamma} \in \mathbf{H}_{q-2},\\ 0, & \text{otherwise.} \end{cases} \tag{8}
$$

Therefore,

$$
{}^{\#}_{1}C_2^{(p)}(\mathbf{S}_{N/4+1}) = \begin{cases} \dfrac{N}{4} + 1, & \text{for } p = 0,\\ 0, & \text{otherwise,} \end{cases} \tag{9}
$$

$$
{}^{\#}_{2}C_2^{(p)}(\mathbf{S}_{N/4+1}) = \begin{cases} 2\left(\dfrac{N}{4} - 1\right) + 1, & \text{for } p = 0,\\ \left(\dfrac{N}{4} - 1\right)\left(\dfrac{N}{8} - 1\right), & \text{for } p = \dfrac{N}{8} - 2,\\ 0, & \text{otherwise.} \end{cases} \tag{10}
$$

Lemma 1 is a straightforward extension of Lemma A.1, in Appendix, introduced from [24, 25].

Lemma 1. *Suppose $\mathbf{D}_b$ is any s-projection of $\{\mathbf{H}_r, \boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_r\}$ with $2^k \le s \le 2^{k+1} - 1$ for some $k\,(0 \le k \le r - 1)$, where $\boldsymbol{\gamma}$ is independent of the columns of $\mathbf{H}_r$.* We have

(i) If $\mathbf{D}_b \cap \{\boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_r\} = \varnothing$, then $\#\{U(\mathbf{D}_b) \cap \mathbf{H}_r\} \ge 2^{k+1} - 1$ and the equality holds when $\mathbf{D}_b$ has $k + 1$ independent columns.

(ii) If $\mathbf{D}_b \cap \{\boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_r\} \neq \varnothing$, then $\#\{U(\mathbf{D}_b) \cap \mathbf{H}_r\} \ge 2^k - 1$ and the equality holds when $\mathbf{D}_b$ has $k + 1$ independent columns.

(iii) If $\mathbf{D}_b \cap \{\boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_r\} \neq \varnothing$, then $\#\{U(\mathbf{D}_b) \cap \{\boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_r\}\} \ge 2^k$ and the equality holds when $\mathbf{D}_b$ has $k + 1$ independent columns.

(iv) If $\mathbf{D}_b \subset \mathbf{H}_{k+1}$, then $U(\mathbf{D}_b) = \mathbf{H}_{k+1}$.

(v) If $\mathbf{D}_b \subset \{\mathbf{H}_k, \boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_k\}$, then $U(\mathbf{D}_b) = \{\mathbf{H}_k, \boldsymbol{\gamma}, \boldsymbol{\gamma}\mathbf{H}_k\}$.

Lemma 2 provides a necessary condition for a 2^{n-m}: 2^s design $\mathbf{D} = (\mathbf{D}_t\colon \mathbf{D}_b)$ with $n = N/4 + 1$ to be a B^2-GMC design.

Lemma 2. *Suppose $\mathbf{D} = (\mathbf{D}_t\colon \mathbf{D}_b)$ is a 2^{n-m}: 2^s design with $n = N/4 + 1$ and $2^k \le s \le 2^{k+1} - 1$ for some $k\,(0 \le k \le q - 3)$; then, $\mathbf{D} = (\mathbf{D}_t\colon \mathbf{D}_b)$ is a B^2-GMC design only if $\mathbf{D}_t = \mathbf{S}_{N/4+1}$.*

Proof. Let $\widetilde{\mathbf{D}} = (\mathbf{S}_{N/4+1}\colon \widetilde{\mathbf{D}}_b)$ be a 2^{n-m}: 2^s design and $\widetilde{\mathbf{D}}_b \in \mathbf{H}_{q-2}$. From (8)–(10), we can obtain

$$
{}^{\#}_{1}C_2(\widetilde{\mathbf{D}}) = \left(\frac{N}{4} + 1, 0, \ldots, 0\right), \tag{11}
$$

which is sequentially maximized by $\widetilde{\mathbf{D}}$, and ${}^{\#}_{2}C_2^{(0)}(\widetilde{\mathbf{D}}) = 2(N/4 - 1) + 1 = N/2 - 1$. Suppose that any $\mathbf{D} = (\mathbf{S}_{N/4+1}\colon \mathbf{D}_b)$ is not a B^2-GMC 2^{n-m}: 2^s design for $n = N/4 + 1$. Then,

by the definition of B^2-GMC criterion, there should be a 2^{n-m}: 2^s design $\mathbf{D}^* = (\mathbf{D}_t^*: \mathbf{D}_b^*)$ outperforming $\widetilde{D} = (\mathbf{S}_{N/4+1}: \widetilde{D}_b)$ in terms of (4). This leads to

$$ {}^{\#}_{1}C_2(\mathbf{D}^*) = \left(\frac{N}{4}+1, 0, \ldots, 0\right), \tag{12} $$

and ${}^{\#}_{2}C_2^{(0)}(\mathbf{D}^*) \geq (N/2) - 1$. Clearly,

$$ {}^{\#}_{1}C_2(\mathbf{D}_t^*) = {}^{\#}_{1}C_2(\mathbf{D}^*) = \left(\frac{N}{4}+1, 0, \ldots, 0\right), \tag{13} $$

and ${}^{\#}_{2}C_2^{(0)}(\mathbf{D}_t^*) \geq {}^{\#}_{2}C_2^{(0)}(\mathbf{D}^*) \geq N/2 - 1$ noting that $\mathbf{D}_t^*$ is the unblocked part of $\mathbf{D}^*$ and $\mathbf{D}_t^*$ has $n = N/4 + 1$ columns.

In fact, the inequality in the formula ${}^{\#}_{2}C_2^{(0)}(\mathbf{D}_t^*) \geq N/2 - 1$ is not valid. Recall that the MaxC2 design $\mathbf{S}_{N/4+1}$ has the largest number, $N/2 - 1$, of clear 2fi's among all the 2^{n-m} designs with $n = N/4 + 1$. Therefore, $\mathbf{D}_t^*$ has at most $N/2 - 1$ clear 2fi's, i.e., ${}^{\#}_{2}C_2^{(0)}(\mathbf{D}_t^*) = N/2 - 1$. Wu and Wu [30] showed that a 2^{n-m} design with $n = N/4 + 1$ is a MaxC2 design if and only if this design has $N/2 - 1$ clear 2fi's. This obtains that $\mathbf{D}_t^* = \mathbf{S}_{N/4+1}$ up to isomorphism. □

With Lemma 2, Theorem 1 provides the constructions of B^2-GMC designs with $n = N/4 + 1$.

Theorem 1. *Suppose* $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ *is a* 2^{n-m}: 2^s *design with* $n = N/4 + 1$ *and* $2^k \leq s \leq 2^{k+1} - 1$ *for some* $k\,(0 \leq k \leq q-3)$; *then,* $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ *is a* B^2*-GMC design if* $\mathbf{D}_t = \mathbf{S}_{N/4+1}$ *and* $\mathbf{D}_b$ *is any s-projection of* $\mathbf{H}_{k+1}$.

Proof. By Lemma 2, if $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ is a B^2-GMC 2^{n-m}: 2^s design with $n = N/4 + 1$, then $\mathbf{D}_t = \mathbf{S}_{N/4+1}$ up to isomorphism. Thus, ${}^{\#}_{1}C_2(\mathbf{D}) = (N/4 + 1, 0, \ldots, 0)$ which is sequentially maximized by $\mathbf{D} = (\mathbf{S}_{N/4+1}: \mathbf{D}_b)$. Let $u_1 = \#\left\{U(\mathbf{D}_b) \cap \left\{(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-2}, \mathbf{q}\mathbf{H}_{q-2}, \mathbf{q}(\mathbf{q}-\mathbf{1})\right\}\right\}$ and $u_2 = \#\left\{U(\mathbf{D}_b) \cap \mathbf{H}_{q-2}\right\}$; then, from (10), we have

$$ {}^{\#}_{2}C_2^{(p)}(\mathbf{D}) = \begin{cases} 2\left(\dfrac{N}{4}-1\right) + 1 - u_1, & \text{for } p = 0, \\ \left(\dfrac{N}{4} - 1 - u_2\right)\left(\dfrac{N}{8} - 1\right), & \text{for } p = \dfrac{N}{8} - 2, \\ 0, & \text{otherwise.} \end{cases} \tag{14} $$

If $\mathbf{D}$ is a B^2-GMC design, then $\mathbf{D}$ must sequentially minimize (u_1, u_2). There are two different ways to choose $\mathbf{D}_b$ from $\mathbf{H}_q \backslash \mathbf{S}_{N/4+1}$:

(i) $\mathbf{D}_b \cap \left\{(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-2}, \mathbf{q}\mathbf{H}_{q-2}, \mathbf{q}(\mathbf{q}-\mathbf{1})\right\} = \varnothing$.

(ii) $\mathbf{D}_b \cap \left\{(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-2}, \mathbf{q}\mathbf{H}_{q-2}, \mathbf{q}(\mathbf{q}-\mathbf{1})\right\} \neq \varnothing$.

It is not hard to verify that $\mathbf{D}_b$ in case (ii) results in $u_1 > 0$ while $\mathbf{D}_b$ in case (i) gives $u_1 = 0$. Therefore, if $\mathbf{D} = (\mathbf{S}_{N/4+1}: \mathbf{D}_b)$ is a B^2-GMC design, then $\mathbf{D}_b$ must be of case (i). Recall that $\mathbf{D}_b \cap \mathbf{S}_{N/4+1} = \varnothing$, and we have $\mathbf{D}_b \subset \mathbf{H}_{q-2}$. By Lemma 1 (i), if $\mathbf{D}_b \subset \mathbf{H}_{q-2}$, then $u_2 = \#\left\{U(\mathbf{D}_b) \cap \mathbf{H}_{q-2}\right\} \geq 2^{k+1} - 1$ and the equality holds when $\mathbf{D}_b \subset \mathbf{H}_{k+1}$ up to isomorphism. This completes the proof. □

The following example illustrates the construction method in Theorem 1.

Example 1. Consider the construction of the B^2-GMC 2^{9-4}: 2^5 design. Here, $n = 9, m = 4, q = 5$, $N = 32$, and $s = 5$ which leads to $k = 2$. According to Theorem 1, let $\mathbf{D}_t = (\mathbf{4}, \mathbf{5}, \mathbf{45H}_3)$ and $\mathbf{D}_b$ be any 5-projection of $\mathbf{H}_3$, say $\mathbf{D}_b = (\mathbf{1}, \mathbf{2}, \mathbf{12}, \mathbf{3}, \mathbf{13})$. Then, $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ is a B^2-GMC 2^{9-4}: 2^5 design.

3.2. B^2-GMC Designs with $N/4 + 1 < n \leq 5N/16$. To sequentially maximize (4), the first part ${}^{\#}_{1}C_2(\mathbf{D})$ of (4) should be first maximized. Recall that ${}^{\#}_{1}C_2(\mathbf{D}) = {}^{\#}_{1}C_2(\mathbf{D}_t)$. If $\mathbf{D}_t$ has resolution IV, then ${}^{\#}_{1}C_2(\mathbf{D})$ must be maximized. According to [31], when $N/4 + 1 < n \leq 5N/16$, the $\mathbf{D}_t$ with resolution IV must be an n-projection of some second-order saturated (SOS) designs. In the following, we first review the concept of SOS design.

A 2^{n-m} design is called an SOS design if all of its degrees of freedom can be used to estimate only the main treatment effects and 2fi's. In terms of coding theory and projective geometry, Davydov and Tombak [32] showed that, given N, only the SOS designs of $N/4 + 1$, $N(2^{w-2} + 1)/2^w$ with $w \geq 4$ and $N/2$ factors exist. Block and Mee [31] further showed that an SOS design of $N(2^{w-2} + 1)/2^w$ factors can be obtained by doubling some smaller SOS $2^{(2^{w-2}+1)-(2^{w-2}+1-w)}$ designs $q - w$ times. Zhang and Cheng [33] showed that the SOS design of $N/2$ factors can be uniquely represented by $\mathbf{S}_{N/2} = (\mathbf{q}, \mathbf{qH}_{q-1})$ up to isomorphism. Let

$$ \mathscr{L}(w) = \left\{D^{q-w}(\mathbf{Y}): \mathbf{Y} \text{ is an SOS } 2^{(2^{w-2}+1)-(2^{w-2}+1-w)} \text{ design}\right\}, \tag{15} $$

denote the collection of all the SOS designs obtained by doubling some SOS $2^{(2^{w-2}+1)-(2^{w-2}+1-w)}$ designs $q - w$ times. Especially, in $\mathscr{L}(w)$, we denote the SOS design obtained by doubling the MaxC2 $2^{(2^{w-2}+1)-(2^{w-2}+1-w)}$ design $\Phi(w) = ((\mathbf{w}-\mathbf{1})_{2^w}, \mathbf{w}_{2^w}, \mathbf{w}_{2^w}(\mathbf{w}-\mathbf{1})_{2^w}\mathbf{H}_{w-2}^w)$ as $\mathbf{S}_{N(2^{w-2}+1)/2^w}$.

With a little algebra, it is easy to verify that

$$ \begin{aligned} D^{q-w}(\mathbf{I}_{2^w}) &= \left(\mathbf{I}_{2^q}, \mathbf{H}_{q-w}^q\right) = \left(\mathbf{I}, \mathbf{H}_{q-w}\right) \\ D^{q-w}(\mathbf{1}_{2^w}) &= (\mathbf{q}-\mathbf{w}+\mathbf{1})_{2^q}\left(\mathbf{I}_{2^q}, \mathbf{H}_{q-w}^q\right) \\ &= (\mathbf{q}-\mathbf{w}+\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-w}\right) \\ D^{q-w}(\mathbf{2}_{2^w}) &= (\mathbf{q}-\mathbf{w}+\mathbf{2})_{2^q}\left(\mathbf{I}_{2^q}, \mathbf{H}_{q-w}^q\right) \\ &= (\mathbf{q}-\mathbf{w}+\mathbf{2})\left(\mathbf{I}, \mathbf{H}_{q-w}\right) \\ &\vdots \\ D^{q-w}((\mathbf{w}-\mathbf{1})_{2^w}) &= (\mathbf{q}-\mathbf{1})_{2^q}\left(\mathbf{I}_{2^q}, \mathbf{H}_{q-w}^q\right) = (\mathbf{q}-\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-w}\right) \\ D^{q-w}(\mathbf{w}_{2^w}) &= \mathbf{q}_{2^q}\left(\mathbf{I}_{2^q}, \mathbf{H}_{q-w}^q\right) = \mathbf{q}\left(\mathbf{I}, \mathbf{H}_{q-w}\right), \end{aligned} \tag{16} $$

where $\mathbf{I}_{2^{q-w}} \otimes \mathbf{1}_{2^w} = (\mathbf{q}-\mathbf{w}+\mathbf{1})_{2^q}$, $\mathbf{I}_{2^{q-w}} \otimes \mathbf{2}_{2^w} = (\mathbf{q}-\mathbf{w}+\mathbf{2})_{2^q}$, $\ldots$, $\mathbf{I}_{2^{q-w}} \otimes (\mathbf{w}-\mathbf{1})_{2^w} = (\mathbf{q}-\mathbf{1})_{2^q}$, and $\mathbf{I}_{2^{q-w}} \otimes \mathbf{w}_{2^w} = \mathbf{q}_{2^q}$. Then,

$$D^{q-w}\left((\mathbf{w}-\mathbf{1})_{2^w}\mathbf{w}_{2^w}\mathbf{H}_{w-2}^{w}\right)=\mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{H}_{q-2}\backslash\mathbf{H}_{q-w}\right). \tag{17}$$

Therefore, $\mathbf{S}_{N(2^{w-2}+1)/2^w}$ can be expressed as

$$\begin{aligned}\mathbf{S}_{N(2^{w-2}+1)/2^w}&=D^{q-w}(\Phi(w))\\&=\Big((\mathbf{q}-\mathbf{1}),(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w},\mathbf{q},\mathbf{q}\mathbf{H}_{q-w},\mathbf{q}(\mathbf{q}-\mathbf{1})\\&\quad\left(\mathbf{H}_{q-2}\backslash\mathbf{H}_{q-w}\right)\Big).\end{aligned} \tag{18}$$

The following lemma provides a necessary condition for a 2^{n-m}: 2^s design $\mathbf{D}=(\mathbf{D}_t:\mathbf{D}_b)$ with $N/4+1<n\le 5N/16$ to be a B^2-GMC design.

Lemma 3. *Suppose* $\mathbf{D}=(\mathbf{D}_t:\mathbf{D}_b)$ *is a* 2^{n-m}: 2^s *design with* $N(2^{w-1}+1)/2^{w+1}<n\le N(2^{w-2}+1)/2^w$ *for* $w\ge 4$ *and* $2^k\le s\le 2^{k+1}-1$ *for some* $k(0\le k\le q-3)$; *then,* $\mathbf{D}=(\mathbf{D}_t:\mathbf{D}_b)$ *is a* B^2*-GMC design only if* $\mathbf{D}_t\subset\mathbf{S}_{N(2^{w-2}+1)/2^w}$.

Proof. As discussed in the first paragraph of this section, if $\mathbf{D}=(\mathbf{D}_t:\mathbf{D}_b)$ is a B^2-GMC 2^{n-m}: 2^s design with $N(2^{w-1}+1)/2^{w+1}<n\le N(2^{w-2}+1)/2^w$, then $\mathbf{D}_t$ must be an n-projection of some SOS designs in $\mathscr{L}(v)$ with $4\le v\le w$ or $\mathbf{S}_{N/2}$.

Let $\mathbf{P}_1$ be an n-projection of $\mathbf{S}_{N/2}$. According to Theorem 3 in [33], we obtain

$${}_2^{\#}C_2^{(p)}(\mathbf{P}_1)=\begin{cases}0, & \text{for } p<n-\dfrac{N}{4}-1,\\[2mm] \dfrac{N}{4}\left(n-\dfrac{N}{4}\right), & \text{for } p=n-\dfrac{N}{4}-1,\\[2mm] 0, & \text{for } n-\dfrac{N}{4}-1<p<\dfrac{N}{8}-1.\end{cases} \tag{19}$$

Let $\mathbf{X}_v=(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_{2^{v-2}+1})$ be any SOS $2^{(2^{v-2}+1)-(2^{v-2}+1-v)}$ design but not a MaxC2 design; then, $D^{q-v}(\mathbf{X}_v)\in\mathscr{L}(v)$. Denote p_i as the number of clear 2fi's of $\mathbf{X}_v$ involving $\mathbf{a}_i$ for $i=1,2,\ldots,2^{v-2}+1$ and c as the total number of clear 2fi's of $\mathbf{X}_v$. Let $\mathbf{P}_v$ be an n-projection of $D^{q-v}(\mathbf{X}_v)$ and $\overline{\mathbf{P}}_v=D^{q-v}(\mathbf{X}_v)\backslash\mathbf{P}_v$. Then, $\#\overline{\mathbf{P}}_v=N(2^{v-2}+1)/2^v-n\le N/2^v$. Applying equations (23) and (24) in the proof of Theorem 3.1 of [29], among all the n-projections of $D^{q-v}(\mathbf{X}_v)$, $\mathbf{P}_v$ sequentially maximizes ${}_2^{\#}C_2$ in (5), only if $\overline{\mathbf{P}}_v\subset D^{q-v}(\mathbf{a}_j)$, where $\mathbf{a}_j$ is the column such that $p_j=\max\{p_1,p_2,\ldots,p_{2^{v-2}+1}\}$. For a $\mathbf{P}_v$ with $\overline{\mathbf{P}}_v\subset D^{q-v}(\mathbf{a}_j)$, according to equation (24) in the proof of Theorem 3.1 of [29], it is obtained that

$${}_2^{\#}C_2^{(p)}(\mathbf{P}_v)=\begin{cases}0, & \text{for } 0\le p<n-\dfrac{N}{4}-1,\\[2mm] \dfrac{p_jN}{2^v}(n-N/4), & \text{for } p=n-\dfrac{N}{4}-1,\\[2mm] 0, & \text{for } n-\dfrac{N}{4}-1<p<\dfrac{N}{2^v}-1,\\[2mm] \left(c-p_j\right)\left(\dfrac{N}{2^v}\right)^2, & \text{for } p=\dfrac{N}{2^v}-1.\end{cases} \tag{20}$$

For a MaxC22$^{(2^{v-2}+1)-(2^{v-2}+1-v)}$ design, it has $2^{v-1}-1$ clear 2fi's and $p_j=2^{v-2}$. Since $\mathbf{X}_v$ is not a MaxC2 design, we have

$$\begin{aligned}p_j&=2^{v-2},\\ c-p_j&<2^{v-1}-2^{v-2}-1,\end{aligned} \tag{21}$$

or

$$\begin{aligned}p_j&<2^{v-2},\\ c&<2^{v-1}-1.\end{aligned} \tag{22}$$

Let $\mathbf{Q}_v$ be an n-projection of $\mathbf{S}_{N(2^{v-2}+1)/2^v}$ obtained by doubling $\Phi(v)$, the MaxC22$^{(2^{v-2}+1)-(2^{v-2}+1-v)}$ design as mentioned above. With a column permutation, rewrite $D^q(\cdot)$ as

$$\begin{aligned}D_{\mathrm{RC}}^{q}(\cdot)=\Big(&\mathbf{H}_{q-v},\mathbf{H}_{q-2}\backslash\mathbf{H}_{q-v},(\mathbf{q}-\mathbf{1})\left(\mathbf{H}_{q-2}\backslash\mathbf{H}_{q-v}\right),\\&\mathbf{q}\left(\mathbf{H}_{q-2}\backslash\mathbf{H}_{q-v}\right),\mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{I},\mathbf{H}_{q-v}\right),\mathbf{S}_{N(2^{v-2}+1)/2^v}\Big),\end{aligned} \tag{23}$$

in a re-changed Yates order. Write $\Phi(v)$ as $\Phi(v)=(\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_{2^{v-2}+1})$, where $\mathbf{s}_1=(\mathbf{v}-\mathbf{1})_{2^v}$, $\mathbf{s}_2=\mathbf{v}_{2^v}$ and $(\mathbf{s}_3,\ldots,\mathbf{s}_{2^{v-2}+1})=\mathbf{v}_{2^v}(\mathbf{v}-\mathbf{1})_{2^v}\mathbf{H}_{v-2}^{v}$. Applying equations (23) and (24) in the proof of Theorem 3.1 of [29], among all the n-projections of $\mathbf{S}_{N(2^{v-2}+1)/2^v}$, $\mathbf{Q}_v$ sequentially maximizes ${}_2^{\#}C_2$ in (3), only if $\overline{\mathbf{Q}}_v\subset D^{q-v}(\mathbf{s}_i)$ with $i=1$ or 2. In the following, we investigate ${}_2^{\#}C_2^{(p)}(\mathbf{Q}_v)$ with $\overline{\mathbf{Q}}_v\subset D^{q-v}(\mathbf{s}_1)$ for which the following analysis and final conclusion are the same as that for ${}_2^{\#}C_2^{(p)}(\mathbf{Q}_v)$ with $\overline{\mathbf{Q}}_v\subset D^{q-v}(\mathbf{s}_2)$. For $\mathbf{Q}_v$ with $\overline{\mathbf{Q}}_v\subset D^{q-v}(\mathbf{s}_1)$, from equations (18) and (23), it is obtained that

$$B_2(\mathbf{Q}_v, \boldsymbol{\gamma}) = \begin{cases} 0, & \text{for } \boldsymbol{\gamma} \in \mathbf{S}_{N(2^{v-2}+1)/2^v}, \\ n - \dfrac{N}{4}, & \text{for } \boldsymbol{\gamma} \in \left\{\mathbf{q}\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-v}\right), \mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-v}\right)\right\}, \\ \dfrac{N}{2^v}, & \text{for } \boldsymbol{\gamma} \in (\mathbf{q}-\mathbf{1})\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-v}\right), \\ \dfrac{N}{8} - \dfrac{N}{2^v}, & \text{for } \boldsymbol{\gamma} \in \mathbf{H}_{q-2} \backslash \mathbf{H}_{q-v}, \\ B_2\left(\overline{Q}_v, \boldsymbol{\gamma}\right) + \dfrac{N}{8}, & \text{for } \boldsymbol{\gamma} \in \mathbf{H}_{q-v}, \end{cases} \tag{24}$$

where the second equality can be easily verified by noting that

$$\begin{aligned} &\left\{\mathbf{q}\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-v}\right), \mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-v}\right)\right\} \\ &= \left\{D^{q-v}(\mathbf{s}_1\mathbf{s}_2), D^{q-v}(\mathbf{s}_1\mathbf{s}_3), \ldots, D^{q-v}(\mathbf{s}_1\mathbf{s}_{2^{v-2}+1})\right\}. \end{aligned} \tag{25}$$

Therefore,

$${}_2^{\#}C_2^{(p)}(\mathbf{Q}_v) = \begin{cases} 0, & \text{for } 0 \le p < n - \dfrac{N}{4} - 1, \\ \dfrac{N}{4}(n - N/4), & \text{for } p = n - \dfrac{N}{4} - 1, \\ 0, & \text{for } n - \dfrac{N}{4} - 1 < p < \dfrac{N}{2^v} - 1, \\ \dfrac{N}{2^v}\left(N/4 - N/2^v\right), & \text{for } p = \dfrac{N}{2^v} - 1, \\ 0, & \text{for } \dfrac{N}{2^v} - 1 < p < \dfrac{N}{8} - \dfrac{N}{2^v} - 1, \\ \left(\dfrac{N}{4} - \dfrac{N}{2^v}\right)\left(\dfrac{N}{8} - \dfrac{N}{2^v}\right), & \text{for } p = \dfrac{N}{8} - \dfrac{N}{2^v} - 1. \end{cases} \tag{26}$$

Denote $\mathbf{D}^* = (\mathbf{D}_t^*: \mathbf{D}_b^*)$, where $\mathbf{D}_t^* \subset \mathbf{S}_{N(2^{w-2}+1)/2^w}$, $\mathbf{S}_{N(2^{w-2}+1)/2^w} \backslash \mathbf{D}_t^* \subset D^{q-w}(\mathbf{s}_1)$ and $\mathbf{D}_b^* \subset \mathbf{H}_{k+1}$. Since $k \le q-3$, we have $U(\mathbf{D}_b^*) \subset \mathbf{H}_{q-2}$. Therefore, from (24) and (26), we obtain

$${}_2^{\#}C_2^{(p)}(\mathbf{D}^*) = \begin{cases} 0, & \text{for } 0 \le p < n - \dfrac{N}{4} - 1, \\ \dfrac{N}{4}(n - N/4), & \text{for } p = n - \dfrac{N}{4} - 1, \\ 0, & \text{for } n - \dfrac{N}{4} - 1 < p < \dfrac{N}{2^w} - 1, \\ \dfrac{N}{2^w}\left(N/4 - N/2^w\right), & \text{for } p = \dfrac{N}{2^w} - 1. \end{cases} \tag{27}$$

Comparing (27) with (19), (20) for $4 \le v \le w$, and (26) for $4 \le v \le w-1$, it is obtained that if $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ is a B^2-GMC design, $\mathbf{D}$ should not be worse than $\mathbf{D}^*$ in terms of (3). Therefore, if $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ is a B^2-GMC design, $\mathbf{D}_t$ should be an n-projection of $\mathbf{S}_{N(2^{w-2}+1)/2^w}$. This completes the proof. □

The lemma below is an extension of Lemma A.2, in Appendix, introduced from [34].

Lemma 4. *Suppose* $\mathbf{D}_t$ *consists of the last* n, $N(2^{w-1}+1)/2^{w+1} < n \le N(2^{w-2}+1)/2^w$, *columns of* $\mathbf{S}_{N(2^{w-2}+1)/2^w}$; *then,* $B_2(\mathbf{D}_t, \boldsymbol{\gamma}_1) \ge B_2(\mathbf{D}_t, \boldsymbol{\gamma}_2)$ *if* $\boldsymbol{\gamma}_1$ *is ahead of* $\boldsymbol{\gamma}_2$ *in* $\mathbf{H}_{q-w}$.

Proof. Suppose $\mathbf{O}$ consists of the last $n - N/4$ columns of $D^{q-w}((\mathbf{w}-\mathbf{1})_{2^w})$. Then, $\mathbf{D}_t = \{\mathbf{O}, \mathbf{S}_{N(2^{w-2}+1)/2^w} \backslash D^{q-w}(\mathbf{w}-\mathbf{1})\}$. From (16) and (18), for any $\boldsymbol{\gamma} \in \mathbf{H}_{q-w}$, $B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = B_2(\mathbf{O}, \boldsymbol{\gamma}) + N/8$. Straightforwardly, from Lemma A.2, it is obtained that $B_2(\mathbf{O}, \boldsymbol{\gamma}_1) \ge B_2(\mathbf{O}, \boldsymbol{\gamma}_2)$ if $\boldsymbol{\gamma}_1$ is ahead of $\boldsymbol{\gamma}_2$ in $\mathbf{H}_{q-w}$. This completes the proof. □

With Lemmas 3 and 4, Theorem 2 provides the construction methods of B^2-GMC $2^{n-m}: 2^s$ designs with $N/4 + 1 < n \le 5N/16$.

Theorem 2. *Let* $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ *be a* $2^{n-m}: 2^s$ *design with* $N(2^{w-1}+1)/2^{w+1} < n \le N(2^{w-2}+1)/2^w$ *for* $w \ge 4$ *and* $2^k \le s \le 2^{k+1} - 1$ *for some* $0 \le k \le q-3$. *Suppose* $2^r \le N(2^{w-2}+1)/2^w - n \le 2^{r+1} - 1$ *for some* $0 \le r \le q-w-1$; *then,* $\mathbf{D}$ *is a* B^2*-GMC design if* $\mathbf{D}_t$ *consists of the last n columns of* $\mathbf{S}_{N(2^{w-2}+1)/2^w}$ *and*

(a) $\mathbf{D}_b$ *is any s-projection of* $\{\mathbf{H}_k, \mathbf{q}-\mathbf{1}, (\mathbf{q}-\mathbf{1})\mathbf{H}_k\}$ *when* $k \le r \le q-w-1$.

(b) $\mathbf{D}_b$ *is any s-projection of* $\mathbf{H}_{k+1}$ *when* $r+1 \le k \le q-w-1$.

(c) $\mathbf{D}_b$ *is any s-projection of* $\mathbf{H}_{q-2}$ *when* $q-w-1 < k \le q-3$.

Proof. By Lemma 3 and its proof, if $\mathbf{D}^* = (\mathbf{D}_t^*: \mathbf{D}_b^*)$ is a B^2-GMC $2^{n-m}: 2^s$ design, there should be $\mathbf{D}_t^* \subset \mathbf{S}_{N(2^{w-2}+1)2^w}$ and $\overline{\mathbf{D}}_t^* = \mathbf{S}_{N(2^{w-2}+1)2^w} \backslash \mathbf{D}_t^* \subset D^{q-w}((\mathbf{w}-\mathbf{1})_{2^w})$. Substituting v with w in (24), we obtain

$$B_2(\mathbf{D}_t^*, \boldsymbol{\gamma}) = \begin{cases} 0, & \text{for } \boldsymbol{\gamma} \in \mathbf{S}_{(N(2^{w-2}+1)/2^w)}, \\ n - \dfrac{N}{4}, & \text{for } \boldsymbol{\gamma} \in \left\{\mathbf{q}\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}\right), \mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-w}\right)\right\}, \\ \dfrac{N}{2^w}, & \text{for } \boldsymbol{\gamma} \in (\mathbf{q}-\mathbf{1})\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}\right), \\ \dfrac{N}{8} - \dfrac{N}{2^w}, & \text{for } \boldsymbol{\gamma} \in \mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}, \\ B_2\left(\overline{D}_t^*, \boldsymbol{\gamma}\right) + \dfrac{N}{8}, & \text{for } \boldsymbol{\gamma} \in \mathbf{H}_{q-w}. \end{cases} \tag{28}$$

Let

$$\begin{aligned} u_1 &= \#\left\{U(\mathbf{D}_b^*) \cap \left\{\mathbf{q}\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}\right), \mathbf{q}(\mathbf{q}-\mathbf{1})\left(\mathbf{I}, \mathbf{H}_{q-w}\right)\right\}\right\}, \\ u_2 &= \#\left\{U(\mathbf{D}_b^*) \cap (\mathbf{q}-\mathbf{1})\left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}\right)\right\}, \\ u_3 &= \#\left\{U(\mathbf{D}_b^*) \cap \left(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}\right)\right\}, \end{aligned} \tag{29}$$

and then

$$
{}_2^{\#}C_2^{(p)}(\mathbf{D}^*) = \begin{cases} \left(n-\dfrac{N}{4}\right)\left(\dfrac{N}{4}-u_1\right), & \text{for } p = n-\dfrac{N}{4}-1, \\ \dfrac{N}{2^w}\left(\dfrac{N}{4}-\dfrac{N}{2^w}-u_2\right), & \text{for } p = \dfrac{N}{2^w}-1, \\ \left(\dfrac{N}{8}-\dfrac{N}{2^w}\right)\left(\dfrac{N}{4}-\dfrac{N}{2^w}-u_3\right), & \text{for } p = \dfrac{N}{8}-\dfrac{N}{2^w}-1, \\ f(\mathbf{D}_t^*, p), & \text{for } p \geq \dfrac{N}{8}-1, \\ 0, & \text{otherwise,} \end{cases} \tag{30}
$$

where $f(\mathbf{D}_t^*, p) = {}_2^{\#}C_2^{(p)}(\mathbf{D}_t^*) - (p+1)\#\{\boldsymbol{\gamma}: \boldsymbol{\gamma} \in \mathbf{H}_{q-w} \cap U(\mathbf{D}_b^*), B_2(\mathbf{D}_t^*, \boldsymbol{\gamma}) = p+1\}$. Therefore, $\mathbf{D}^*$ sequentially maximizes ${}_2^{\#}C_2$ among all the possible 2^{n-m}: 2^s designs, only if $\mathbf{D}^*$ sequentially minimizes

$$
(u_1, u_2, u_3). \tag{31}
$$

For ease of presenting, let $\mathbf{A} = \{\mathbf{q}(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}), \mathbf{q}(\mathbf{q}-\mathbf{1})(\mathbf{I}, \mathbf{H}_{q-w})\}$, $\mathbf{B} = (\mathbf{q}-\mathbf{1})(\mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w})$, and $\mathbf{C} = \mathbf{H}_{q-2} \backslash \mathbf{H}_{q-w}$.

For (a), when $k \leq q-w-1$, there are two different ways to choose $\mathbf{D}_b^*$ from $\mathbf{H}_q$: (i) $\mathbf{D}_b^* \cap \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} = \varnothing$ and (ii) $\mathbf{D}_b^* \cap \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \neq \varnothing$. Recall that if $\mathbf{D}^* = (\mathbf{D}_t^*: \mathbf{D}_b^*)$ is a B^2-GMC design, then $\mathbf{D}_t^* \subset \mathbf{S}_{N(2^{w-2}+1)2^w}$ with $\overline{D}_t^* = \mathbf{S}_{N(2^{w-2}+1)2^w} \backslash \mathbf{D}_t^* \subset D^{q-w}((\mathbf{w}-\mathbf{1})_{2^w})$. Case (i) implies that $\mathbf{D}_b^* \subset \{\mathbf{H}_{q-w}, \overline{D}_t^*\}$, and thus $\mathbf{D}_b^* \subset \{\mathbf{H}_{q-w}, \mathbf{q}-\mathbf{1}, (\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w}\}$. Therefore, choosing $\mathbf{D}_b^*$ according to case (i) results in $u_1 = u_2 = u_3 = 0$. If $\mathbf{D}_b^*$ is chosen according to case (ii), then $U(\mathbf{D}_b^*) \cap \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \neq \varnothing$ which results in $u_1 > 0$, $u_2 > 0$, or $u_3 > 0$. Clearly, if $\mathbf{D}^*$ is a B^2-GMC design, $\mathbf{D}_b^*$ should be chosen as (i). In the following, we consider only $\mathbf{D}_b^* \cap \{\mathbf{A}, \mathbf{B}, \mathbf{C}\} = \varnothing$, i.e., $\mathbf{D}_b^* \subset \{\mathbf{H}_{q-w}, (\mathbf{q}-\mathbf{1}), (\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w}\}$.

For $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ in (a), according to (v) in Lemma 1, if $\mathbf{D}_b \subset \{\mathbf{H}_k, (\mathbf{q}-\mathbf{1}), (\mathbf{q}-\mathbf{1})\mathbf{H}_k\}$ with $k \leq q-w-1$, then $U(\mathbf{D}_b) = \{\mathbf{H}_k, (\mathbf{q}-\mathbf{1}), (\mathbf{q}-\mathbf{1})\mathbf{H}_k\}$. Denote $\mathbf{b}$ as the last column of $\mathbf{H}_k$ in Yates order, and $B_2(\mathbf{D}_t, \mathbf{b}) = b$; then, $b \geq N/8$. According to Lemma 4, for $\boldsymbol{\gamma} \in \mathbf{H}_{q-w}$, if $\boldsymbol{\gamma}$ is ahead of $\mathbf{b}$ in $\mathbf{H}_{q-w}$, then $B_2(\mathbf{D}_t, \boldsymbol{\gamma}) \geq b$ by (18). For any $\boldsymbol{\gamma} \in \{(\mathbf{q}-\mathbf{1}), (\mathbf{q}-\mathbf{1})\mathbf{H}_k\}$, we have $B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = 0$. For any $\boldsymbol{\gamma}$ such that $B_2(\mathbf{D}_t, \boldsymbol{\gamma}) < b$, we have $\boldsymbol{\gamma} \in \mathbf{H}_q \backslash \mathbf{H}_k$. Therefore, for $p \leq b-2$,

$$
\begin{aligned} {}_2^{\#}C_2^{(p)}(\mathbf{D}) &= (p+1)\#\{\boldsymbol{\gamma} \in \mathbf{H}_q: \boldsymbol{\gamma} \notin U(\mathbf{D}_b), B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = p+1\} \\ &= (p+1)\#\{\boldsymbol{\gamma} \in \mathbf{H}_q: B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = p+1\} \\ &= {}_2^{\#}C_2^{(p)}(\mathbf{D}_t), \end{aligned} \tag{32}
$$

and for $p \geq b$,

$$
\begin{aligned} {}_2^{\#}C_2^{(p)}(\mathbf{D}) &= (p+1)\#\{\boldsymbol{\gamma} \in \mathbf{H}_q: \boldsymbol{\gamma} \notin U(\mathbf{D}_b), B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = p+1\} \\ &= (p+1)\#\{\boldsymbol{\gamma} \in \mathbf{H}_q: B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = p+1\} - (p+1)\#\{\boldsymbol{\gamma} \in \mathbf{H}_k: B_2(\mathbf{D}_t, \boldsymbol{\gamma}) = p+1\} \\ &= 0. \end{aligned} \tag{33}
$$

From (33), it is obtained that

$$
\begin{aligned} &\left({}_2^{\#}C_2^{(0)}(\mathbf{D}), \ldots, {}_2^{\#}C_2^{(N/8-1)}(\mathbf{D}), {}_2^{\#}C_2^{(N/8)}(\mathbf{D}), \ldots, {}_2^{\#}C_2^{(b-2)}(\mathbf{D})\right) \\ &= \left({}_2^{\#}C_2^{(0)}(\mathbf{D}_t), \ldots, {}_2^{\#}C_2^{(N/8-1)}(\mathbf{D}_t), {}_2^{\#}C_2^{(N/8)}(\mathbf{D}_t), \ldots, {}_2^{\#}C_2^{(b-2)}(\mathbf{D}_t)\right). \end{aligned} \tag{34}
$$

Note that $\mathbf{D}_t$ is a GMC design [29]; then, $\mathbf{D}$ maximizes

$$
\left({}_2^{\#}C_2^{(0)}(\cdot), \ldots, {}_2^{\#}C_2^{(N/8-1)}(\cdot), {}_2^{\#}C_2^{(N/8)}(\cdot), \ldots, {}_2^{\#}C_2^{(b-2)}(\cdot)\right), \tag{35}
$$

among all the possible 2^{n-m}: 2^s designs.

Suppose $\mathbf{D}$ is not a B^2-GMC design; then, there should be a $\mathbf{D}^* = (\mathbf{D}_t^*: \mathbf{D}_b^*)$ which outperforms $\mathbf{D} = (\mathbf{D}_t: \mathbf{D}_b)$ in terms of (3). This implies that there exists some $p_1 \geq b-1$ such that

$$\begin{aligned}&\left({}_2^{\#}C_2^{(0)}(\mathbf{D}^*),\ldots,{}_2^{\#}C_2^{(N/8-1)}(\mathbf{D}^*),{}_2^{\#}C_2^{(N/8)}(\mathbf{D}^*),\ldots,{}_2^{\#}C_2^{(b-2)}(\mathbf{D}^*)\right)\\&=\left({}_2^{\#}C_2^{(0)}(\mathbf{D}),\ldots,{}_2^{\#}C_2^{(N/8-1)}(\mathbf{D}),{}_2^{\#}C_2^{(N/8)}(\mathbf{D}),\ldots,{}_2^{\#}C_2^{(b-2)}(\mathbf{D})\right),\end{aligned}\tag{36}$$

$${}_2^{\#}C_2^{(p_1-1)}(\mathbf{D}^*)>{}_2^{\#}C_2^{(p_1-1)}(\mathbf{D}).\tag{37}$$

Recalling the definitions of ${}_2^{\#}C_2^{(p)}(\mathbf{D})$ and $B_2(\mathbf{D}_t,\boldsymbol{\gamma})$, we have

$$\begin{aligned}\sum_{p=0}^{P}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D})}{(p+1)}&=\sum_{p=0}^{P}\#\left\{\boldsymbol{\gamma}\in\mathbf{H}_q\colon\boldsymbol{\gamma}\notin U(\mathbf{D}_b),B_2(\mathbf{D}_t,\boldsymbol{\gamma})=p+1\right\}\\&=\sum_{p=0}^{P}\#\left\{\boldsymbol{\gamma}\in\left(\mathbf{H}_q\backslash\mathbf{S}_{N(2^{w-2}+1)/2^w}\right)\colon B_2(\mathbf{D}_t,\boldsymbol{\gamma})=p+1\right\}\\&\quad-\sum_{p=0}^{P}\#\left\{\boldsymbol{\gamma}\in U(\mathbf{D}_b)\colon B_2(\mathbf{D}_t,\boldsymbol{\gamma})=p+1\right\}\\&=\#\left(\mathbf{H}_q\backslash\mathbf{S}_{N(2^{w-2}+1)/2^w}\right)-\#\mathbf{H}_k\\&=2^q-\frac{N\left(2^{w-2}+1\right)}{2^w}-2^k,\end{aligned}\tag{38}$$

where the second and third equalities are due to $B_2(\mathbf{D}_t,\boldsymbol{\gamma})=0$ for any $\boldsymbol{\gamma}\in\mathbf{S}_{N(2^{w-2}+1)/2^w}$. Similarly,

$$\begin{aligned}\sum_{p=0}^{P}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D}^*)}{(p+1)}&=\sum_{p=0}^{P}\#\left\{\boldsymbol{\gamma}\in\mathbf{H}_q\colon\boldsymbol{\gamma}\notin U(\mathbf{D}_b^*),B_2(\mathbf{D}_t^*,\boldsymbol{\gamma})=p+1\right\}\\&=\sum_{p=0}^{P}\#\left\{\boldsymbol{\gamma}\in\left(\mathbf{H}_q\backslash\mathbf{S}_{N(2^{w-2}+1)/2^w}\right)\colon B_2(\mathbf{D}_t^*,\boldsymbol{\gamma})=p+1\right\}\\&\quad-\sum_{p=0}^{P}\#\{\boldsymbol{\gamma}\in U(\mathbf{D}_b^*)\colon B_2(\mathbf{D}_t^*,\boldsymbol{\gamma})=p+1\},\\&=\#\left(\mathbf{H}_q\backslash\mathbf{S}_{N(2^{w-2}+1)/2^w}\right)-\#U(\mathbf{D}_b^*),\\&=2^q-\frac{N\left(2^{w-2}+1\right)}{2^w}-\#U(\mathbf{D}_b^*).\end{aligned}\tag{39}$$

From (33) and (38), we have

$$\begin{aligned}\sum_{p=0}^{P}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D})}{(p+1)}&=\sum_{p=0}^{b-2}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D})}{(p+1)}+\frac{{}_2^{\#}C_2^{(b-1)}(\mathbf{D})}{b}\\&=2^q-\frac{N\left(2^{w-2}+1\right)}{2^w}-2^k.\end{aligned}\tag{40}$$

By (36) and (37), it is obtained that

$$\sum_{p=0}^{P}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D}^*)}{(p+1)}\geq\sum_{p=0}^{b-2}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D}^*)}{(p+1)}+\frac{{}_2^{\#}C_2^{(b-1)}(\mathbf{D}^*)}{b}$$
$$>\sum_{p=0}^{b-2}\frac{{}_2^{\#}C_2^{(p)}(\mathbf{D})}{(p+1)}+\frac{{}_2^{\#}C_2^{(b-1)}(\mathbf{D})}{b}. \tag{41}$$

Then, according to (39)–(41), we obtain $\#U(\mathbf{D}_b^*)<2^k-1$ which contradicts Lemma 1 (i) and (ii).

For (b), similar to (a), when $k\leq q-w-1$, there are two different ways to choose $\mathbf{D}_b^*$ from $\mathbf{H}_q$: (i) $\mathbf{D}_b^*\cap\{\mathbf{A},\mathbf{B},\mathbf{C}\}=\varnothing$ and (ii) $\mathbf{D}_b^*\cap\{\mathbf{A},\mathbf{B},\mathbf{C}\}\neq\varnothing$. With a similar argument to (a), case (i) results in $u_1=u_2=u_3=0$ while case (ii) results in $u_1>0, u_2>0$ or $u_3>0$. Therefore, if $\mathbf{D}^*$ is a B^2-GMC design, $\mathbf{D}_b^*$ should be chosen according to case (i), i.e., $\mathbf{D}_b^*\subset\{\mathbf{H}_{q-w},(\mathbf{q}-\mathbf{1}),(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w}\}$. When choosing $\mathbf{D}_b^*$ according to (i), there should be $\mathbf{D}_b^*\cap\{\mathbf{q}-\mathbf{1},(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w}\}=\varnothing$, i.e., $\mathbf{D}_b^*\subset\mathbf{H}_{q-w}$; otherwise, $\mathbf{D}_t^*\cap U(\mathbf{D}_b^*)\neq\varnothing$ noting that $\#\{U(\mathbf{D}_b^*)\cap(\mathbf{q}-\mathbf{1},(\mathbf{q}-\mathbf{1})\mathbf{H}_{q-w})\}\geq 2^k>2^{r+1}-1$ as shown in Lemma 1 (iii). The remainder of the proof is similar to that of (a).

For (c), when $k>q-w-1$, there are two different ways to choose $\mathbf{D}_b^*$ from $\mathbf{H}_q$: (i) $\mathbf{D}_b^*\cap\{\mathbf{B},\mathbf{C}\}=\varnothing$ and (ii) $\mathbf{D}_b^*\cap\{\mathbf{B},\mathbf{C}\}\neq\varnothing$. With a similar argument to (a), it is not hard to verify that case (ii) results in $u_1>0$ or $u_2>0$ while case (i) results in $u_1=u_2=0$. Therefore, if $\mathbf{D}^*$ is a B^2-GMC design, $\mathbf{D}_b^*$ should be an s-projection of $\mathbf{H}_{q-2}=\{\mathbf{B},\mathbf{C}\}$. The remainder of the proof is similar to that of (a). □

Remark 1. Theorem 2 shows that when constructing the B^2-GMC designs with $N/4+1<n\leq 5N/16$, we can first partition the range $(N/4+1,5N/16)$ into $q-4$ sequential subranges as $(N(2^{w-1}+1)/2^{w+1},N(2^{w-2}+1)/2^w)$ with $w=4,5,\ldots,q-1$ and then obtain the B^2-GMC designs according to Theorem 2.

In the following, an example is provided to illustrate the construction method in Theorem 2 and Remark 1.

Example 2. Consider the constructions of B^2-GMC 2^{19-13}: 2^s designs for $s=1,3,4$. Here, $n=19, m=13, q=6$, and $N=64$. The values of the parameters N and n satisfy $N/4+1<n\leq 5N/16$, i.e., $n\in(17,20]$. From Remark 1, we partition the range $(17,20]$ into two subranges $(17,18]$ because $w=5$ and $(18,20]$ because $w=4$. Since $19\in(18,20]$, according to Theorem 2, $\mathbf{D}_t$ should be the last 19 columns of $\mathbf{S}_{5N/16}$, where $\mathbf{S}_{5N/16}=D^{q-4}(\Phi(4))=(\mathbf{5},\mathbf{5H}_2,\mathbf{6},\mathbf{6H}_2,\mathbf{56}(\mathbf{H}_4\backslash\mathbf{H}_2))$ from (18). Therefore, $\mathbf{D}_t=(\mathbf{5H}_2,\mathbf{6},\mathbf{6H}_2,\mathbf{56}(\mathbf{H}_4\backslash\mathbf{H}_2))$. Note that $5N/16-n=20-19=1$; then, $r=0$. Next, we choose $\mathbf{D}_b$ for $s=1,3,4$.

For $s=1$, we have $k=0$. According to Theorem 2 (a), $\mathbf{D}_b=\mathbf{5}$ as $\mathbf{H}_0=\varnothing$.

For $s=3$, we have $k=1$. According to Theorem 2 (b), $\mathbf{D}_b$ should be a 3-projection of $\mathbf{H}_2$, i.e., $\mathbf{D}_b=\mathbf{H}_2=(\mathbf{1},\mathbf{2},\mathbf{12})$.

For $s=4$, we have $k=2$. According to Theorem 2 (c), $\mathbf{D}_b$ should be a 4-projection of $\mathbf{H}_4$. Without loss of generality, we choose $\mathbf{D}_b=(\mathbf{1},\mathbf{2},\mathbf{12},\mathbf{3})$.

4. Concluding Remarks

The regular 2^{n-m} designs have wide applications in engineering, manufacturing industry, agriculture and medicine, and so on. When the size of experimental units is large, the inhomogeneity of experimental units results in unwanted variances of estimations of treatment effects. An essential way to solve this problem is to partition the experimental units into blocks.

There are two kinds of blocking problems as pointed out in [1]. One is called the single block variable problem which considers only a single block variable, and the other is called multi-block variable problem which considers two or more block variables. As stated in Section 1, experiments which involve multi-block variables are more widely concerned in practice than those which involve only a single block variable. However, due to the complexity of multi-block variable problem, the studies on constructing optimal designs with multi-block variables are relatively rare.

In this paper, we aim at exploring the theories and constructions of optimal blocked 2^{n-m} designs with multi-block variables. The prevalent B^2-GMC criterion is adopted. This criterion is preferable when there is some prior knowledge on the importance ordering of the treatment effects. The systematical construction methods of the B^2-GMC 2^{n-m}: 2^s designs with $N/4+1\leq n\leq 5N/16$ are developed. The construction methods are concise and easy to implement as indicated by the examples provided.

Appendix

Lemma A.1 is some result from Lemmas A.2, A.4, and A.5 in [21] and Lemma 1 in [25].

Lemma A.1. *Suppose* $\mathbf{O}$ *is any s-projection of* $\mathbf{H}_q$ *with* $2^k\leq s\leq 2^{k+1}-1$ *and* $k\leq q-2$.

(i) If $\mathbf{O}\cap\{\mathbf{q},\mathbf{qH}_{q-1}\}=\varnothing$, *then* $\#\{U(\mathbf{O})\cap\mathbf{H}_{q-1}\}\geq 2^{k+1}-1$, *and when* $\mathbf{O}$ *has* $k+1$ *independent columns, the equality holds.*

(ii) If $\mathbf{O}\cap\{\mathbf{q},\mathbf{qH}_{q-1}\}\neq\varnothing$, *then* $\#\{U(\mathbf{O})\cap\mathbf{H}_{q-1}\}\geq 2^k-1$, *and when* $\mathbf{O}$ *has* $k+1$ *independent columns, the equality holds.*

(iii) If $\mathbf{O}\cap\{\mathbf{q},\mathbf{qH}_{q-1}\}\neq\varnothing$, *then* $\#\{U(\mathbf{O})\cap\{\mathbf{q},\mathbf{qH}_{q-1}\}\}\geq 2^k$, *and when* $\mathbf{O}$ *has* $k+1$ *independent columns, the equality holds.*

(iv) If $\mathbf{O}\subset\mathbf{H}_{k+1}$, *then* $U(\mathbf{O})=\mathbf{H}_{k+1}$.

(v) If $\mathbf{O}\subset\{\mathbf{H}_k,\mathbf{q},\mathbf{qH}_k\}$, *then* $U(\mathbf{O})=\{\mathbf{H}_k,\mathbf{q},\mathbf{qH}_k\}$.

Lemma A.2 is some result of Lemma 1 in [34].

Lemma A.2. *Suppose* $\mathbf{O}$ *consists of the last n columns of* $\{\mathbf{q},\mathbf{qH}_{q-1}\}$; *then,* $B_2(\mathbf{O},\boldsymbol{\gamma}_1)\geq B_2(\mathbf{O},\boldsymbol{\gamma}_2)$ *if* $\boldsymbol{\gamma}_1$ *is ahead of* $\boldsymbol{\gamma}_2$ *in* $\mathbf{H}_{q-1}$.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant no. 11801331).

References

[1] S. Bisgaard, "A note on the definition of resolution for blocked 2 k-p designs," *Technometrics*, vol. 36, no. 3, pp. 308–311, 1994.

[2] R. R. Sitter, J. Chen, and M. Feder, "Fractional resolution and minimum aberration in blocked 2 n-k designs," *Technometrics*, vol. 39, no. 4, pp. 382–390, 1997.

[3] H. Chen and C. S. Cheng, "Theory of optimal blocking of s^{n-m}designs," *Annals of Statistics*, vol. 27, no. 6, pp. 1948–1973, 1999.

[4] R. Zhang and D. Park, "Optimal blocking of two-level fractional factorial designs," *Journal of Statistical Planning and Inference*, vol. 91, no. 1, pp. 107–121, 2000.

[5] S. W. Cheng and C. F. J. Wu, "Choice of optimal blocking schemes in two-level and three-level designs," *Technometrics*, vol. 44, no. 3, pp. 269–277, 2002.

[6] H. Xu, "Blocked regular fractional factorial designs with minimum aberration," *Annals of Statistics*, vol. 8, no. 5, pp. 2534–2553, 2006.

[7] H. Xu and R. W. Mee, "Minimum aberration blocking schemes for 128-run designs," *Journal of Statistical Planning and Inference*, vol. 140, no. 11, pp. 3213–3229, 2010.

[8] S. Zhao, P. Li, and R. Karunamuni, "Blocked two-level regular factorial designs with weak minimum aberration," *Biometrika*, vol. 100, no. 1, pp. 249–253, 2013.

[9] S.-L. Zhao and P.-F. Li, "Construction of minimum aberration blocked two-level regular factorial designs," *Communications in Statistics - Theory and Methods*, vol. 45, no. 17, pp. 5028–5036, 2016.

[10] B.-J. Chen, P.-F. Li, M.-Q. Liu, and R.-C. Zhang, "Some results on blocked regular 2-level fractional factorial designs with clear effects," *Journal of Statistical Planning and Inference*, vol. 136, no. 12, pp. 4436–4449, 2006.

[11] S. Zhao, P. Li, and M.-Q. Liu, "On blocked resolution IV designs containing clear two-factor interactions," *Journal of Complexity*, vol. 29, no. 5, pp. 389–395, 2013.

[12] R. C. Zhang and R. Mukerjee, "General minimum lower-order confounding in block designs using complementary sets," *Statistica Sinica*, vol. 19, no. 4, pp. 1787–1802, 2009.

[13] S. Zhao, D. K. J. Lin, and P. Li, "A note on the construction of blocked two-level designs with general minimum lower order confounding," *Journal of Statistical Planning and Inference*, vol. 172, pp. 16–22, 2016.

[14] S.-L. Zhao and Q.-Q. Zhao, "Some results on constructing two-level block designs with general minimum lower order confounding," *Communications in Statistics-Theory and Methods*, vol. 47, no. 9, pp. 2227–2237, 2018.

[15] R. C. Zhang, P. Li, S. L. Zhao, and M. Y. Ai, "A general minimum lower-order confounding criterion for two-level regular designs," *Statistica Sinica*, vol. 18, no. 4, pp. 1689–1705, 2008.

[16] S. Zhao, P. Li, R. Zhang, and R. Karunamuni, "Construction of blocked two-level regular designs with general minimum lower order confounding," *Journal of Statistical Planning and Inference*, vol. 143, no. 6, pp. 1082–1090, 2013.

[17] Y. Zhao, S. Zhao, and M.-Q. Liu, "A theory on constructing blocked two-level designs with general minimum lower order confounding," *Frontiers of Mathematics in China*, vol. 11, no. 1, pp. 207–235, 2016.

[18] S.-L. Zhao and Q. Sun, "On constructing general minimum lower order confounding two-level block designs," *Communications in Statistics-Theory and Methods*, vol. 46, no. 3, pp. 1261–1274, 2017.

[19] J. Wei, P. Li, and R. Zhang, "Blocked two-level regular designs with general minimum lower order confounding," *Journal of Statistical Theory and Practice*, vol. 8, no. 1, pp. 46–65, 2014.

[20] C. F. J. Wu and M. S. Hamada, *Experiments: Planning, Analysis and Optimization*, Wiley, Hoboken, NJ, USA, 2 edition, 2009.

[21] Q. Zhao and S. Zhao, "Some results on two-level regular designs with multi block variables containing clear effects," *Statistical Papers*, vol. 60, no. 5, pp. 1569–1582, 2019.

[22] S. Zhao and Q. Zhao, "Minimum aberration blocked designs with multiple block variables," *Metrika*, vol. 84, no. 2, pp. 121–140, 2021.

[23] R. Zhang, P. Li, and J. Wei, "Optimal two-level regular designs with multi block variables," *Journal of Statistical Theory and Practice*, vol. 5, no. 1, pp. 161–178, 2011.

[24] Y. N. Zhao, S. L. Zhao, and M. Q. Liu, "On constructing optimal two-level designs with multi block variables," *Journal of Systems Science and Complexity*, vol. 31, pp. 1–14, 2018.

[25] Y. N. Zhao and S. L. Zhao, "Construction of optimal blocked designs with multi block variables," *AIMS mathematics*, vol. 6, no. 6, pp. 6293–6308, 2021.

[26] B. Tang and C. F. J. Wu, "Characterization of minimum aberration 2^{n-k}designs in terms of their complementary designs," *Annals of Statistics*, vol. 25, no. 6, pp. 2549–2559, 1996.

[27] C. F. J. Wu and Y. Chen, "A graph-aided method for planning two-level experiments when certain interactions are important," *Technometrics*, vol. 34, no. 2, pp. 162–175, 1992.

[28] G. E. P. Box and J. S. Hunter, "The 2^{k-p}fractional factorial designs I and II," *Technometrics*, vol. 3, no. 4, pp. 311–351, 1961.

[29] Y. Cheng and R. Zhang, "On construction of general minimum lower order confounding 2n−m designs with ,$(N/4)+1 \le n \le 9N/32$," *Journal of Statistical Planning and Inference*, vol. 140, no. 9, pp. 2384–2394, 2010.

[30] H. Wu and C. F. J. Wu, "Clear two-factor interactions and minimum aberration," *Annals of Statistics*, vol. 30, no. 5, pp. 1496–1511, 2002.

[31] R. M. Block and R. W. Mee, "Second order saturated resolution IV designs," *Journal of Statistical Theory and Applications*, vol. 2, no. 2, pp. 96–112, 2003.

[32] A. A. Davydov and L. M. Tombak, "Quasi-perfect linear binary codes with distance 4 and complete caps in projective geometry," *Problems of Information Transmission*, vol. 25, no. 4, pp. 265–275, 1990.

[33] R. Zhang and Y. Cheng, "General minimum lower order confounding designs: an overview and a construction theory," *Journal of Statistical Planning and Inference*, vol. 140, no. 7, pp. 1719–1730, 2010.

[34] Q. Zhou, N. Balakrishnan, and R. Zhang, "The factor aliased effect number pattern and its application in experimental planning," *Canadian Journal of Statistics*, vol. 41, no. 3, pp. 540–555, 2013.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6650237, 13 pages
https://doi.org/10.1155/2021/6650237

Research Article

Weibull-Exponential Distribution and Its Application in Monitoring Industrial Process

Muhammad Bilal,[1,2] Muhammad Mohsin,[2] and Muhammad Aslam [3]

[1]*Department of Statistics and Computer Science, University of Veterinary and Animal Sciences, Lahore, Pakistan*
[2]*Department of Statistics, COMSATS University Islamabad, Lahore Campus, Lahore 54000, Pakistan*
[3]*Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia*

Correspondence should be addressed to Muhammad Aslam; aslam_ravian@hotmail.com

Received 4 December 2020; Revised 12 February 2021; Accepted 6 March 2021; Published 26 March 2021

Academic Editor: Hussein Abulkasim

This paper presents a new Weibull family of distributions. The compatibility of the newly developed class is justified through its application in the field of quality control using Weibull-exponential distribution, a special case of the proposed family. In this paper, an attribute control chart using Weibull-exponential distribution is developed. The estimations of the model parameters and the proposed chart parameters are performed through the methods of maximum likelihood and average run-length. The significance of the proposed model is demonstrated using a simulation study and real-life problems. The results of the monitoring process and quick detection are compared with exponential distribution.

1. Introduction

The Weibull (W) distribution is extensively used to model failure time data. Numerous distributions have been developed and applied to describe various phenomena in engineering and quality control. Several generalizations of the family of distribution have been studied in the last two decades. Eugene et al. [1] proposed a new beta generated family of distributions for skewed and bimodal data. Many practitioners discussed a technique to generate different distributions Nadarajah and Kotz [2], Nadarajah and Gupta [3], Nadarajah and Kotz [4], and Nadarajah et al. [5]. Alzaatreh et al. [6] generalized the technique of Eugene et al. [1] to develop distributions defined over any domain. A lot of work is available in the recent literature regarding Alzaatreh et al. [6], Cordeiro et al. [7], Afify et al. [8], Alizadeh et al. [9], Afify et al. [8, 10], and Nofal et al. [11]. Chahkandi and Ganjali [12] established an exponential power series (EPS) family of distributions, which included a new mixture of an exponential and binomial distribution (see Barreto-Souza et al. [13]). Lu and Shi [14] introduced the W-geometric (WG) and W-Poisson (WP) distributions were expanded as exponential geometric and exponential Poisson distributions, respectively. A comprehensive account of statistical methods can be traced in quality control since its commencement. In 1924, Walter A. Shewhart developed the concept of statistical quality control [15]. In early days, the control charts were extensively applied to monitor production processes only, but nowadays these control charts are applied in various fields like health care [16], education [17], coal monitoring [18], nuclear engineering [17], and veterinary medicine [19, 20]. Roberts [21] depicted a graphical technique to generate geometric moving average by selecting weight for most recent values. Brook and Evans [22] suggested the technique of adopting a Markov chain approach. By using this method, they attained a transition probability matrix along with moments, percentage points of run-length distribution, and exact probability of run-length. Lucas [23] developed a scheme by combining the Shewhart control chart with the cumulative sum control chart (CUSM), which detected the small shifts rapidly. Borror et al. [24] suggested the Markov chain approach to assess the average run-length (ARL). Khoo [25] designed a new approach to monitor the fraction of

nonconforming products. They constructed a moving average (MA) control chart for a fraction of nonconforming using the ARL profile of simulation results and mathematical calculations.

Sukparungsee and Novikov [26] devolved analytical approximation of exponentially weighted moving average (EWMA) using a martingale procedure for the expectation of exit times. Areepong and Novikov [27] derived ARL and average delay time (AD) utilizing the martingale approach. Noorossana et al. [28] worked on different methods to monitor rare health events utilizing the ZIB-EWMA chart. Santiago and Smith [29] put forward the idea of the probabilistic-based procedure to create a t-chart to observe the stability of the process. They yielded exponential distribution and defined supplementary rules like ARL and shift detection. Ho and Quinino [30] used the "np" control chart to monitor the process variability instead of the mean chart. They explored that the "np" control chart was better in performance than R (range chart) and S^2 control chart. Aslam et al. [31] implemented an ordinary sampling plan and group sampling plan. They developed sampling plans for the multistage process, which depended upon time-truncated tests and gave nonlinear optimized solutions to decrease the average sample number. Rakitzis and Castagliola [32] studied the Shewhart control chart for zero-inflated probability models. They suggested pragmatic rules for the statistical design of inspected charts for a fixed sample size. Azam et al. [33] designed a control chart under a repetitive sampling technique using an exponential distribution. They analyzed two existing control charts with the new proposed chart using average run-length. Azam et al. [34] unfolded an attribute chart for the W distribution using the discriminant limits. They established a control chart to estimate parameters in light of ARL. Wu and Jiao [35] performed attribute inspection to track the mean value of variable characteristic for normal distribution. This chart produced an out-of-control signal, if the interval between the two suspect samples is lower than a control limit. Kim and Lee [36] used a fixed sampling interval scheme (FSI) for the detection of small shifts to moderate shifts by increasing fractional parameters for zero-inflated binomial process. The effectiveness of the FSI scheme over the VSI (variable sampling interval) scheme was justified through a numerical study. For further achievements to the statistical process monitoring literature on the Shewhart-CUSUM and Shewhart-EWMA, for instance, Aslam et al. [37] improved the GWMA (generally weighted moving average) monitoring scheme for the detection of large shift in the process. For this purpose, they proposed the phase-II composite Shewhart-GWMA scheme using the Mann–Whitney U statistic. Shah et al. [38] proposed a new monitoring scheme has been developed for time between events under exponential and gamma distribution. Shaheen et al. [39] developed a control chart using the lognormal distribution to monitor the variation under repetitive sampling scheme. Aslam et al. [40] developed and designed the modification of multiple dependent states' sampling plan for satisfying a mean lifetime of the products under Birnbaum-Saunders and Weibull distributions.

In modeling of real-data problems, the common failure time distribution needs not to be the same but can be a mixture of different lifetime distributions. Each of these distinct lifetime distributions can represent a different type of failure cause for the population, like weather change depending on different parameters such as temperature, humidity, wind flow, and precipitation. Each parameter follows different distributions, which together model the weather change. In this article, a new class of Weibull-G distributions is derived by using the method given by Alzaatreh et al. [6]. In fact, we follow a certain methodology, presented in equations (1) and (2). According to this methodology, we need $H(g(x))$ and $r(t)$, and for it, we use exponential and Weibull distributions. The exponential is a special case of Weibull distribution, but when we use the given methodology, the resulting distribution appears as the combination of exponential and Weibull distribution. Indeed, it is the combination of these two distributions, not the mixture of distributions. The benefit of using the proposed distribution is that it provides an extraparameter, which controls the kurtosis and provides better efficiency in detecting early shift. This can be seen from the comparison of proposed distribution with mere simple exponential and simple Weibull distributions. Thus, the four parameter distributions are more useful to explain several real-life phenomena much better than two or five parameters (see Berk et al. [41]). In quality control, this feature helps to detect out-of-control process much early.

We develop an attribute control chart scheme using WEx distribution. The parameter estimation of the WEx distribution is carried out through the maximum likelihood estimation method. The designed scheme is used to monitor the process out-of-control and found more suitable than existing schemes. We compare the results of the simulated data and real data of urinary tract infection in the health care department using the proposed scheme with the existing method based on ARL and control charts.

2. Weibull-G Class of Distributions

In this section, we derive the cumulative distribution function (cdf) and the probability density function (pdf) of the Weibull-G class of distributions. For this purpose, we use the method of Alzaatreh et al. [6]. The cdf of the proposed method is given as

$$F(y) = \int_{a}^{H(G(x))} r(t)\mathrm{d}t, \tag{1}$$

and the pdf is given as follows:

$$f(x) = \left\{\frac{\mathrm{d}}{\mathrm{d}x} H(G(x))\right\} r(H(G(x))), \tag{2}$$

where $r(t)$ is the pdf of the parent distribution and $H(.)$ should satisfy the following conditions:

(i) $H(G(x)) \in [a, b]$.

(ii) $H(G(x))$ is differentiable and monotonically nondecreasing.

(iii) $H(G(x)) \longrightarrow a$ as $x \longrightarrow -\infty$, and $H(G(x)) \longrightarrow b$ as $x \longrightarrow \infty$.

Using $H(G(x)) = -\log[(1-G(x))^{\alpha}]$ and $r(t)$ of Weibull distribution, the cdf of the WG class of distributions is obtained as

$$F(x) = \int_a^{-\log\left[(1-G(x))^{\alpha}\right]} \frac{c}{\gamma}\left(\frac{t}{\gamma}\right)^{c-1} \exp\left(-\left(\frac{t}{\gamma}\right)^{c}\right) dt = 1 - \exp\left(-\left(\frac{-\log\left((1-G(x))^{\alpha}\right)}{\gamma}\right)^{c}\right), \tag{3}$$

and pdf is obtained by differentiating (3) as

$$f(x) = \alpha\left(\frac{c}{\gamma}\right)\left[\frac{-\log\ [1-G(x)]^{\alpha}}{\gamma}\right]^{c-1} \cdot e^{-\left[\left(-\log\ [1-G(x)]^{\alpha}/\gamma\right)\right]^{c}} \frac{g(x)}{1-G(x)}. \tag{4}$$

Now, several new distributions can be generated for the different values of $G(x)$.

3. Weibull-Exponential Distribution

In this section, we derive Weibull-exponential distribution using the proposed class of distributions.

If x is an exponential distribution with cdf and pdf, then it is

$$G(x) = 1 - e^{-\lambda x}, \tag{5}$$

$$g(x) = \lambda e^{-\lambda x}, \tag{6}$$

respectively. The cdf and the pdf of the WEx distribution is generated by inserting (5) and (6) in (3) as

$$F(x) = 1 - e^{-(\alpha\lambda x/\gamma)^{c}},$$
$$f(x) = c\left(\frac{\alpha\lambda}{\gamma}\right)^{c} x^{c-1} e^{-(\alpha\lambda x/\gamma)^{c}}, \quad x>0, \alpha>0, \gamma, \lambda>0. \tag{7}$$

Some special cases of the proposed distribution are given in Appendix (Table 1).

4. The Proposed Control Chart

This article proposes the new scheme of attribute characteristics control charts for the manufacturing process using WEx distribution. For a fixed sampling interval, where n is considered as a sample for each subgroup, we discuss the lifetime of the product when random variable follows the proposed model. If the selected unit falls within control limits, we accept it; otherwise, it would be rejected. Therefore, we present the lower control limit (LCL) and upper control limit (UCL) as follows.

The LCL and UCL are defined as

$$P_r\left(X < \mathrm{LCL}|c_0\right) = P_L,$$
$$P_r\left(X < \mathrm{UCL}|c_0\right) = P_U, \tag{8}$$

Table 1: Special cases of the proposed distribution.

Special cases	Parameters	Distribution
Case 1	$(\alpha\lambda/\gamma) = \theta$	$f(x) \sim W(\theta, c)$
Case 2	If $c = 1, \lambda = 1, \alpha = 1$	$f(x) \sim \mathrm{Exp}(\gamma)$
Case 3	If $\gamma = 1, \lambda = 1, \alpha = 1$	$f(x) \sim W(c)$
Case 4	If $c = 1, \alpha = 1$	$f(x) \sim \mathrm{Exp}(\lambda, \gamma)$
Case 5	If $c = 2$	$f(x) \sim \mathrm{Rayleigh}(\alpha, \lambda, \gamma)$

respectively, where c_0 is the shape parameter. The LCL and UCL for the control process by following Azam et al. [34] are given by

$$\mathrm{LCL} = \frac{\gamma}{\alpha\lambda}\left[\ln\frac{1}{1-P_L}\right]^{(1/c)}, \tag{9}$$

$$\mathrm{UCL} = \frac{\gamma}{\alpha\lambda}\left[\ln\frac{1}{1-P_U}\right]^{(1/c)}. \tag{10}$$

Ho and Quinino [30] presented the derivation of control charts.

Firstly, a sample is chosen at random from each subgroup and its quality characteristics X are measured. The selected unit is accepted if it lies within defined limits and otherwise declared as rejected.

Secondly, the process is in control if a item is accepted at first and out-of-control if b item is rejected first.

When the process is in control, the probability that the process is out-of-control is defined as

$$P_0 = P\left(\mathrm{LCL} < X < \mathrm{UCL}|c = c_0\right) = 1 - P_U + P_L. \tag{11}$$

Moreover, for the in-control process, the probability of the out-of-control process is demonstrated through control charts is given as

$$P_{\mathrm{out}}^{0} = 1 - \sum_{x=a}^{a+b-1}\binom{x-1}{a-1} P_o^{a}\left(1-P_0\right)^{x-a}. \tag{12}$$

Furthermore, the shape parameter is converted to c_1. Then, the probability of the rejected item is derived as

$$P_1 = P\left(\mathrm{LCL} < X < \mathrm{UCL}|c = c_1\right). \tag{13}$$

The probability that the process is found to be out-of-control due to the shifted process using control chart is obtained as

$$P_{\mathrm{out}}^{1} = 1 - \sum_{x=a}^{a+b-1}\binom{x-1}{a-1} P_1^{a}\left(1-P_1\right)^{x-a}. \tag{14}$$

Now, the shape parameter is considered for the shifted process, which is given as follows:

$$c = f * c_1, \tag{15}$$

where f is a shifted constant and lies between 0.25 and 1. Therefore, (11) is formed as

$$P_1 = \exp\left[\left(\frac{\alpha\lambda \mathrm{UCL}}{\gamma}\right)^{fc_1}\right] + 1 - \exp\left[\left(\frac{\alpha\lambda \mathrm{LCL}}{\gamma}\right)^{fc_1}\right],$$

$$P_1 = 1 + \exp\left[-\left(\frac{\alpha\lambda}{\gamma} * \frac{\gamma}{\alpha\lambda}\right)\left(\left(\ln\frac{1}{1-P_U}\right)^{\frac{1}{c_1}}\right)^{fc_1}\right] - \exp\left[-\left(\frac{\alpha\lambda}{\gamma} * \frac{\gamma}{\alpha\lambda}\right)\left(\left(\ln\frac{1}{1-P_L}\right)^{(1/c_1)}\right)^{fc_1}\right], \tag{16}$$

$$P_1 = 1 + \exp\left[-\ln\left(\frac{1}{1-P_U}\right)^{f}\right] - \exp\left[-\ln\left(\frac{1}{1-P_L}\right)^{f}\right].$$

The ARL is evaluated using P_0 and P_1 which are obtained through the parameters of the proposed distribution. Note here that, the equations of P_0 and P_1 are independent of scale parameters. Therefore, we fixed only the shape parameter.

Now, the ARL of in-control and ARL of out-of-control have been calculated as

$$\mathrm{ARL}_0 = \frac{1}{P_{\mathrm{out}}^0} \text{ and } \mathrm{ARL}_1 = \frac{1}{P_{\mathrm{out}}^1}. \tag{17}$$

5. ARL Behavior of the Newly Proposed Control Chart

In this section, we discuss the ARL behavior of the suggested control chart. A simulation study is conducted in the R-language program using the above-estimated equations of probabilities. In simulation study, 10,000 samples are generated from the well-known exponential, Weibull, and the proposed WEx distributions. The ML estimates of the parameters of the suggested distribution are obtained by using SANN (simulated annealing) given in Belisle [42]. SANN is an optimization technique, which is more recent and flexible as compared to Newton–Raphson. The SANN method is available in R-package maxLik Henningsen and Toomet [43]. We fixed the shape parameters to find the ARLs.

These ML estimates are used to find ARLs for $\mathrm{ARL}_0 = 200, 300$, and 370. Furthermore, ARLvalues are explained by calculating for different drift levels. It is important to mention that the shift unit is equal to the in-control ARL. The performance of ARL can be understood by reading the R-code program. Simulated data of 10,000 samples are generated by choosing the random values of a, b, and n. For in-control processes, the ARL is calculated by using these given random values. For the selection of a and b parameters, see details in Ho and Quinino [30]. Moreover, we compare the exponential and Weibull distribution with the proposed WEx distribution for the efficiency of ARL.

6. Comparative Study

In this section, we compare the performance of the proposed WEx distribution with the exponential distribution under ARL. In Table 2, we report the estimated values of the proposed model parameters. Various values of the parameters of exponential, Weibull, and the proposed WEx distributions are placed in Tables 3–5 for various combinations of a and b at $\mathrm{ARL}_0 = 200, 300$, and 370. From Table 3, it is seen that various values of a and b at $f = 1$, i.e., there is no shift in the values of ARLs of the exponential, Weibull, and the proposed WEx distributions, are the same; i.e., 201.98, 303.33, and 371.99 at $\mathrm{ARL}_0 = 200, 300$, and 370. On the other hand, it is observed that, at a different shift level $f = 1$ to 0.25, the values of ARL decrease for the exponential and the proposed models.

One can clearly observe in Tables 3–5 that the proposed model suppressed the exponential and Weibull distribution in terms of shift detection. From Table 3, for instance, the value of ARL is 17.04 for the WEx distribution and 65.5 for the exponential distribution when $f = 0.85$, $a = 9, b = 20$, and $\mathrm{ARL}_0 = 200$. Furthermore, Tables 4 and 5 also represent the values of ARLs of the exponential and the proposed models at $\mathrm{ARL}_0 = 200$, 300, and 370 with $P_L = 0.1$ and $P_U = 0.9$ and $P_L = 0.2$ and $P_U = 0.8$ for various combinations of a and b at different levels of shifts. It can be noticed that the values of ARLs decrease for both models when the values of a and b decrease. By comparing both models, it is concluded that the proposed model gives quick detection of out-of-control process at different levels of shifts. Tables 4 and 5 are presented in Appendix.

Figures 1–9 demonstrate the graphical representations of the performance of ARLs at 200, 300, and 370 for several combinations of $a\&b$ and $P_L\&P_U$. One can see from Figures 1–9 as shift level f decreases the curves of ARLs, the exponential and the proposed WEx distributions also decrease. Also, the ARL curves show that the behavior of the proposed model is better than that of the compared model. Figures 4–9 are also presented in Appendix.

7. Real-Life Example

In this section, the proposed control chart is implemented in real-life data set from the health sector. The data of urinary tract infection (UTI) are taken from the hospital system. These data are attained from Azam et al. [34] which was initially used by Santiago and Smith [29]. Data are collected from a hospital with a high risk of urinary tract infections particularly to identify risk rate

Table 2: Maximum likelihood estimation of the proposed WEx distribution parameter.

Maximum likelihood estimation SANN maximization, 10000 iterations Return code: 0 successful convergence Log-likelihood: −42.88574				
Parameter estimates	Estimate	Std. error	t value	P value
α	2.74	0.06969	39.3169	$2e-16$
C	1.328	4.4492	0.29484	0.653
γ	7.01	7.1518	0.9801	0.613
λ	4.88	11.3607	0.4293	0.697

Table 3: Comparison of ARL (SDRL) using exponential and Weibull distributions and WEx distribution with $p_L = 0.05$ and $p_U = 0.95$.

f	$ARL_0 = 200$ $a = 9, b = 20$			$ARL_0 = 300$ $a = 7, b = 10$			$ARL_0 = 370$ $a = 13, b = 37$		
	Exp	Weibull	WEx	Exp	Weibull	WEx	Exp	Weibull	WEx
1	201.98 (201.48)	201.98 (201.48)	201.98 (201.48)	303.33 (302.83)	303.33 (302.83)	303.33 (302.83)	371.99 (371.49)	371.99 (371.49)	371.99 (371.49)
0.99	189.71 (189.21)	228.04 (227.54)	167.50 (166.99)	287.91 (287.41)	335.52 (335.02)	259.48 (258.98)	341.46 (340.96)	439.25 (438.75)	288.13 (287.62)
0.95	144.93 (144.43)	351.89 (351.39)	81.72 (81.22)	229.87 (229.37)	480.20 (479.70)	141.65 (141.15)	236.66 (236.16)	798.29 (797.79)	109.4 (108.90)
0.9	99.43 (98.93)	513.59 (513.09)	35.81 (35.30)	167.32 (166.82)	(654.78) (654.28)	69.45 (68.95)	142.27 (141.77)	1348.95 (1348.45)	36.91 (36.41)
0.85	65.5 (64.99)	586.54 (586.04)	17.04 (16.54)	117.23 (116.73)	729.78 (729.28)	35.81 (35.30)	81.49 (80.99)	1623.17 (1622.66)	14.37 (13.86)
0.8	41.7 (41.20)	507.34 (506.84)	8.84 (8.32)	79.35 (78.85)	648.25 (647.75)	19.44 (18.94)	45.00 (44.49)	1326.14 (1325.64)	6.49 (5.97)
0.75	25.88 (25.38)	337.90 (337.40)	5.01 (4.48)	52.14 (51.63)	464.42 (463.92)	11.14 (10.63)	24.32 (23.82)	754.76 (754.26)	3.41 (2.87)
0.7	15.83 (15.32)	183.37 (182.87)	3.11 (2.56)	33.47 (32.97)	279.86 (279.36)	6.74 (6.22)	13.11 (12.60)	325.98 (325.47)	2.09 (1.51)
0.65	9.66 (9.15)	87.14 (86.64)	2.12 (1.54)	21.15 (20.65)	149.60 (149.10)	4.33 (3.79)	7.21 (6.69)	119.20 (118.70)	1.48 (0.838)
0.6	5.97 (5.45)	38.70 (38.20)	1.58 (0.953)	13.28 (12.77)	74.35 (73.84)	2.94 (2.39)	4.14 (3.60)	40.83 (40.33)	1.19 (0.476)
0.55	3.80 (3.26)	16.89 (16.39)	1.28 (0.599)	8.37 (7.85)	35.52 (35.02)	2.13 (1.55)	2.55 (1.99)	14.22 (13.71)	1.06 (0.259)
0.5	2.53 (1.97)	7.54 (7.02)	1.12 (0.371)	5.35 (4.82)	16.70 (16.19)	1.64 (1.03)	1.72 (1.12)	5.40 (4.87)	1.02 (1.13)
0.25	1.01 (0.098)	1.001 (0.0296)	1 (0.005)	1.15 (0.409)	1.04 (0.211)	1 (0.082)	1 (0.012)	1.00 (0.0013)	1 (0.00004)

and the frequency of the UTI patients being discharged. In this example, male patients are focused. The data show the number of days between admissions and discharge of the patients having UTIs. The simulated data for the above scenario are generated from the proposed WEx distribution with $P_L = 0.1, P_U = 0.9$, $a = 29$, and $b = 59$. Firstly, we generate 40 observations for the in-control process and then the next 40 observations for the shifted process with $f = 0.90$, as presented in Tables 6 and 7. In addition, we estimate the values of LCL and UCL as 0.03206 and 0.95051, respectively, and display in Figure 10. Figures 10 and 11 show the LCL and UCL for in-control and out-of-control data. It is clearly shown that 40th value is detected as out-of-control for the proposed model.

8. Comparisons of the Exponential and the Weibull-Exponential Distribution Using Real-Life Example

In this section, we compare the control charts of the exponential and the proposed WEx distributions using real data of UTI patients. It can be easily seen in Figure 12 that the control chart of the exponential distribution is in-control and no value is detected as out-of-control,

Table 4: Comparison of ARL (SDRL) using exponential and Weibull distributions and WEx distribution with $\mathbf{p_L} = 0.1$ and $\mathbf{p_U} = 0.9$

f	$ARL_0 = 200$, $a = 29, b = 59$			$ARL_0 = 300$, $a = 23, b = 40$			$ARL_0 = 370$, $a = 34, b = 70$		
	Exp	Weibull	WEx	Exp	Weibull	WEx	Exp	Weibull	WEx
1	200.01 (199.51)	200.01 (199.51)	200.01 (199.51)	304.99 (304.49)	304.99 (304.49)	304.99 (304.49)	374.91 (374.41)	374.91 (374.41)	374.91 (374.41)
0.99	181.27 (180.77)	242.46 (241.96)	150.1 (149.60)	279.28 (278.78)	362.15 (361.65)	235.77 (235.27)	334.49 (333.99)	468.70 (468.20)	268.78 (268.28)
0.95	119.61 (119.11)	494.23 (493.76)	51.55 (51.05)	192.14 (191.64)	680.54 (680.03)	89.12 (88.62)	206.62 (206.12)	1071.88 (1071.38)	78.115 (77.61)
0.9	68.05 (67.55)	967.05 (966.52)	16.29 (15.78)	115.04 (114.54)	1225.84 (1225.34)	30.13 (29.62)	107.63 (107.13)	2340.44 (2339.94)	20.78 (20.28)
0.85	37.3 (36.79)	1294.41 (1293.91)	6.35 (5.83)	66.01 (65.51)	1580.35 (1579.85)	11.85 (11.34)	53.79 (53.29)	3286.59 (3286.09)	7.12 (6.60)
0.8	20.02 (19.51)	1062.93 (1062.43)	3.06 (2.51)	36.71 (36.20)	1331.21 (1330.71)	5.45 (4.92)	26.31 (25.81)	2612.75 (2612.25)	3.15 (2.60)
0.75	10.74 (10.23)	522.02 (521.52)	1.82 (1.22)	20.08 (19.57)	714.11 (713.61)	2.94 (2.39)	12.91 (12.40)	1142.25 (1141.75)	1.79 (1.18)
0.7	5.90 (5.38)	166.63 (166.13)	1.31 (0.633)	11.00 (10.48)	258.97 (258.47)	1.86 (1.26)	6.55 (6.03)	303.37 (302.87)	1.27 (0.583)
0.65	3.42 (2.88)	41.37 (40.87)	1.1 (0.328)	6.16 (5.64)	72.69 (72.19)	1.36 (0.702)	3.56 (3.02)	60.61 (60.11)	1.08 (0.284)
0.6	2.15 (1.57)	10.09 (9.58)	1.02 (0.155)	3.62 (3.08)	18.88 (18.38)	1.14 (0.392)	2.14 (1.56)	12.03 (11.52)	1.01 (0.122)
0.55	1.51 (0.872)	3.07 (2.53)	1 (0.061)	2.29 (1.72)	5.47 (4.95)	1.04 (0.206)	1.46 (0.8244)	3.16 (2.61)	1 (0.041)
0.5	1.19 (0.476)	1.42 (0.770)	1 (0.018)	1.59 (0.9692)	2.10 (1.520)	1.01 (0.095)	1.16 (0.4275)	1.38 (0.7208)	1 (0.010)
0.25	1 (0.0004)	1.00 (---)	1 (0.000)	1 (0.0079)	1.00 (0.000001)	1 (0.00002)	1 (0.0001)	1.00 (0.00000)	1 (0.000)

Table 5: Comparison of ARLs (SDRL) using exponential and Weibull distribution and WEx distribution with $\mathbf{p_L} = 0.2$ and $\mathbf{p_U} = 0.8$.

f	$ARL_0 = 200$, $a = 63, b = 57$			$ARL_0 = 300$, $a = 54, b = 45$			$ARL_0 = 370$, $a = 42, b = 31$		
	Exp	Weibull	WEx	Exp	Weibull	WEx	Exp	Weibull	WEx
1	200.21 (199.71)	200.21 (199.71)	200.21 (119.71)	300.51 (300.01)	300.51 (300.01)	300.51 (300.01)	370.82 (370.32)	370.82 (370.32)	370.82 (370.32)
0.99	182.13 (181.63)	241.05 (240.55)	152.62 (152.12)	274.70 (274.20)	358.29 (357.78)	232.2 (231.70)	342.73 (342.23)	432.55 (432.05)	295.64 (295.14)
0.95	122.3 (121.80)	483.80 (483.30)	55.17 (54.67)	188.00 (187.50)	690.67 (690.17)	87.45 (86.95)	245.41 (244.91)	765.86 (765.35)	124.17 (123.67)
0.9	71.44 (70.94)	961.52 (961.02)	18.14 (17.63)	112.25 (111.75)	1312.93 (1312.43)	29.35 (28.84)	155.25 (154.75)	1333.46 (1332.96)	45.98 (45.48)
0.85	40.3 (39.80)	1357.57 (1357.07)	7.18 (6.66)	64.44 (63.94)	1809.98 (1809.48)	11.45 (10.94)	94.33 (93.83)	1756.73 (1756.23)	18.96 (18.45)
0.8	22.24 (21.74)	1181.92 (1181.42)	3.44 (2.90)	35.94 (35.44)	1591.17 (1590.67)	5.24 (4.71)	55.43 (54.92)	1572.92 (1572.42)	8.76 (8.24)
0.75	12.22 (11.71)	578.96 (578.46)	2.00 (1.41)	19.74 (19.23)	817.31 (816.81)	2.82 (2.27)	31.79 (31.28)	886.02 (885.52)	4.56 (4.03)
0.7	6.82 (6.30)	162.28 (161.77)	1.39 (0.740)	10.86 (10.34)	246.16 (245.66)	1.79 (1.19)	18.00 (17.49)	311.23 (310.73)	2.69 (2.13)
0.65	3.96 (3.42)	31.11 (30.60)	1.14 (0.393)	6.10 (5.58)	50.03 (49.52)	1.32 (0.654)	10.22 (9.70)	74.98 (74.48)	1.80 (1.202)
0.6	2.45 (1.89)	5.83 (5.31)	1.04 (0.195)	3.59 (3.05)	9.22 (8.71)	1.12 (0.358)	5.92 (5.39)	15.37 (14.86)	1.36 (0.702)
0.55	1.67 (1.06)	1.71 (1.098)	1.01 (0.083)	2.27 (1.70)	2.33 (1.75)	1.03 (0.182)	3.57 (3.03)	3.68 (3.14)	1.15 (0.409)
0.5	1.27 (0.589)	1.05 (0.2310)	1 (0.028)	1.58 (0.954)	1.15 (0.4118)	1.01 (0.079)	2.29 (1.72)	1.44 (0.7909)	1.05 (0.225)
0.25	1 (0.0009)	1.00 (0.00000)	1 (0.0000)	1 (0.006)	1.00 (0.00000)	1 (0.00001)	1 (0.042)	1.00 (----)	1 (0.0004)

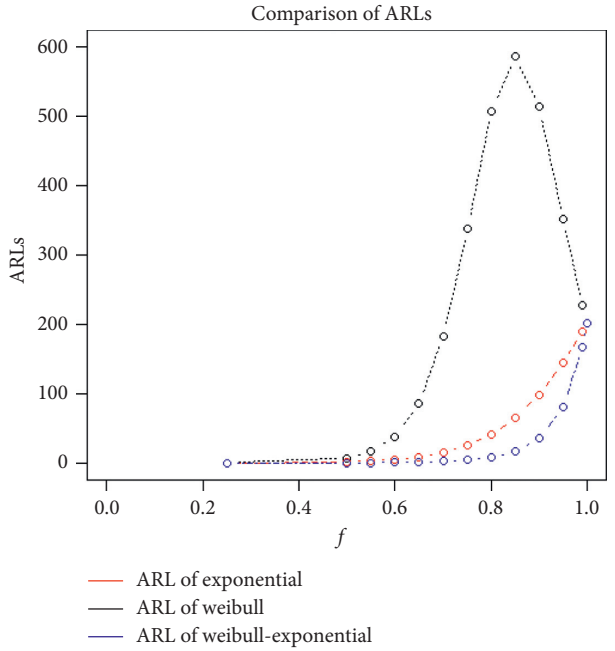


Figure 1: ARLs of exponential, Weibull, and proposed WEx at 200 with $\mathbf{p_L} = 0.05$ and $\mathbf{p_U} = 0.95$.

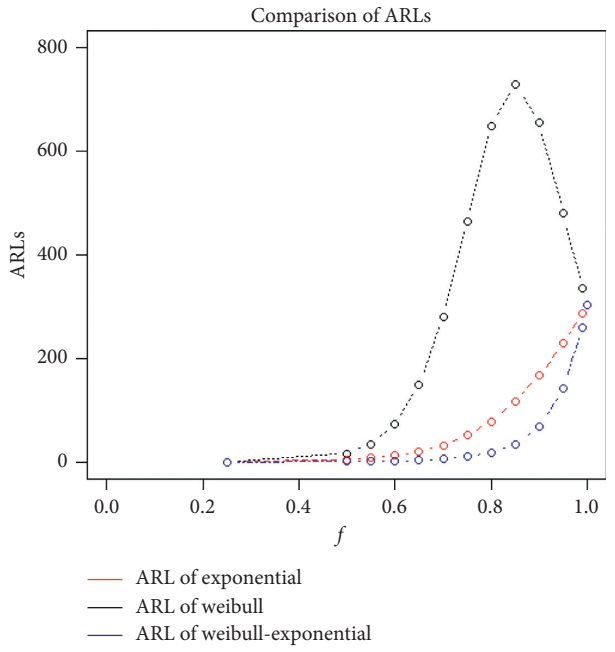


Figure 2: ARLs of exponential, Weibull, and proposed WEx at 300 with $p_L = 0.05$ and $p_U = 0.95$.

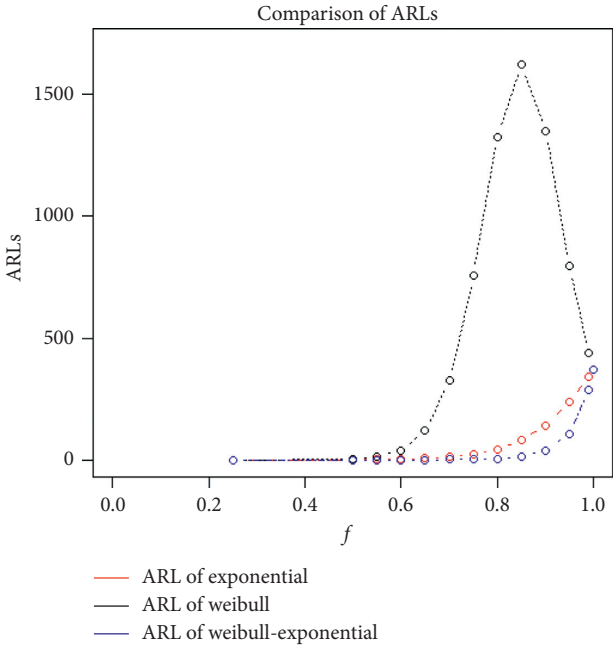


Figure 3: ARLs of exponential, Weibull, and proposed WEx at 370 with $\mathbf{p_L} = 0.05$ and $\mathbf{p_U} = 0.95$.

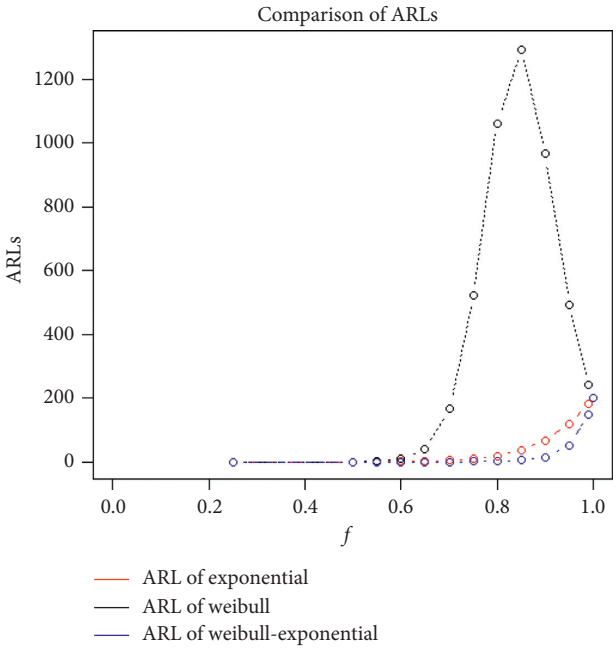


Figure 4: ARLs of exponential, Weibull, and proposed WEx at 200 with $p_L = 0.1$ and $p_U = 0.9$.

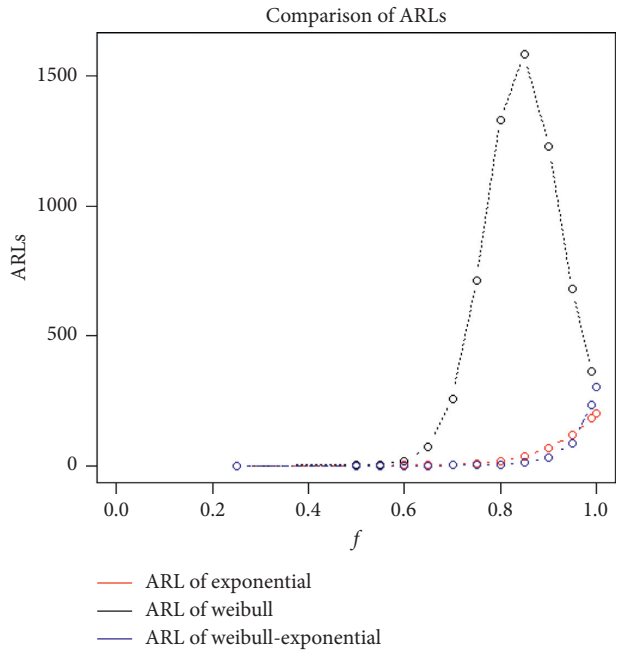


Figure 5: ARLs of exponential, Weibull, and proposed WEx at 300 with $p_L = 0.1$ and $p_U = 0.9$.

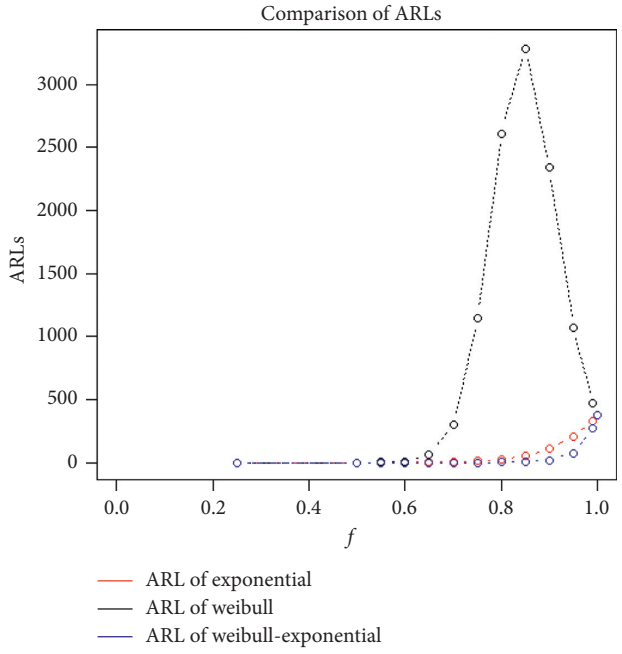


Figure 6: ARLs of exponential, Weibull, and proposed WEx at 370 with $\mathbf{P_L} = 0.1$ **an d** $\mathrm{P_U} = 0.9$.

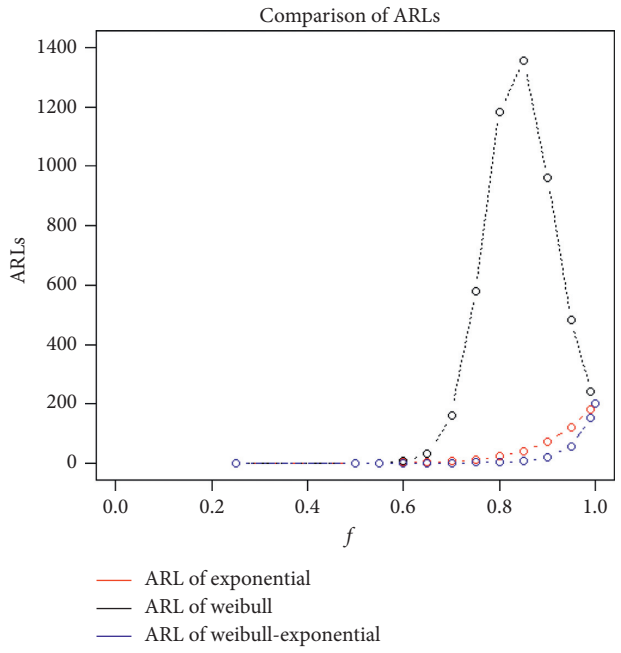


Figure 7: ARLs of exponential, Weibull, and proposed WEx at 200 with $p_L = 0.2$ and $p_U = 0.8$.

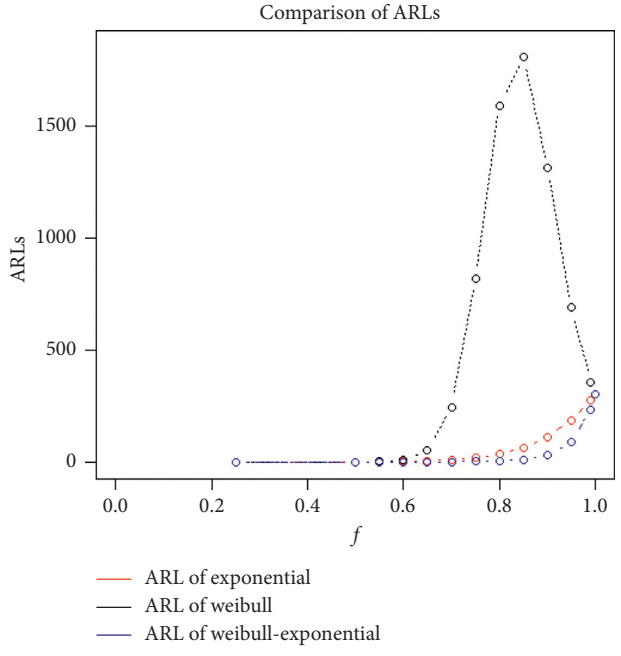


Figure 8: ARLs of exponential, Weibull, and proposed WEx at 300 with $p_L = 0.2$ and $p_U = 0.8$.

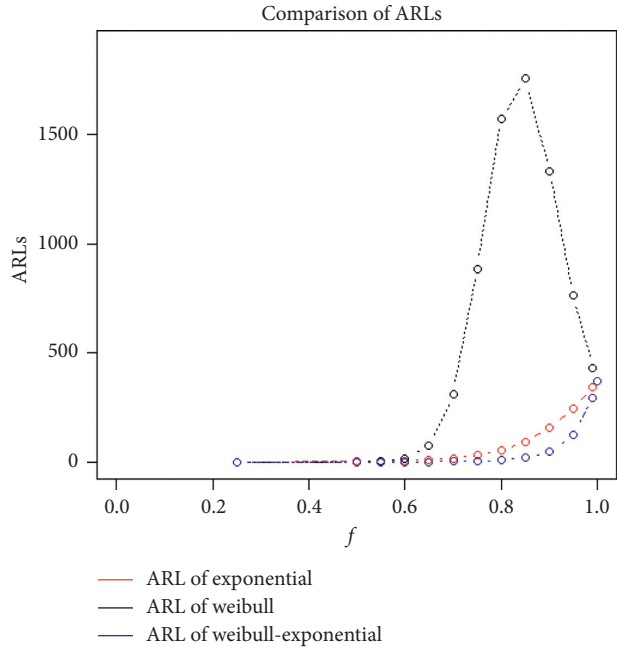


Figure 9: ARLs of exponential, Weibull, and proposed WEx at 370 with $\mathbf{p}_L = 0.2$ and $\mathbf{p}_U = 0.8$.

Table 6: Simulated in-control data.

0.220499	0.467107	0.165884	0.478009	0.186779	0.058493
0.115906	0.088989	0.450255	0.287988	0.867485	0.253406
0.4247	0.714495	0.356458	0.26482	0.503415	0.529773
0.278757	0.08675	0.204802	0.181542	0.417945	0.103082
0.269525	0.543777	0.096818	0.04666	0.070134	0.17702
0.166788	0.632749	0.482233	0.208581	0.354741	0.242477
0.592133	0.837021	0.513781	0.086579		

Table 7: Simulated shifted data.

0.650604	0.144811	0.585181	0.1867	0.613353	0.501533
0.092963	0.733554	0.192515	0.154301	0.913754	0.573958
0.462511	0.190089	0.58542	0.297584	0.162154	0.080671
0.066036	0.746633	0.545823	0.061	0.198106	0.92125
0.057984	0.692892	0.26246	0.463523	0.204097	0.330734
0.454024	0.765241	0.51748	0.676308	0.341994	0.470963
0.289353	0.391563	0.502378	0.011261		

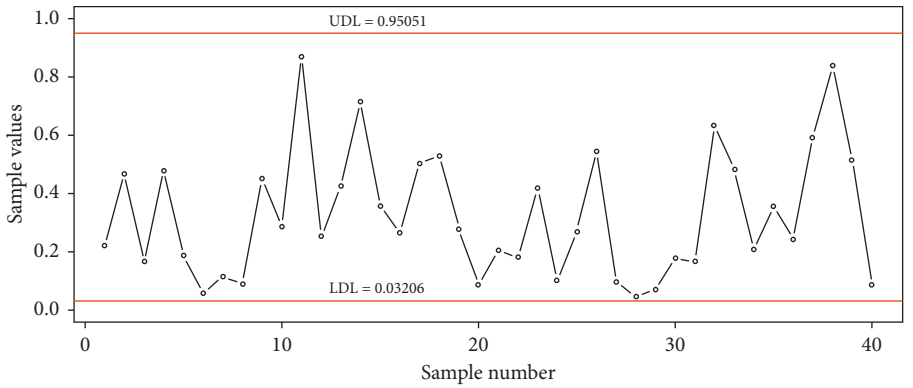


Figure 10: Plot for simulated in-control data.

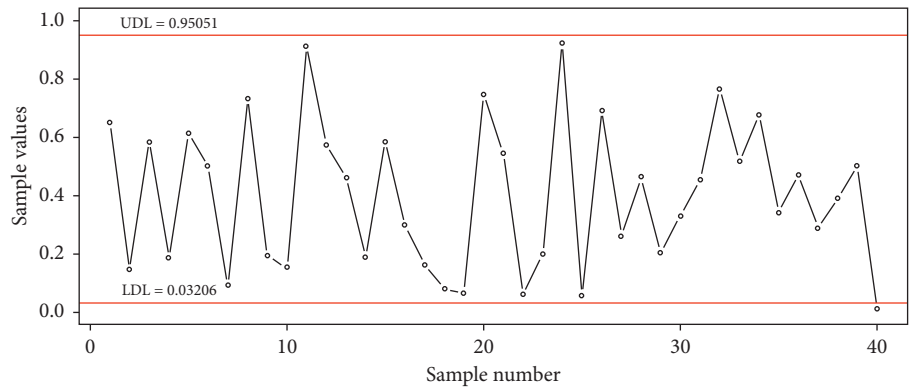


Figure 11: Plot for simulated shift data with $f = 0.9$, $p_L = 0.1$, $p_U = 0.9$, and $\mathrm{ARL}_0 = 200$.

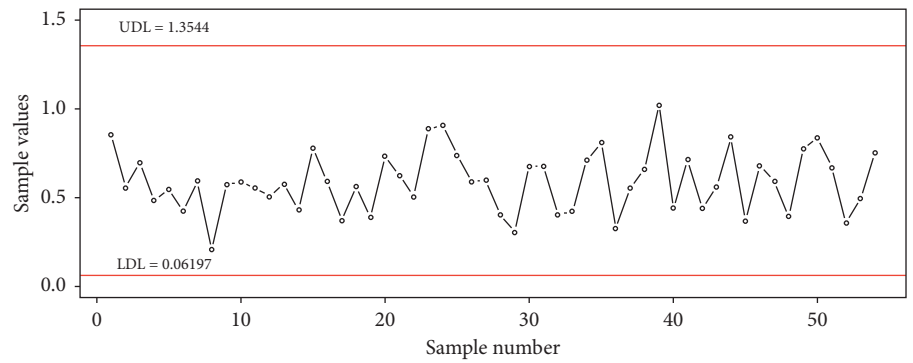


Figure 12: Plot of UTI using exponential distribution control chart at $p_L = 0.1$ and $p_U = 0.9$.

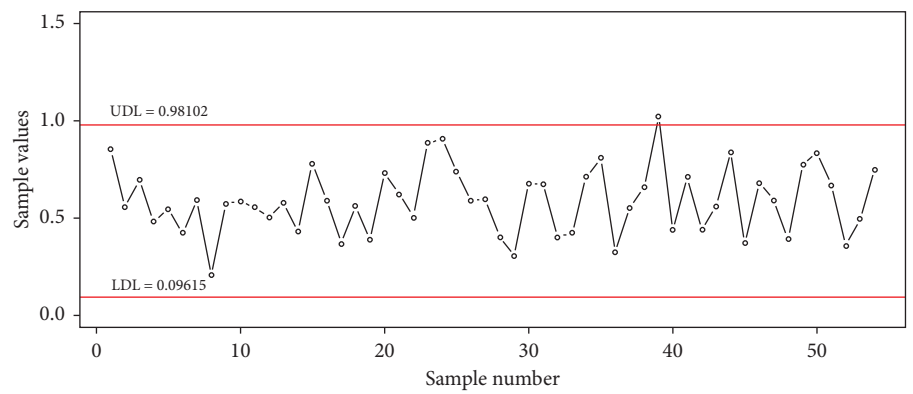


Figure 13: Plot of UTI using proposed (WEx) control chart at $p_L = 0.1$ and $p_U = 0.9$.

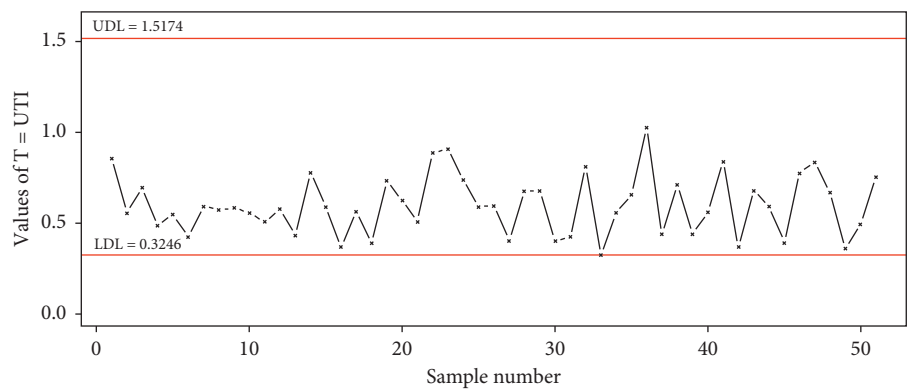


Figure 14: Plot of UTI using Weibull distribution control chart at $p_L = 0.1$ and $p_U = 0.9$.

whereas the control chart of the proposed distribution which can be seen in Figure 13 that detected that the process is out-of-control at the 40th value. It may be concluded that our proposed distribution using real data set of UTI shows better performance to detect the out-of-control process quickly.

9. Conclusion

In this paper, we develop a new Weibull class of distributions using the distribution generator technique. The newly proposed WEx distribution is applied in the attribute characteristics control chart for monitoring the manufacturing process. Furthermore, we compare the proposed WEx distribution with exponential distribution to monitor the quick detection of the item in the out-of-control process. We conclude that, under various shift levels, our proposed model is more efficient than the exponential distribution. A simulation study is also conducted to known the behavior of ARL using the proposed WEx, Weibull, and exponential distributions. The implementation of the proposed model in a real-life example from the health care department and the comparison of the two models endorse the scope of WEx distribution. We hope that this proposed control chart can also be used for other accepting sampling schemes and other lifetime models (Figure 14).

Notations

W: Weibull
WEx: Weibull-exponential
R: Range chart
S^2 chart: Variance chart
W-G: Weibull-generalized
P_L, P_U: Percentiles
UDL: Upper discriminant limit
LDL: Lower discriminant limit
a: Accepted item
b: Rejected item
f: Shift level
$\alpha, \gamma, c, \lambda$: Parameters of Weibull-exponential distribution
ARL: Average run-length
P^0_{out}: Probability of out-of-control process
P^1_{out}: Probability of out-of-control process of shifted process
P_0: Probability of rejected item when process is in-control
P_1: Probability of rejected item when process is shifted
UTI: Urinary tract infection.

Appendix

A. R-Code

Derivation of Equation (9) (Algorithms 1 and 2)

Input: sample sizes (ss), P_L, P_U, and r_0
$P_0 = 1 - P_U + P_L$
a = sample(1 : 70, 1, T); b = sample(1 : 70, 1, T)
$u = a + \text{b} - 1$; $l = a - 1$; $n = a + b - 1$
Pout0 = sum(dbinom(l:u, n, $p0$))
ARL0[i] = 1/Pout0
ARL_0 [i] = 1/Pout0; aa[i] = a; bb[i] = b; nn[i] = n}
Impose condition $\text{ARL}_0 \geq r_0$
Output: obtain corresponding values of a, b, n, and ARL_0

Algorithm 1: Computation of a, b, n, and ARL_0.

Input: different values of shift (f)
$f1 = \log(1/(1 - p_U))$; $f2 = \log(1/(1 - p_L))$
$p1 = 1 + \exp(-f1 * f) - \exp(-f2 * f)$
$u = a + b - 1$; $l = a - 1$
beta = sum(dbinom(l:u, n, $p1$))
ARL_1 = 1/beta; SDRL = sqrt((1−beta)/((beta)^2))
Output: ARL for different values of shift.

Algorithm 2: Computation of ARL for different shifts.

$$P_r\left(X < \text{LCL} | c_0\right) = P_L$$

$$1 - \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right) = P_L \tag{A.1}$$

$$1 - P_L = \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right).$$

Taking ln on both sides,

$$\ln\left(1 - P_L\right) = \ln \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right)$$

$$-\ln\left(1 - P_L\right) = \left(\frac{\alpha\lambda x}{\gamma}\right)^c$$

$$\left[-\ln\left(1 - P_L\right)\right]^{(1/c)} = \left(\frac{\alpha\lambda x}{\gamma}\right) \tag{A.2}$$

$$\left[\ln\left(\frac{1}{1 - P_L}\right)\right]^{(1/c)} = \left(\frac{\alpha\lambda x}{\gamma}\right)$$

$$\frac{\gamma}{\alpha\lambda}\left[\ln\left(\frac{1}{1 - P_L}\right)\right]^{(1/c)} = x.$$

The LCL is $(\gamma/\alpha\lambda)[\ln(1/1 - P_L)]^{(1/c)}$.

Derivation of Equation (10)

$$P_r\left(X < \mathrm{UCL} | c_0\right) = P_U,$$
$$1 - \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right) = P_U, \tag{A.3}$$
$$1 - P_U = \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right).$$

Taking ln on both sides,

$$\ln\left(1 - P_U\right) = \ln\ \exp\left(-\left(\frac{\alpha\lambda x}{\gamma}\right)^c\right)$$
$$-\ln\left(1 - P_U\right) = \left(\frac{\alpha\lambda x}{\gamma}\right)^c$$
$$\left[-\ln\left(1 - P_u\right)\right]^{(1/c)} = \left(\frac{\alpha\lambda x}{\gamma}\right) \tag{A.4}$$
$$\left[\ln\left(\frac{1}{1 - P_U}\right)\right]^{(1/c)} = \left(\frac{\alpha\lambda x}{\gamma}\right)$$
$$\frac{\gamma}{\alpha\lambda}\left[\ln\left(\frac{1}{1 - P_U}\right)\right]^{(1/c)} = x.$$

The UCL is $(\gamma/\alpha\lambda)[\ln(1/1 - P_L)]^{(1/c)}$.

Equation (12) Derivation

For the in-control process, the probability of the out-of-control process is demonstrated through control charts is given as

$$P_{\text{out}}^0 = 1 - \sum_{x=a}^{a+b-1}\binom{x-1}{a-1}P_o^a\left(1 - P_0\right)^{x-a}. \tag{A.5}$$

The authors compute the probability of out-of-control process as $P(X \geq b)$.

Now,

$$P(X \geq b) = 1 - P(X < b). \tag{A.6}$$

Since "X" follows the negative binomial distribution, then

$$P(X < b) = \sum_{x=a}^{a+b-1}\binom{x-1}{a-1}p_0^a\left(1 - p_0\right)^{x-a}. \tag{A.7}$$

Equation (14) Derivation

The probability that the process is found out-of-control due to the shifted process using control chart is obtained as

$$P_{\text{out}}^1 = 1 - \sum_{x=a}^{a+b-1}\binom{x-1}{a-1}P_1^a\left(1 - P_1\right)^{x-a}. \tag{A.8}$$

The authors compute the probability of out-of-control process as $P(X \geq b)$.

Now,

$$P(X \geq b) = 1 - P(X < b). \tag{A.9}$$

Since "X" follows the negative binomial distribution, then

$$P(X < b) = \sum_{x=a}^{a+b-1}\binom{x-1}{a-1}p_1^a\left(1 - p_1\right)^{x-a}. \tag{A.10}$$

Data Availability

The data used to support the findings of this study are given in the paper.

Disclosure

No funds were obtained for this paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] N. Eugene, C. Lee, and F. Famoye, "Beta-normal distribution and its applications," *Communications in Statistics-Theory and Methods*, vol. 31, no. 4, pp. 497–512, 2002.

[2] S. Nadarajah and S. Kotz, "The beta Gumbel distribution," *Mathematical Problems in Engineering*, vol. 2004, no. 4, pp. 323–332, 2004.

[3] S. Nadarajah and A. K. Gupta, "The beta Fréchet distribution," *Far East Journal of Theoretical Statistics*, vol. 14, no. 1, pp. 15–24, 2004.

[4] S. Nadarajah and S. Kotz, "The beta exponential distribution," *Reliability Engineering & System Safety*, vol. 91, no. 6, pp. 689–697, 2006.

[5] S. Nadarajah, G. M. Cordeiro, and E. M. M. Ortega, "The zografos-balakrishnan-GFamily of distributions: mathematical properties and applications," *Communications in Statistics - Theory and Methods*, vol. 44, no. 1, pp. 186–215, 2015.

[6] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continuous distributions," *Metron*, vol. 71, no. 1, pp. 63–79, 2013.

[7] G. M. Cordeiro, E. M. M. Ortega, B. V. Popović, and R. R. Pescim, "The Lomax generator of distributions: properties, minification process and regression model," *Applied Mathematics and Computation*, vol. 247, pp. 465–486, 2014.

[8] A. Z. Afify, M. Alizadeh, H. M. Yousof, G. Aryal, and M. Ahmad, "The transmuted geometric-g family OF distributions: THEORY and applications," *Pakistan Journal of Statistics*, vol. 32, no. 2, pp. 139–160, 2016.

[9] M. Alizadeh, M. H. Tahir, G. M. Cordeiro, M. Mansoor, M. Zubair, and G. G. Hamedani, "The Kumaraswamy Marshal-Olkin family of distributions," *Journal of the Egyptian Mathematical Society*, vol. 23, no. 3, pp. 546–557, 2015.

[10] A. Z. Afify, G. M. Cordeiro, H. M. Yousof, A. Alzaatreh, and Z. M. Nofal, "The Kumaraswamy transmuted-G family of distributions: properties and applications," *Journal of Data Science*, vol. 14, no. 2, pp. 245–270, 2016.

[11] Z. M. Nofal, A. Z. Afify, H. M. Yousof, and G. M. Cordeiro, "The generalized transmuted-G family of distributions,"

Communications in Statistics-Theory and Methods, vol. 46, no. 8, pp. 4119–4136, 2017.

[12] M. Chahkandi and M. Ganjali, "On some lifetime distributions with decreasing failure rate," *Computational Statistics and Data Analysis*, vol. 53, no. 12, pp. 4433–4440, 2009.

[13] W. Barreto-Souza, A. H. S. Santos, and G. M. Cordeiro, "The beta generalized exponential distribution," *Journal of Statistical Computation and Simulation*, vol. 80, no. 2, pp. 159–172, 2010.

[14] W. Lu and D. Shi, "A new compounding life distribution: the Weibull-Poisson distribution," *Journal of Applied Statistics*, vol. 39, no. 1, pp. 21–38, 2012.

[15] D. C. Montgomery, *Statistical Quality Control*, Vol. 7, Wiley, New York, NY, USA, 2009.

[16] W. H. Woodall, "The use of control charts in health-care and public-health surveillance," *Journal of Quality Technology*, vol. 38, no. 2, pp. 89–104, 2006.

[17] S.-L. Hwang, J.-T. Lin, G.-F. Liang, Y.-J. Yau, T.-C. Yenn, and C.-C. Hsu, "Application control chart concepts of designing a pre-alarm system in the nuclear power plant control room," *Nuclear Engineering and Design*, vol. 238, no. 12, pp. 3522–3527, 2008.

[18] Z. Wang and R. Liang, "Discuss on applying SPC to quality management in university education," in *Proceedings of the 9th International Conference for Young Computer Scientists*, pp. 2372–2375, Hunan, China, November 2008.

[19] M. Pastell and H. Madsen, "Application of CUSUM charts to detect lameness in a milking robot," *Expert Systems with Applications*, vol. 35, no. 4, pp. 2032–2040, 2008.

[20] A. Ivanova, C. Xiao, and Y. Tymofyeyev, "Two-stage designs for phase 2 dose-finding trials," *Statistics in Medicine*, vol. 31, no. 24, pp. 2872–2881, 2012.

[21] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.

[22] D. Brook and D. A. Evans, "An approach to the probability distribution of CUSUM run length," *Biometrika*, vol. 59, no. 3, pp. 539–549, 1972.

[23] J. M. Lucas, "Combined Shewhart-CUSUM quality control schemes," *Journal of Quality Technology*, vol. 14, no. 2, pp. 51–59, 1982.

[24] C. M. Borror, J. B. Keats, and D. C. Montgomery, "Robustness of the time between events CUSUM," *International Journal of Production Research*, vol. 41, no. 15, pp. 3435–3444, 2003.

[25] M. B. C. Khoo, "A moving average control chart for monitoring the fraction non-conforming," *Quality and Reliability Engineering International*, vol. 20, no. 6, pp. 617–635, 2004.

[26] S. Sukparungsee and A. A. Novikov, "Analytical approximations for detection of a change-point in case of light-tailed distributions," *Journal of Quality Measurement and Analysis*, vol. 4, no. 2, pp. 49–56, 2008.

[27] Y. Areepong and A. A. Novikov, "Martingale approach to EWMA control chart for changes in Exponential distribution," *Journal of Quality Measurement and Analysis*, vol. 4, pp. 197–203, 2008.

[28] R. Noorossana, A. A. Fatahi, P. Dokouhaki, and M. Babakhani, "ZIB-EWMA control chart for monitoring rare health events," *Journal of Mechanics in Medicine and Biology*, vol. 11, no. 04, pp. 881–895, 2011.

[29] E. Santiago and J. Smith, "Control charts based on the exponential distribution: adapting runs rules for the tChart," *Quality Engineering*, vol. 25, no. 2, pp. 85–96, 2013.

[30] L. L. Ho and R. C. Quinino, "An attribute control chart for monitoring the variability of a process," *International Journal of Production Economics*, vol. 145, no. 1, pp. 263–267, 2013.

[31] M. Aslam, M. Azam, and C.-H. Jun, "Acceptance sampling plans for multi-stage process based on time-truncated test for Weibull distribution," *The International Journal of Advanced Manufacturing Technology*, vol. 79, no. 9–12, pp. 1779–1785, 2015.

[32] A. C. Rakitzis and P. Castagliola, "The effect of parameter estimation on the performance of one-sided Shewhart control charts for zero-inflated processes," *Communications in Statistics - Theory and Methods*, vol. 45, no. 14, pp. 4194–4214, 2016.

[33] M. Azam, M. N. Aslam, and C.-H. Jun, "An EWMA control chart for the exponential distribution using repetitive sampling plan," 2017.

[34] M. Azam, L. Ahmad, M. Aslam, and C.-H. Jun, "An attribute control chart using discriminant limits for monitoring process under the Weibull distribution," *Production Engineering*, vol. 12, no. 5, pp. 659–665, 2018.

[35] Z. Wu and J. Jiao, "A control chart for monitoring process mean based on attribute inspection," *International Journal of Production Research*, vol. 46, no. 15, pp. 4331–4347, 2008.

[36] H. Kim and S. Lee, "On the VSI CUSUM chart for count processes and its implementation with R package attrCUSUM," *Industrial Engineering & Management Systems*, vol. 17, no. 1, pp. 91–101, 2018.

[37] K. Mabude, J. C. Malela-Majika, M. Aslam, Z. L. Chong, and S. C. Shongwe, "Distribution-free composite Shewhart-GWMA Mann-Whitney charts for monitoring the process location," *Quality and Reliability Engineering International*, 2020.

[38] M. T. Shah, M. Azam, M. Aslam, and U. Sherazi, "Time between events control charts for gamma distribution," *Quality and Reliability Engineering International*, vol. 37, no. 2, pp. 785–803, 2020.

[39] U. Shaheen, M. Azam, and M. Aslam, "A control chart for monitoring the lognormal process variation using repetitive sampling," *Quality and Reliability Engineering International*, vol. 36, no. 3, pp. 1028–1047, 2020.

[40] M. Aslam, P. Jeyadurga, S. Balamurali, M. Azam, and A. AL-Marshadi, "Economic determination of modified multiple dependent state sampling plan under some lifetime distributions," *Journal of Mathematics*, vol. 2021, Article ID 7470196, , 2021.

[41] R. H. Berk, N. L. Johnson, S. Kotz, and N. Balakrishnan, "Continuous univariate distributions," *Technometrics*, vol. 2, p. 752, 1996.

[42] C. J. P. Bélisle, "Convergence theorems for a class of simulated annealing algorithms on $\mathbb{R}^d$," *Journal of Applied Probability*, vol. 29, no. 4, pp. 885–895, 1992.

[43] A. Henningsen and O. Toomet, "Maxlik: a package for maximum likelihood estimation in R," *Computational Statistics*, vol. 26, pp. 443–458, 2011.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 8897111, 8 pages
https://doi.org/10.1155/2021/8897111

Research Article

Analytical Investigation of Magnetic Field on Unsteady Boundary Layer Stagnation Point Flow of Water-Based Graphene Oxide-Water and Graphene Oxide-Ethylene Glycol Nanofluid over a Stretching Surface

Ali Rehman and Zabidin Salleh

Department of Mathematics, Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Kuala Nerus 21030, Terengganu, Malaysia

Correspondence should be addressed to Zabidin Salleh; zabidin@umt.edu.my

Received 24 September 2020; Revised 3 February 2021; Accepted 15 March 2021; Published 23 March 2021

Academic Editor: Hussein Abulkasim

This study explains the effect of magnetic field of the stagnation point flow of a water-based nanofluid graphene oxide-water (GO-W) and graphene oxide-ethylene glycol (GO-EG). Heat transfer analyses are discussed by converting the given partial differential equation into a nonlinear ordinary differential equation using the similarity transformation and solved using an approximate analytical method, namely, the optimal homotopy analysis method (OHAM), to obtain an approximate analytical solution of the nonlinear problem that analyzes the problem. The BVPh 2.0 package function of Mathematica is used to obtain the numerical results. The results of important parameters such as the magnetic field parameter, unsteady parameter, stretching parameter, Prandtl number, Eckert number, and kinematic parameter for both velocity and temperature profiles are plotted and discussed. The convergence control parameter of the approximate analytical method is obtained up to the 25th iteration using the BVPh 2.0 package. The skin friction coefficient and Nusselt number are explained in tabular form.

1. Introduction

In the history of fluid mechanics, the derivation of the boundary layer equation and its solution using similarity transformation is an important area for researchers. With the help of boundary layer theory, both Newtonian and non-Newtonian fluids can be modeled. The results obtained using the boundary layer theory are more similar to the experimental results. The industrial application of non-Newtonian fluids is higher than that of Newtonian fluids because of the use of non-Newtonian fluids in petroleum drilling polymer engineering and manufacturing of food and paper. Recently, the more active area of graphene oxide as a nanofluid is a rotating disk model problem. In a hexagonal lattice, graphene is a single layer of carbon atoms. Because graphene is a single layer of carbon and graphene oxide is a two-dimensional material, compared with graphene, researchers are interested in graphene oxide. In 1859, graphene oxide was introduced by Sir Second Baronent Benjamin Colline using the chemical method, Brodie, through the oxidation of bulk graphite with potassium chlorate and nitric acid. Lerf and Klinowsky used a modified Hummers' method to study the natural structure of graphene oxide. Graphene oxide has some key applications in engineering and industry, such as computer storage systems, turbomachinery hard disk jet motors, and turbine systems. It is known that the thermal conductivity of the base fluid is lower than that of the nanofluid, and to increase the heat transfer ratio, a different method is used; for example, the reduced heat transfer time and heat exchanger size can be minimized. It is known that the thermal conductivity of ethylene glycol is greater than that of the base fluid; therefore, ethylene glycol is used as a cooling fluid and antifreezing agent.

Many nonlinear relations were observed for the stress and the rate of strain for non-Newtonian fluids. It is difficult to express all the properties of a non-Newtonian fluid in a single equation. The flow owing to the stretching sheet of the boundary layer of non-Newtonian fluids has important applications in several manufacturing processes, such as molten polymers through a slit die for the production of plastic sheets, hot rolling wire and fiber coating, processing of food stuffs, metal spinning glass fiber production, and paper production. In mechanical applications, non-Newtonian fluid models play an important role. Numerous materials have incredible significance in ventures; for example, liquid plastic, nourishment items, divider paint, oils, lubricant oils, and penetrating mud have non-Newtonian fluid conduct. According to scientists, non-Newtonian liquids are the most pertinent to exploration. One of the precarious subclasses is Sisko fluid, which has more centrality as an event in many designing activities. The common sense application has delivered an enthusiasm for examining the feasibility of differential conditions overseeing streams of non-Newtonian fluids, which have various uses in the field of design, applied mathematics, and computer engineering. Munir et al. [1] used an extended sheet to study the bidirectional stream of the Sisko fluid. Olanrewaju et al. [2] used a level plate to study the Sisko precarious free convective liquid stream in the presence of heat exchange. Khan et al. [3] considered the Sisko fluid, which is consistent with the heat movement in an annular pipe. Khan et al. [4] used an extended surface to study the Sisko limit-layer liquid course. Patel et al. [5] studied a laminar Sisko fluid limit stream. Darji et al. [6] studied the Sisko fluid precarious common convective limit-layer stream. Siddiqui et al. [7] proposed a Sisko fluid film on a vertical belt for a waste framework.

Khan et al. [8] used an annular pipe to study the Sisko fluid. Sar et al. [9] reported the limit-layer conditions of a Sisko fluid. Marinca et al. [10] examined Maxwell fluid courses using permeable extending plates. Moallemi et al. [11] discovered a definite answer for the Sisko fluid stream in a pipe. Hayat et al. [12] investigated Maxwell nanofluids. Malik et al. [13] investigated an Erying-Powell nanofluid through an extension of the blended convective stream of magnetohydrodynamics (MHD). Nadeem et al. [14] investigated Maxwell fluid transfer through a vertical extending surface in the event of nanoparticles. Raju et al. [15] investigated an MHD nano non-Newtonian fluid through a cone with free convective heat and mass exchange. Rokni et al. [16] investigated a nanoliquid stream with heat exchange through plates. Nadeem et al. [13] investigated the numerical arrangement of a nanofluid stream on an extended sheet. Shehzad et al. [17] investigated the Jaffrey nanofluid MHD stream in the event of convective limit situations. Sheikholeslami et al. [18] used a nanofluid stream with an attractive field in the presence of heat exchange. Mahmoodi et al. [19] investigated a nanofluid stream with heat exchange for cooling applications. Recently, Shah et al. [20, 21] investigated nanofluid streams with the effects of heat radiation and lobby current with a turning framework. The most recent theoretical study and investigation of Sheikholeslami on nanofluids utilizing disparate marvels with current application, assets, and properties with the use of different methodologies can be considered in [22–25]. Rehman et al. [26] studied analytically the Marangoni convection of a thin-film spray of CNT nanofluids using a stretching cylinder. In this study, the effect of a magnetic field on the unsteady boundary layer stagnation point flow of a water-based nanofluid graphene oxide-water (GO-W) and graphene oxide-ethylene glycol (GO-EG) with stretching surfaces is explained. Rehman et al. [27] studied analytically a thin-film unsteady nanofluid using a stretching sheet. Gul et al. [28] discussed the performance of graphene oxide nanofluids flowing in an upright channel through a permeable medium. Gul et al. [29] discussed the impact of Marangoni and thermal radiation convection on a graphene-oxide-water-based and ethylene-glycol-based nanofluid. Next, Gul et al. [30] discussed integer and noninteger order studies of GO-W and GO-EG nanofluids. Gul et al. [31] used rotating disks to study the stable dispersion of a graphene oxide nanofluid analytically.

In this paper, we will explain the approximate analytical solution of the unsteady boundary layer stagnation point flow of water-based GO-W and GO-EG nanofluids with stretching surfaces in the presence of a magnetic field, and variable thermal conductivity is obtained. The given partial differential equation is converted to a nonlinear ordinary differential equation using the similarity transformation and solved using the approximate analytical method (OHAM). Liao et al. [32] used this method for the first time and showed that the convergence control parameter of this method is faster than the approximate solution of the nonlinear problem as compared with other analytical methods. Using this method, we found a series solution of the velocity and temperature profiles. The impacts of different parameters on both the velocity and temperature profiles were plotted and discussed.

2. Mathematical Formulation

Let us assume an unsteady two-dimensional incompressible flow of nanofluids over a stretching sheet. The stretching sheet velocity in x direction is denoted by $U_\omega = bx/1 - \gamma t$, where b and γ are constants, both of which are non-negative and t represents time. B_0 represents a uniform magnetic field, which was assumed from the Faradays. $B(t)$ represents the magnetic field and is defined as $B(t) = B_0(1-\gamma t)^{-(1/2)}$, which depends on time t, where $T_\omega(x,t) = T_0 - T_r(bx^2/2\nu)(1-\gamma t)^{-(3/2)}$ represents the temperature distribution on the surface that changes in the x direction (Figure 1).

Under the above conditions, the boundary layer equation for unsteady flow toward the stretching surface can be written as follows:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \tag{1}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = U_s\frac{dU_s}{dx} + \nu\frac{\partial^2 u}{\partial y^2} - \frac{\sigma B_0^2}{\rho_{nf}}. \tag{2}$$

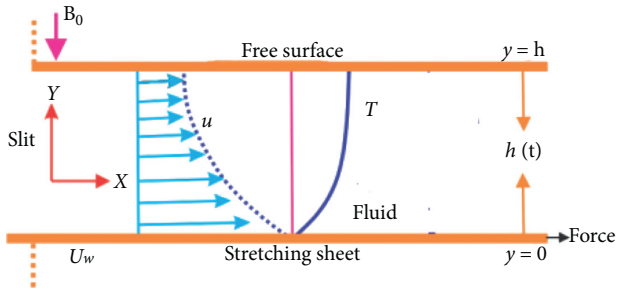


Figure 1: Geometry of the problem.

The temperature distribution in the flow field is as follows:

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial t} + v\frac{\partial T}{\partial t} = \frac{k}{\rho C_p}\frac{\partial^2 T}{\partial y^2} + \frac{\mu_s}{(\rho C_p)_f}\left(\frac{\partial u}{\partial y}\right)^2. \tag{3}$$

The temperature is denoted by T, k represents the thermal conductivity, ρ is the fluid density, C_p is the specific heat, and the dynamic viscosity of the nanofluid is denoted by μ_s. u and v are the velocity components along the x and y directions, respectively, x denotes the distance along the sheet, y represents the distance perpendicular to the sheet, the velocity of the stagnation point is denoted by $U_s = ax$ with $a > 0$, where ν represents the kinematic viscosity of the fluid, the magnetic field is denoted by M and is defined as $M = \sigma B_0^2/\rho_{nf}$, the unsteady parameter is denoted by S and is defined as $S = \gamma/b$, Eckert number is defined as $\mathrm{Ec} = U_\varpi^2/C_p(T_\varpi - T_0)$, and $A = b/c$ is the stretching parameter.

The boundary condition for velocity component is given as follows:

$$u = U_w, v = 0 \text{ at } y = 0;\ u \longrightarrow U_s \text{ as } y \longrightarrow \infty. \tag{4}$$

The stretching velocity of the sheet is denoted by $U_w = cx$, where c is the stretching constant. The boundary conditions for the temperature distribution were as follows:

$$T = T_w \text{ at } y = 0;\ T \longrightarrow T_\infty \text{ as } y \longrightarrow \infty. \tag{5}$$

In equation (5), T_w represents the temperature at the sheet and T_∞ represents the free stream temperature, which is assumed to be constant. The stream function is defined as follows:

$$u = \frac{\partial \psi}{\partial y}, v = -\frac{\partial \psi}{\partial x}. \tag{6}$$

The similarity transformation is defined as follows:

$$\psi = \sqrt{\frac{a\nu_f}{1-\alpha t}}x f(\eta),\ \eta = y\sqrt{\frac{a}{\nu_f(1-\alpha t)}}, \text{ and}$$
$$\theta(\eta) = \frac{T - T_\infty}{T_w - T_\infty}. \tag{7}$$

The similarity transformation from equation (7) is used in equations (2) and (3) to convert the PDE into a nonlinear ordinary differential equation.

$$f''' - (1-\phi)^{2.5}\left((1-\phi) + \phi\frac{\rho_s}{\rho_f}\right)f''f - f'^2 - s(1-\phi)^{2.5}M\left(\frac{\eta}{2}f'' + f'\right) = 0, \tag{8}$$

with boundary condition $f(\eta) = 0$ and $f'(\eta) = c/a$ at $\eta = 0$, and

$$\frac{k_{nf}}{k_f}\theta'' - \mathrm{Pr}\left((1-\phi) + \phi\frac{(\rho C_p)_s}{(\rho C_p)_f}\right)(\theta' f) - \frac{s}{2}\mathrm{Ec}(\eta\theta') = 0, \tag{9}$$

with boundary condition $\theta(\eta) = 1$ at $\eta = 0; \theta(\eta) \longrightarrow 0$ as $\eta \longrightarrow \infty$.

The skin friction coefficient C_{nf} is defined as $C_{nf} = 2\tau_\varpi/\rho U_\varpi^2$, and the local Nusselt number $Nu_x = (q_\varpi/k(T_\varpi - T_0))x$.

3. Method of Solution

The nondimensionless equations (8) and (9) are solved analytically by an approximate analytical method, namely, the OHAM, which is given as follows:

$$L(u(x)) + N(u(x)) + g(x) = 0,\ B(u(x)), \tag{10}$$

where L represents a linear operator, x represents an independent variable, $g(x)$ represents an unknown function, N represents a nonlinear operator, and $B(u)$ represents a boundary operator. Using this method, we first obtain a family of equations:

$$\begin{aligned} H(\phi(x), p) &= (1-p)[L, (\phi(x,p)) + g(x)] \\ &\quad - H(p)[L(\phi(x,p)) + g(x) \\ &\quad + N(\phi(x,p))] = 0, \\ B(\phi(x,p)) &= 0, \end{aligned} \tag{11}$$

where p represents an embedding parameter and lies in $[0,1]$, $H(p)$ represents a nonzero auxiliary function for $p \neq 0$ and $H(0) \neq 0$, and $\phi(x,p)$ is an unknown function. The initial estimates obtained from equations (8) and (9) are as follows:

$$\begin{aligned} f_0(\eta) &= \frac{c}{a}\eta, \\ \theta_0(\eta) &= 1, \\ L_f &= \frac{d^3 f}{d\eta^3}, \\ L_\theta &= \frac{d^2\theta}{d\eta^2}, \end{aligned} \tag{12}$$

with constant properties

$$L_f(C_1 + C_2\eta + C_3\eta^2 + C_4\eta^3) = 0, L_\theta(C_5 + C_6\eta) = 0. \tag{13}$$

For equations (7) and (8), the average squared residual error is obtained from the following equation:

$$\varepsilon_m^f = \frac{1}{n+1}\sum_{j=1}^{n}\left[K_f\left(\sum_{j=1}^{n} f(\eta)_{\eta=j\delta\eta}\right)\right],$$
$$\varepsilon_m^\theta = \frac{1}{n+1}\sum_{j=1}^{n}\left[K_\theta\left(\sum_{j=1}^{n} f(\eta)_{\eta=j\delta\eta}, \sum_{j=1}^{n}\theta(\eta)_{\eta=j\delta\eta}\right)\right], \quad (14)$$
$$\varepsilon_m^t = \varepsilon_m^f + \varepsilon_m^\theta.$$

4. Result and Discussion

This section explains the consequences of different nondimensional parameters such as kinematic viscosity, magnetic field, stretching parameter, Prandtl, Eckert number, and unsteady parameter on velocity and temperature distribution ($\nu, M, A, \mathrm{Pr}, \mathrm{Ec}$, and S). The thermophysical properties of the base fluid and nanofluid are presented in Tables 1 and 2, respectively. The influence of skin friction and Nusselt number is presented for both GO-W and GO-EG in Tables 3 and 4. Table 3 shows that the skin friction coefficient is a decreasing function of the magnetic parameter and unsteady parameter; that is, by increasing the value of the magnetic parameter and the unsteady parameter, the skin friction coefficient decreases in both GO-W and GO-EG. Table 4 shows that the Nusselt number coefficient is a decreasing function of the Prandtl number and Eckert number; that is, higher values of the Prandtl number and Eckert number decrease the Nusselt number coefficient in both GO-W and GO-EG. Tables 5 and 6 show the convergence of the given approximate analytical method for both GO-W and GO-EG nanofluids up to the 25th iteration, and it is observed from Tables 5 and 6 that with an increase in the number of iterations, a strong convergence is attained. Figure 2 shows the influence of the magnetic field parameters on the velocity profile. The relationship between $f'(\eta)$ and M is inversely related. The velocity profile is a decreasing function of the magnetic field parameter, that is, a higher value of the magnetic field parameter decreases the velocity profile. With the increase in magnetic field resistance, certain types of forces known as Lorentz force are produced, which oppose the motion of the fluid particles in both GO-EG and GO-W; therefore, the velocity profile decreases, as presented in Figure 2. Figure 3 shows the influence of the unsteady parameters on the velocity profile. The values of $f'(\eta)$ and M were inversely related. The velocity profile is a decreasing function of the unsteady parameter; that is, a higher value of the unsteady parameter decreases the velocity profile. By increasing the unsteady parameter, the motion of the fluid particles in both GO-EG and GO-W decreases, resulting in a decrease in the velocity profile, as shown in Figure 3. The influence of the stretching parameter on the velocity profile is shown in Figure 4. The relation between the velocity field $f'(\eta)$ and the stretching parameter A is direct. The velocity is the increasing function of the stretching parameter; that is, the increasing values of the stretching parameter A increase

Table 1: Thermo physical properties.

Physical properties	Thermal conduct (kW/mK)	Specific heat (CpJ/kg K)	Density (ρkg/m^3)
Base fluid water	0.135	2190	753
GO-W	6500	2400	2500
GO-EG	3000	1700	1500

Table 2: Thermal conductivity values at different volume fractions.

Volume fraction φ	0.0	0.01	0.02	0.03	0.04
Knf (GO-W)	0.135	0.137	0.214	0.215	0.256
Knf (GO-EG)	0.135	0.162	0.3	0.218	0.247

Table 3: Comparison of the skin friction for the two nanofluids when $\mathrm{Pr} = 15.6$, $\nu = 0.9$, and $A = 0.1$.

S	M	$f''(0)$ GO-W $\phi = 0.01$	$f''(0)$ GO-EG $\phi = 0.01$
0.7	1	0.929705	0.39471
0.8		0.92164	0.379547
0.9		0.91147	0.37264
	2	0.9100877	0.36994
	3	0.77608	0.33724
		0.768432	0.3246
		0.75818	0.31641

Table 4: Comparison of the Nusselt number $(\mathrm{Re}_x^{-(1/2)}Nu_x)$ for the two nanofluids when $\nu = 0.5$, $M = 2$, and $A = 0.7$.

Pr	Ec	$f''(0)$ GO-W $\phi = 0.01$	$f''(0)$ GO-EG $\phi = 0.01$
3	5	0.51231	0.49077
5		0.49341	0.48237
7		0.47451	0.47397
	7	0.45614	0.45647
	9	0.43776	0.44897
		0.42795	0.43121
		0.35021	0.41346

Table 5: Convergence of the method for GO-EG when $\mathrm{Pr} = 6$, $M = 10$, $\mathrm{Ec} = 1$, $\nu = 1$, and $A = 1$.

m	ε_m^f GO-EG	ε_m^θ GO-EG
5	1.36438×10^{-1}	2.86775×10^{-1}
10	7.14094×10^{-3}	1.48738×10^{-2}
15	5.209443×10^{-7}	1.07298×10^{-4}
20	4.37298×10^{-9}	8.54131×10^{-5}
25	3.95787×10^{-11}	7.94423×10^{-6}

Table 6: Convergence of the approximate analytical method for GO-W when $\mathrm{Pr} = 6$, $M = 10$, $\mathrm{Ec} = 1$, $\nu = 1$, and $A = 1$.

m	ε_m^f GO-W	ε_m^θ GO-W
5	1.07991×10^{-1}	2.88574×10^{-1}
10	5.65266×10^{-2}	1.0759×10^{-3}
15	4.12383×10^{-3}	1.0759×10^{-5}
20	3.4616×10^{-4}	8.55721×10^{-7}
25	3.133×10^{-5}	8.006632×10^{-9}

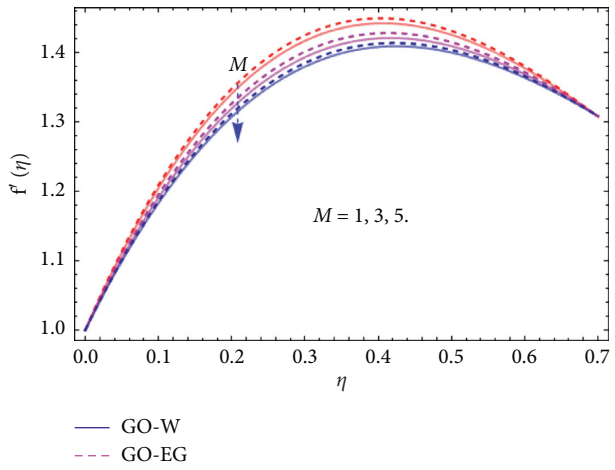


Figure 2: Effect of magnetic field on velocity profile.

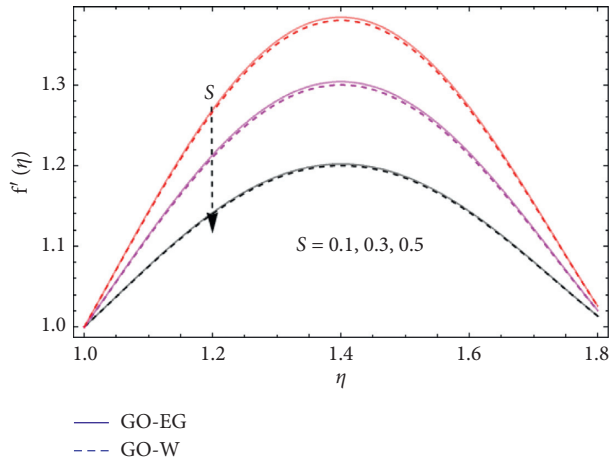


Figure 3: Effect of unsteady parameter on velocity profile.

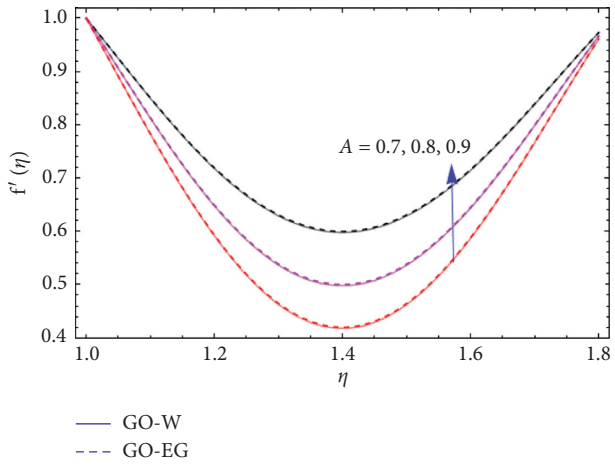


Figure 4: Effect of stretching parameter on velocity profile.

the velocity field $f'(\eta)$. Physically, by increasing the stretching parameter A, the position of the fluid particle changes; thus, the moment of the particle increases, and the velocity field is enhanced by increasing the stretching parameter A, as shown in Figure 4. The influence of the kinematic viscosity on the velocity profile is shown in Figure 5. The relationship between the velocity profile and kinematic viscosity is inverse, or the velocity profile is a decreasing function of the kinematic viscosity; that is, higher values of the kinematic viscosity α decrease the velocity profile $f'(\eta)$, as shown in Figure 5. By increasing the kinematic viscosity, viscous forces that oppose the motion of fluid particles in both GO-EG and GO-W are produced; therefore, the velocity profile decreases with increasing kinematic viscosity. Figure 6 shows the effect of the magnetic field on the temperature profile. It can be observed that by increasing the magnetic field, the temperature profile increases or the temperature profile is a decreasing function of the magnetic parameter. Such a state occurs because of the increase in resistance forces, which opposes the motion of fluid particles in both GO-EG and GO-W, resulting in friction forces that increase the temperature profile. The influence of the Prandtl number on the temperature profile is shown in Figure 7. From Figure 7, it can be observed that the temperature profile is a decreasing function of the Prandtl number; that is, a higher value of Pr decreases the

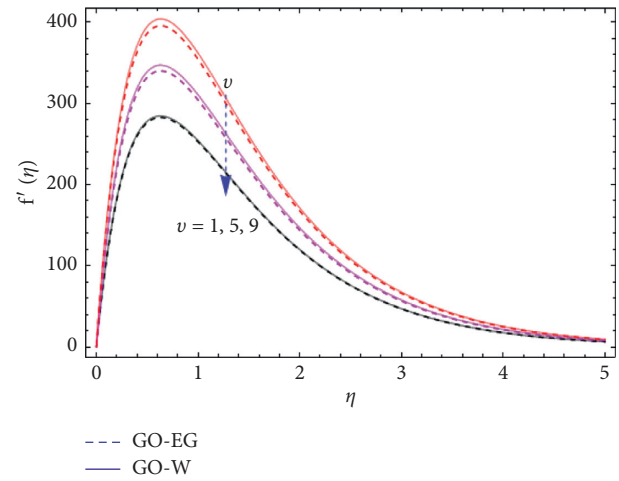


Figure 5: Effect of kinematic viscosity on velocity profile.

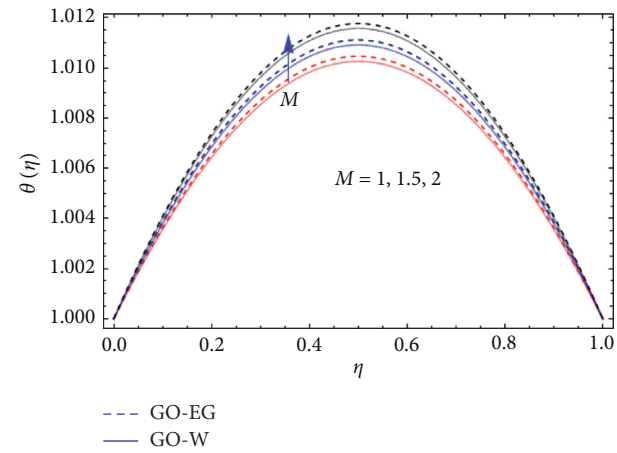


Figure 6: Effect of magnetic field on temperature profile.

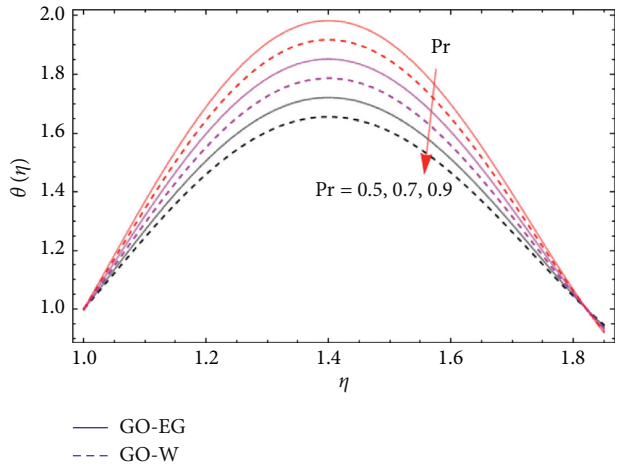


Figure 7: Effect of Prandtl number on temperature profile.

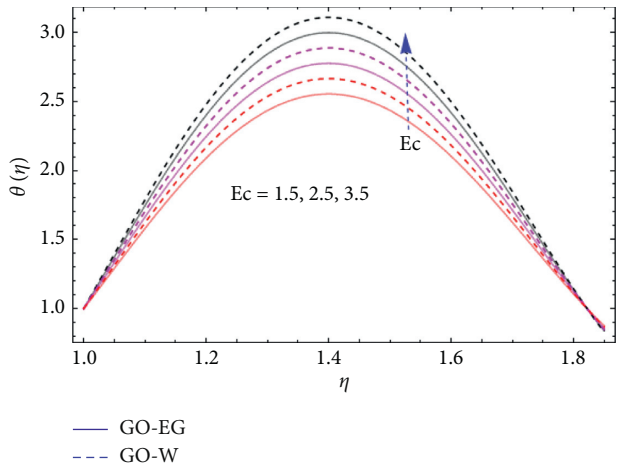


Figure 8: Effect of Eckert number on temperature profile.

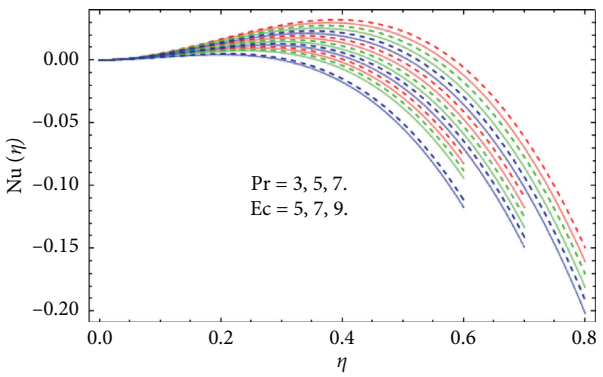


Figure 9: Comparison of the Nusselt number for the two nanofluids.

temperature profile in both GO-EG and GO-W, as presented in Figure 7. This effect is because the thickness of the momentum boundary layer is greater than that of the thermal boundary layer, and as a result, a higher value of the Prandtl number decreases the thermal boundary layer. The influence of the Eckert number on the temperature profile is shown in Figure 8. From Figure 8, it can be seen that the temperature profile is an increasing function of the Eckert number; that is, the increasing value of Eckert number increases the temperature field, and thermal diffusion is higher viscous diffusion; therefore, the improved value of Ec increases the temperature profile. The relationship between Nusselt number, Prandtl number, and Eckert number is shown in Figure 9. We see that Nusselt number is the decreasing function of Prandtl number and Eckert number in both nanofluids. Figure 10 shows the relationship between skin friction coefficient, unsteady parameter, and magnetic field parameter. We see that skin friction coefficients are the decreasing function of unsteady parameter and M (magnetic field parameter) in both nanofluids.

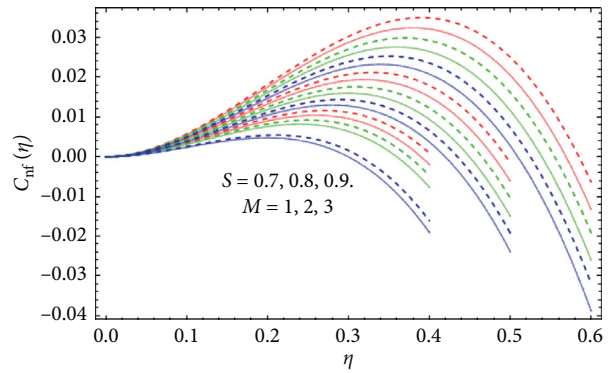


Figure 10: Comparison of the skin friction for the two nanofluids.

5. Conclusion

In this paper, we explain the approximate analytical solution of the unsteady boundary layer stagnation point flow of water-based nanofluids GO-W and GO-EG with stretching surfaces in the presence of a magnetic field, and variable thermal conductivity is obtained. The given partial differential equation is converted to a nonlinear ordinary differential equation by using the similarity transformation and solved by an approximate analytical method, namely, the OHAM. The practical application of this study is to produce cooling and heating effects, which also have significant applications in engineering and industrial machinery; the key objective of which is the circular pressure gradient and stability of centrifugal forces. The numerical results were obtained using the BVPh 2.0 package function of Mathematica. The results of important parameters such as the magnetic field parameter, unsteady parameter, stretching parameter, Prandtl number, Eckert number, and kinematic parameter were obtained. The effects of the skin friction coefficient and Nusselt number coefficient are explained in table form. The outputs obtained were as follows:

(i) The velocity profile is the decreasing function of the unsteady parameter S; that is, it decreases the velocity field by increasing the unsteady parameter

(ii) The velocity profile is an increasing function of the stretching parameter A; that is, the velocity profile

increases with an increase in the stretching parameter

(iii) The temperature profile is the decreasing function of the Prandtl number; that is, the temperature profile decreases with the increasing value of the Prandtl number

(iv) The temperature profile is an increasing function of the Eckert number; that is, the temperature profile increases with increasing Eckert number

(v) The velocity profile is a decreasing function of the magnetic field M; that is, the velocity profile decreases with an increase in the magnetic field M [33–35]

Nomenclature

S: Unsteady parameter
C_p: Specific heat
μ_{nf}: Dynamic viscosity of the nanofluid axial directions
W_w: Stretching velocity
ρ_{nf}: Density of the nanofluids
η: Similarity variable
Ec: Eckert number
φ: Solid particle volume fraction
Q: Heat generation/absorption parameter
σ^*: Stefan–Boltzmann constant
ν_{nf}: Kinematic viscosity of the nanofluids
C_f: Skin friction
T: Local temperature
δ: Thickness of the liquid film
Pr: Prandtl number
U_w: Suction, injection velocity
T_δ: Temperature at the free surface
k^*: Mean absorption coefficient
k^*: Mean absorption coefficient
q_r: Thermal radiative heat flux
β: Nondimensional thickness of the film liquid
k_{nf}: Thermal conductivity of the nanoparticles
Nu: Nusselt number
β_{nf}: Thermal expansion coefficient of solid particles
(u, w): Velocity components.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest for the publication of this paper.

Acknowledgments

This research was partially supported by the Universiti Malaysia Terengganu under the Postgraduate Research Grant (PGRG) vote no. 55193/3. The authors would like to thank Editage (http://www.editage.com) for English language editing.

References

[1] M. Khan, S. Munawar, and S. Abbasbandy, "Steady flow and heat transfer of a Sisko fluid in annular pipe," *International Journal of Heat and Mass Transfer*, vol. 53, no. 7-8, pp. 1290–1297, 2010.

[2] M. Khan and A. Shahzad, "On boundary layer flow of a Sisko fluid over a stretching sheet," *Quaestiones Mathematicae*, vol. 36, no. 1, pp. 137–151, 2013.

[3] M. Patel, J. Patel, and M. G. Timol, "Laminar boundary layer flow of Sisko fluid," *Applications and Applied Mathematics: An International Journal*, vol. 10, no. 2, pp. 909–918, 2015.

[4] R. M. Darji and M. G. Timol, "Similarity analysis for unsteady natural convective boundary layer flow of Sisko fluid," *International Journal of Advances in Applied Mathematics and Mechanics*, vol. 13, pp. 22–36, 2014.

[5] A. M. Siddiqui, H. Ashraf, T. Haroon, and A. Walait, "On the analytic solution for the steady drainage of magnetohydrodynamic (MHD) Sisko fluid film down a vertical belt," *International Journal of Ayurveda & Alternative Medicine*, vol. 8, pp. 465–480, 2017.

[6] M. Khan, Q. Abbas, and K. Duru, "Magnetohydrodynamic flow of a Sisko fluid in annular pipe: a numerical study," *International Journal for Numerical Methods in Fluids*, vol. 62, pp. 1169–1180, 2010.

[7] G. Sar, M. Pakdemirli, T. Hayat, and Y. Aksoy, "New perturbation iteration solutions for Fredholm and Volterra integral equations," *Journal of Applied Mathematics*, vol. 2013, Article ID 682537, 9 pages, 2013.

[8] V. Marinca, R.-D. Ene, B. Marinca, and R. Negrea, "Different approximations to the solution of upper-convected Maxwell fluid over a porous stretching plate," *Abstract and Applied Analysis*, vol. 2014, Article ID 139314, 12 pages, 2014.

[9] N. Moallemi, I. Shafieenejad, and A. B. Novinzadeh, "Exact solutions for flow of a Sisko fluid in pipe," *Bulletin of the Iranian Mathematical Society*, vol. 37, pp. 49–60, 2011.

[10] S.-J. Liao, "An explicit, totally analytic approximate solution for Blasius' viscous flow problems," *International Journal of Non-Linear Mechanics*, vol. 34, no. 4, pp. 759–778, 1999.

[11] S.-J. Liao, "On the analytic solution of magnetohydrodynamic flows of non-Newtonian fluids over a stretching sheet," *Journal of Fluid Mechanics*, vol. 488, pp. 189–212, 2003.

[12] M. Y. Malik, K. Imad, H. Arif, and T. Salahuddin, "Mixed convection flow of MHD Eyring-Powell nanofluid over a stretching sheet," *Neural Computing & Applications*, vol. 5, pp. 117-118, 2015.

[13] S. Nadeem, R. U. Haq, and Z. H. Khan, "Numerical solution of non-Newtonian nanofluid flow over a stretching sheet," *Applied Nanoscience*, vol. 4, no. 5, pp. 625–631, 2014.

[14] S. Nadeem, R. U. Haq, and Z. H. Khan, "Numerical study of MHD boundary layer flow of a Maxwell fluid past a stretching sheet in the presence of nanoparticles," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 45, no. 1, pp. 121–126, 2014.

[15] C. S. K. Raju, P. Sanjeevi, M. C. Raju, S. M. Ibrahim, G. Lorenzini, and E. Lorenzini, "The flow of magnetohydrodynamic Maxwell nanofluid over a cylinder with Cattaneo-Christov heat flux model," *Continuum Mechanics and Thermodynamics*, vol. 29, no. 6, pp. 1347–1363, 2017.

[16] H. B. Rokni, D. M. Alsaad, and P. Valipour, "Electrohydrodynamic nanofluid flow and heat transfer between two plates," *Journal of Molecular Liquids*, vol. 216, pp. 583–589, 2016.

[17] S. A. Shehzad, T. Hayat, and A. Alsaedi, "MHD flow of Jeffrey nanofluid with convective boundary conditions," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 37, no. 3, pp. 873–883, 2015.

[18] M. Sheikholeslami, M. Hatami, and D. D. Ganji, "Nanofluid flow and heat transfer in a rotating system in the presence of a magnetic field," *Journal of Molecular Liquids*, vol. 190, pp. 112–120, 2014.

[19] M. Mahmoodi and S. H. Kandelousi, "Kerosene-alumina nanofluid flow and heat transfer for cooling application," *Journal of Central South University*, vol. 23, no. 4, pp. 983–990, 2016.

[20] Z. Shah, T. Gul, A. M. Khan, I. Ali, and S. Islam, "Effects of hall current on steady three dimensional non-Newtonian nanofluid in a rotating frame with Brownian motion and thermophoresis effects," *Journal of Engineering Technology*, vol. 6, pp. 280–296, 2017.

[21] M. Sheikholeslami, M. Shamlooei, and R. Moradi, "Fe_3O_4-ethylene glycol nanofluid forced convection inside a porous enclosure in existence of Coulomb force," *Journal of Molecular Liquids*, vol. 249, pp. 429–437, 2018.

[22] M. Sheikholeslami, "Numerical investigation of nanofluid free convection under the influence of electric field in a porous enclosure," *Journal of Molecular Liquids*, vol. 249, pp. 1212–1221, 2018.

[23] M. Sheikholeslami and H. B. Rokni, "Numerical simulation for impact of coulomb force on nanofluid heat transfer in a porous enclosure in presence of thermal radiation," *International Journal of Heat and Mass Transfer*, vol. 118, pp. 823–831, 2018.

[24] Z. Shah, E. Bonyah, S. Islam, W. Khan, and M. Ishaq, "Radiative MHD thin film flow of Williamson fluid over an unsteady permeable stretching sheet," *Applied Nanoscience*, vol. 4, no. 10, pp. 825–830, 2018.

[25] A. Munir, S. Towers, A. Shahzad, and M. Khan, "Mass media and the contagion of fear: the case of Ebola in America," *PLoS One*, vol. 10, no. 6, pp. 1–13, 2015.

[26] A. Rehman, T. Gul, and Z. Salleh, "Effect of the Marangoni convection in the unsteady thin fico spray of CNT nanofluids," *Processes*, vol. 7, no. 6, p. 392, 2019.

[27] A. Rehman, Z. Salleh, T. Gul, and Z. Zaheer, "The impact of viscous dissipation on the thin film unsteady flow of GO-EG/GO-W nanofluids," *Mathematics*, vol. 7, no. 7, p. 653, 2019.

[28] T. Gul, M. Z. Ullah, A. K. Alzahrani, and I. S. Amiri, "Thermal performance of the graphene oxide nanofluids flow in an upright channel through a permeable medium," *IEEE Access*, vol. 7, pp. 102345–102355, 2019.

[29] T. Gul, W. Noman, M. Sohail, and M. A. Khan, "Impact of the Marangoni and thermal radiation convection on the graphene-oxide-water-based and ethylene-glycol-based nanofluids," *Advances in Mechanical Engineering*, vol. 11, no. 6, 2019.

[30] T. Gul, H. Anwar, M. A. Khan, I. Khan, and P. Kumam, "Integer and non-integer order study of the GO-W/GO-EG nanofluids flow by means of Marangoni convection," *Symmetry*, vol. 11, no. 5, p. 640, 2019.

[31] T. Gul and K. Firdous, "The experimental study to examine the stable dispersion of the graphene nanoparticles and to look at the GO-H_2O nanofluid flow between two rotating disks," *Applied Nanoscience*, vol. 8, no. 7, pp. 1711–1727, 2018.

[32] S. Liao, "An optimal homotopy-analysis approach for strongly nonlinear differential equations," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 8, p. 2003, 2010.

[33] T. Hayat, T. Muhammad, S. A. Shehzad, and A. Alsaedi, "Three-dimensional boundary layer flow of Maxwell nanofluid: mathematical model," *Applied Mathematics and Mechanics*, vol. 36, no. 6, pp. 747–762, 2015.

[34] Z. Shah, T. Gul, S. Islam et al., "Three dimensional third grade nanofluid flow in a rotating system between parallel plates with Brownian motion and thermophoresis effects," *Results in Physics*, vol. 10, pp. 36–45, 2018.

[35] Z. Shah, S. Islam, T. Gul, E. Bonyah, and M. Altaf Khan, "The electrical MHD and Hall current impact on micropolar nanofluid flow between rotating parallel plates," *Results in Physics*, vol. 9, pp. 1201–1214, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6671403, 12 pages
https://doi.org/10.1155/2021/6671403

Research Article

On-Line Interpretation and Real-Time Diagnosis of Rocket's Single Equipment

Erbao Xu,[1] **Yan Li**,[1] **Lining Peng**,[1] **Yuxi Li**,[2] **and Mingshun Yang**[1]

[1]*School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an 710048, China*
[2]*Xi'an Modern Control Technology Research Institute, Xi'an 710048, China*

Correspondence should be addressed to Erbao Xu; baojams@126.com

Received 8 December 2020; Revised 9 February 2021; Accepted 24 February 2021; Published 13 March 2021

Academic Editor: Bekir Sahin

The work state of a launch vehicle is generally interpreted automatically on software. However, the sheer number of target parameters makes it difficult to realize real-time interpretation, and abnormal interpretation result does not necessarily mean that the vehicle is in abnormal state. This paper introduces the edge computing to achieve on-line interpretation and real-time diagnosis of a single launch vehicle. Firstly, the parameters to be interpreted were subjected to thresholding, leaving only those with high interpretation value. Next, the interpretation server layer of the real-time diagnosis model was built based on the attribute and value reduction algorithm of variable precision rough set (VPRS). Moreover, the higher-grade criteria were written in criterion modeling language (CML) and used to interpret the various higher-grade interpretation data pushed by the edge layer in real time. On this basis, the outputs of the edge layer and interpretation server layer were integrated to achieve the real-time diagnosis of single vehicle faults. Finally, the proposed model was proved feasible through the application in a launch vehicle.

1. Introduction

During the "Twelfth Five-Year Plan" and "Thirteenth Five-Year Plan" period, China's launch vehicle and weapon models entered a high-density launch period, and the research model entered a critical period of development. Compared with the past, the launch task is heavy and the test cycle is significantly shortened. In the process of test and launch, the measurement data interpretation is an important part, which is directly related to the function and performance judgment and evaluation of each system of the aircraft [1, 2].

In the current telemetry and other parameter interpretations of launch vehicle, the parameter data received by the system in real time are generally saved. After the test is finished, all the data of all parameters stored in the disk are processed uniformly, and then, in the interpretation server, the automatic interpretation software is used for interpretation, and generally a fixed threshold comparison is used. The statistical results show that the time allocation of the launch vehicle to perform a total inspection test at the shooting range is 30% for test preparation and 40% for power-on test, and the remaining 30% is for postprocessing and interpretation [1].

However, CZ3, cz5, cz7, and other new rocket models adopt a high bit rate measurement system, with many fast-changing parameters and large volume of data, resulting in heavy interpretation workload, long interpretation time, and increased risk of misjudgment. Especially for attitude control and guidance profession, the workload of interpretation is larger. Therefore, the amount of parameter data to be interpreted is extremely large, and the interpretation server needs to carry out complex analysis processing such as path selection for the whole package of data before automatic interpretation [2]. The hysteresis nature of interpretation results caused by after-event interpretation makes it impossible to find and locate problems in the first place and on the spot and then to grasp the real-time status of the rocket, which has seriously affected the quality and progress of the rocket range test. After the test, the state of the rocket based on the interpretation afterwards is also meaningless due to the lack of timeliness and accuracy. Therefore, only if the

time for data processing and interpretation is infinitely shortened, real-time on-line interpretation is achieved, and the speed and accuracy of condition diagnosis are improved can it meet the needs of modern rocket range testing [3].

It can be seen that the problems of the current auto-interpretation system of the launch vehicle include the following: the hysteresis nature is caused by the large amount of data as well as the use of after-event interpretation; most of the criterions are threshold criterions, which are single and lack intelligence; the computing resources of the interpretation server are mostly used for threshold comparison of large amounts of normal data parameters, as well as lack of focus and special attention on key parameters or over limit data that truly reflect the fault status of the equipment. Therefore, at present, although the auto-interpretation software is a great workload and long period by hand, it is inefficient and inaccurate. Domestic and foreign scholars have also conducted extensive research on this issue.

In Guo et al.'s research [4], for several types of parameters of the rocket range, the median filtering algorithm is used to achieve smooth filtering of the step parameters and pulse parameters. A bilateral multipoint threshold judgment method is designed to automatically and accurately identify the step parameters and realize automatic auxiliary interpretation function, but this method is limited to step parameters and pulse parameters, and it is not applicable to a large number of slowly varying parameters that are difficult to interpret; Shi et al. [5] believed that an open interpretation function compiling platform is proposed based on the research background of on-orbit fault diagnosis. From the perspective of the architecture design of the telemetry parameter interpretation platform and from the perspective of the architecture design of the telemetry parameter interpretation platform, the composition, implementation technical details, and interpretation rules of the platform are discussed, but the analysis of the telemetry parameter interpretation method is lacking. Dong et al. [6] proposed a novel method of automatic interpretation using linear prediction method and prediction method based on GM (1, 1) model to solve the problem of slow-varying telemetry parameters changing. To some extent, this method can realize the interpretation of slow variation parameter but lacks the mining of historical data of telemetry parameters. Based on the statistical characteristics of historical data, Hu et al. [7] put forward a more comprehensive content and a view calendar in view of the weak link of the existing rocket flight data comparison and achieved the same parameter history data comparison function and also achieved the slicing viewing function for the step type signal of severe jump; however, this research mainly focuses on the historical telemetry parameter data and lacks the consideration of massive real-time data interpretation and real-time diagnosis; Li et al. [8] believed that a new method of automatic interpretation of telemetry slow variation parameter based on the statistical characteristics of historical data is proposed, which can effectively identify abnormal parameters in the slow variation parameter. Compared with the traditional manual interpretation method, this method can effectively improve the efficiency and accuracy of the interpretation of telemetry parameters but still does not achieve a real real-time interpretation. Guo et al. [9] proposed an intelligent fault diagnosis method for aeroengine sensors combining a deep learning algorithm (CNN) with time-frequency analysis, wherein the signal recognition problem is transformed into an image recognition problem. In this paper, a large amount of historical data is used for CNN model training, which is still after diagnosis and does not take real-time diagnosis as the research goal. Glowacz [10] described fault diagnosis method based on analysis of thermal images. It was found that $RE_{EID\text{-}3\text{-}\ CLASSES}$ of the backpropagation neural network (97.91%) was slightly lower than the result of the NN (100%). The results of the experimental analysis showed usefulness of the proposed approach. This approach can find application for protection of rotating machinery and engines.

In recent years, with the universal application of the Internet of Things technology and the gradual rise of edge computing technology, it is urgent to use these new technologies to technically reconstruct the autointerpretation system of the launch vehicle. In this paper, a "double-layer interpretation" architecture of edge-interpretation server layer is proposed, which has achieved good results in actual rocket range tests. It has realized on-line automatic interpretation and real-time diagnosis, which verifies the effectiveness and practicability of the method.

The remainder of this article is arranged as follows: Section 2 describes the concept and construction process of "double-layer interpretation" and gives its network topology diagram and software framework diagram; Section 3 describes the role and the specific data flow of edge interpretation and clarifies the significance and innovation of introducing the edge interpretation layer; Section 4 describes the work undertaken by the interpretation server layer, introduces the original CML criterion syntax and the format of higher-grade criterion, and gives the process of extracting the simplest diagnostic rules by VPRS step by step. In Section 5, the effectiveness of the proposed strategy is verified by a project example. Firstly, the construction process of the edge interpretation layer is described in detail according to Section 2. Then, the specific process of higher-grade interpretation is shown through the software interface in Section 3. Finally, a specific example is given to show the detailed steps of using VPRS to extract the simplest diagnostic rules for real-time interpretation. In Section 6, we get our final conclusion; that is, the "double-layer interpretation" strategy can realize on-line interpretation and real-time diagnosis of single equipment of launch vehicle. The block diagram of the research is shown in Figure 1.

2. On-Line Interpretation Architecture

The "double-layer interpretation" of the edge-interpretation server layer mainly includes two aspects: Firstly, at the edge, the source of data generation usually refers to the rocket or the general inspection workshop, and the existing sensors and other data acquisition equipment upgrade hardware to make it an edge computing node, giving it the ability to

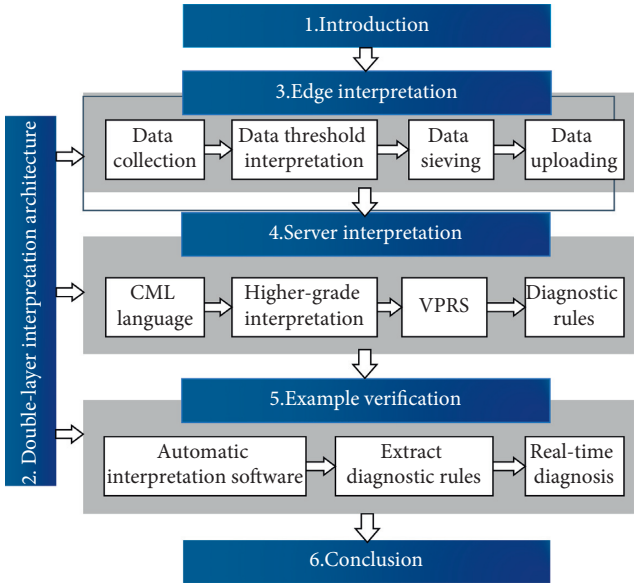


Figure 1: Block diagram of the research.

perform simple calculations, small-scale data storage, and data transmission. For devices that are not convenient to upgrade, a data collection box can be deployed separately as an edge computing node. The edge computing nodes distributed in different subsystems or data sources constitute the edge-interpretation layer. Each node collects the telemetry parameters of each corresponding equipment in real time and compares the threshold directly in the node. Most of the normal data that have not exceeded the limit will be selectively transmitted out after local aggregation within the edge node according to the set time interval, and all data that exceed or are about to exceed the limit will be specially marked and immediately transmitted out. Secondly, a new higher-grade criterion is added to the traditional interpretation server. Meanwhile, a diagnosis model is established by VPRS data mining method. The relationship between parameter overrun and equipment status is mined from the previous fault data of equipment, the importance of different interpretation parameters is analyzed quantitatively, and accurate and highly simplified diagnosis rules are deduced. Finally, for the high-value parameter interpretation data from the edge interpretation layer in real time, the higher-grade criterion is used for in-depth interpretation. If the interpretation result is abnormal, it is substituted into the diagnosis model for state diagnosis.

The "double-layer interpretation" network deployment topology is shown in Figure 2, and its software architecture is shown in Figure 3.

3. Edge Interpretation

Compared with the remote computing mode of cloud or dedicated server, edge computing refers to localized data collection, instant computing, real-time on-line diagnosis, timely response, and precise control at the data source close to the physical environment [11, 12]. Edge computing integrates the functions of computing, storage, and transmission into the equipment, so that data can be processed in a timely and effective manner near the source of the collection without being transferred to a cloud server, which will greatly improve data processing efficiency and value density and reduce cloud server concentration load of processing data [13, 14].

In the "double-layer interpretation" architecture, the edge-interpretation layer undertakes important functions such as raw data collection, data cleaning, data threshold interpretation, data sieving, and uploading. According to the distribution of data sources, edge nodes are allocated and deployed reasonably, and each edge node collects the parameters of the corresponding subsystem or stand-alone equipment. In the edge node, the threshold criterion corresponding to the acquisition parameters is stored in advance. At the same time, an over limit recognition area is added based on the fixed upper and lower limit of the threshold criterion. The range of the recognition area is slightly larger than the threshold range. The massive parameter of real-time data is first compared with the corresponding threshold in the edge node. For most normal data that are within the threshold and do not enter the over-limit identification area, they are temporarily stored in the internal buffer of the edge node, and then the maximum and minimum values of the data in this period are selected at a certain interval and sent to the interpretation server. For a few abnormal data that exceed the threshold range, they are immediately uploaded to the interpretation server. For data that enter the over-limit identification area but do not exceed the threshold for the time being, they are closely followed by measures to increase the collection frequency. The flow chart of edge-interpretation layer data is shown in Figure 4.

The main purpose of the edge-interpretation layer is to transfer the simple but heavy work of parameter threshold comparison from server to multiple edge computing nodes for distributed parallel interpretation. According to the interpretation results, more valuable or abnormal data is screened into the interpretation server. In this way, not only can the amount of data in interpretation server be greatly reduced, but also the efficiency of interpretation can be effectively improved, and real-time on-line interpretation of data can be realized.

4. Interpretation Server

At present, most of the autointerpretation systems use threshold criterion or fixed value upper and lower limit criterion [15]. When the judged parameter data exceeds the threshold, it is regarded as a parameter anomaly. This kind of criterion is mainly applied to the interpretation of a large number of slow variation parameters such as voltage, current, temperature, and pressure. The advantage of the threshold criterion is that the calculation is simple and the reading speed is fast, but it can only simply reflect whether the parameter is within the set threshold range. However, if the parameter exceeds the threshold value, it does not mean that the single equipment is faulty. The fuzzy mapping relationship between the two usually needs to be defined by an expert system or even manually, which often causes the

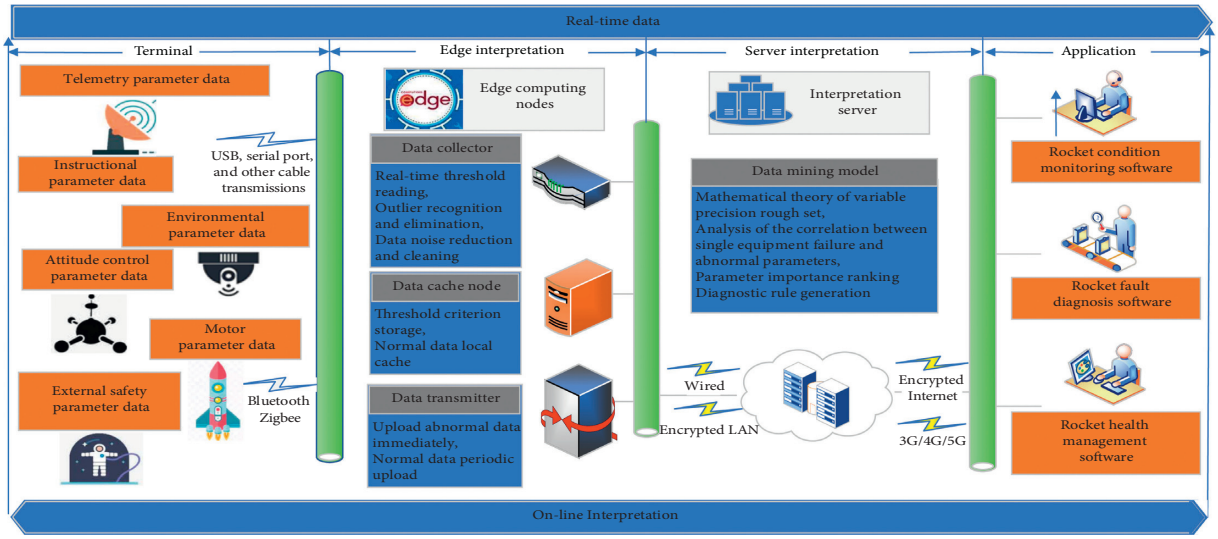


Figure 2: Network topology diagram of double-layer interpretation.

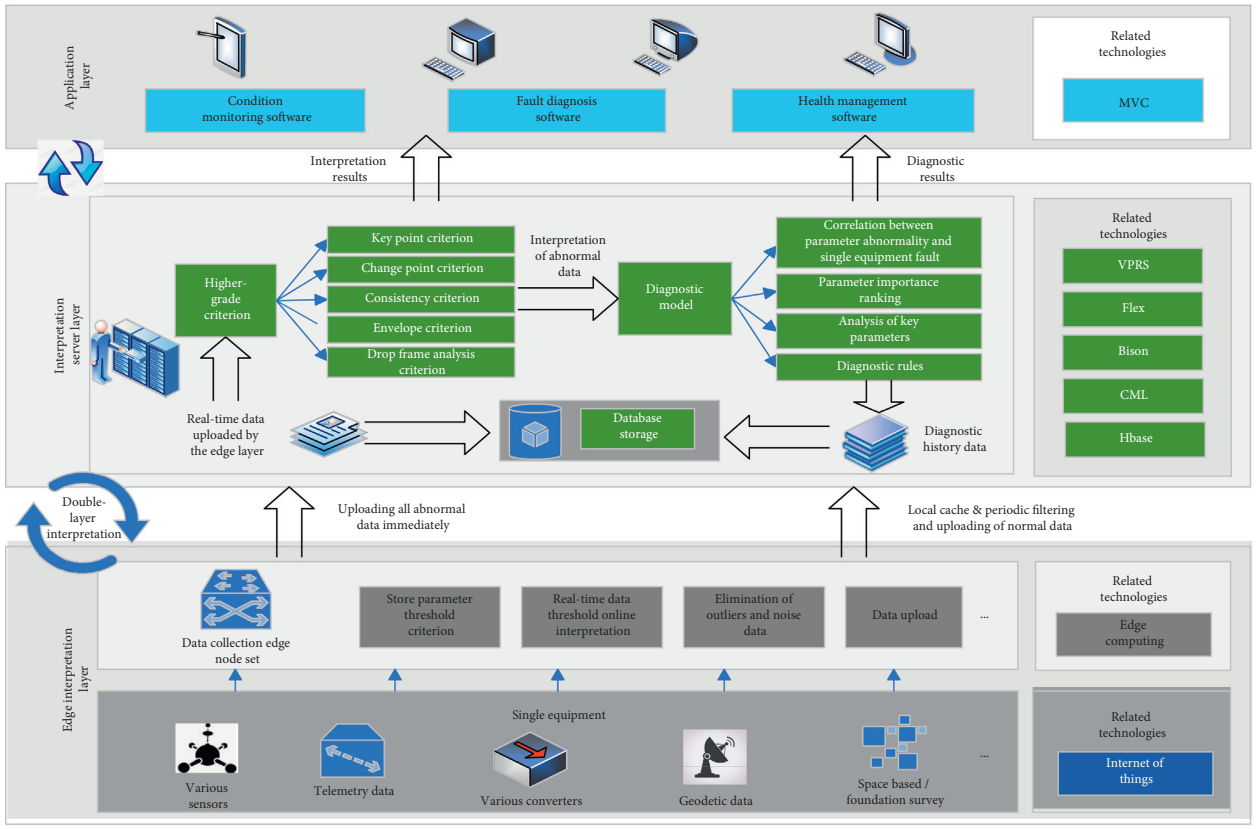


Figure 3: Architecture diagram of double-layer interpretation.

problem of "judgment is not allowed," so that the interpretation work only stays in the judgment of whether the parameter is abnormal, and loses the function of diagnosing the fault state of the equipment.

In view of this, at the interpretation server layer, a new criterion description language is first defined to describe more complex and intelligent higher-grade criteria. The parameters are selected by the edge layer threshold criterion; according to their characteristics, write corresponding higher-grade criteria, perform higher-grade interpretation and obtain the interpretation results, and store them in the HBase distributed database. Then, make full use of the historical parameter data of the single equipment's faults and establish a diagnostic model.

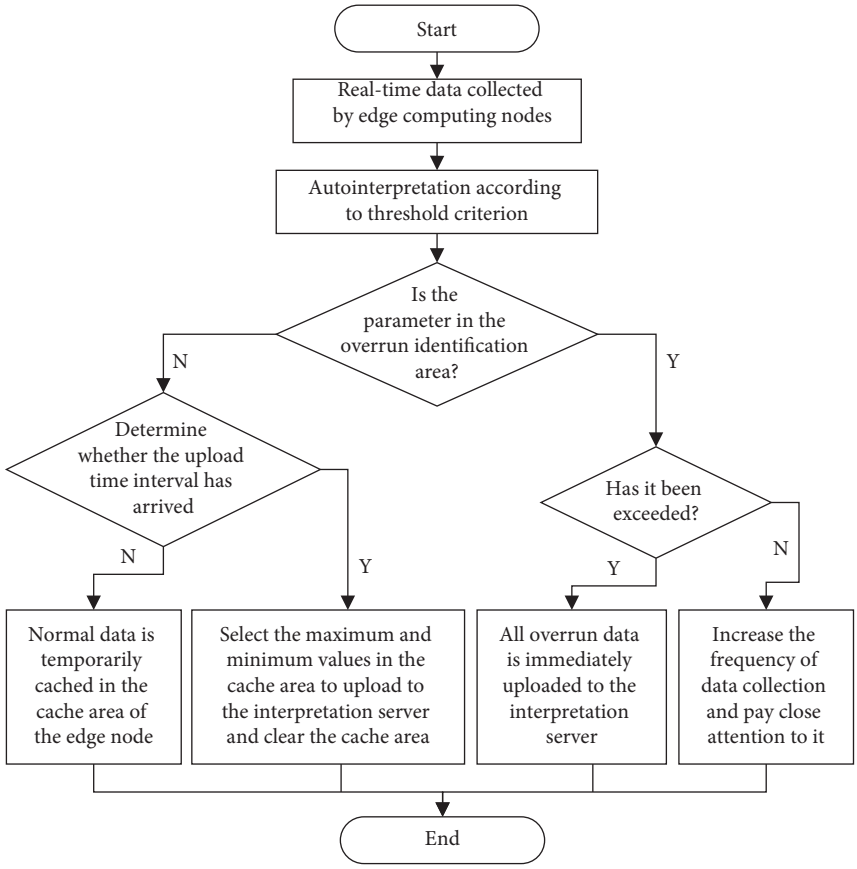


Figure 4: Flow chart of data interpretation.

There are many methods to establish the diagnostic model [10]. Considering the need to extract the most simplified diagnostic rules, the algorithm is required to have the ability of parameter reduction and value reduction. So, we adopt variable precision rough set (VPRS) theory and use its attribute dependency and importance calculation methods to quantitatively analyze the parameter interpretation results and the equipment status association relationship and degree of correlation and dig out the key parameters that affect the working status of the equipment. With the help of attribute reduction and value reduction algorithms of VPRS, diagnostic rules are extracted and finally accumulated into a diagnostic rule base. Therefore, the fault status of the device can be inferred intuitively based on the parameter interpretation results, breaking the barrier between autointerpretation and fault diagnosis and realizing the on-line interpretation and real-time status diagnosis of the single equipment.

4.1. CML Criterion Description Language. At present, most of the criteria are described by natural language. How to accurately convert the criteria to computer language is the first key technology to realize autointerpretation. In order to solve the above problems, a computer advanced language CML (criterion modeling language) containing grammar and semantics is designed before the interpretation platform to implement attitude control, guidance, power, electrical overall, and measurement system or professional criterion description, mathematical modeling, and autointerpretation, while automatically generating data interpretation analysis report. As shown in Table 1, CML includes identifier, expression, and criterion syntax.

4.1.1. Identifier. The identifier is used to identify the valid character sequence of the variables and functions. It is stipulated that the identifier can only consist of three characters: letters, numbers, and underscores, and the first character must be a letter or an underscore. In CML, there are two main types of identifiers:

(1) Keywords. 32 keywords are specified, such as diff, mp, CP, and envelope. They have specific uses and meanings and cannot be used as variable names.

(2) System predefined identifier. Some examples are the names of library functions provided by the system, such as sin and cos.

4.1.2. Expression. Expressions are the basic unit of a sentence. Expressions are used in time periods, processing formulas, criterion formulas, and errors. Commonly used expressions are shown in Table 1.

Table 1: Table of expressions in CML.

Expression name	Expression meaning
Operator expression	Expression 1 operator expression 2, operators include + − ∗ /arithmetic operator and &&;||not logical operators
Function call expression	Function name (expression 1, expression 2, . . ., expression n)," functions include system built-in mathematical functions and user-defined functions
Take parameter value expression	P (parameter code [point index]) P (Ujd)−P (Ujd[−1]) calculates the difference between the current point value and the previous point value of parameter Ujd
Take parameter time expression	T (parameter code [point index]) T(Ujd)−T(Ujd [1]) calculating the difference between the current time and the next time of parameter Ujd.
Instruction action time expression	Parameter code switch symbol the sequence of occurrence "↑" indicates ON, "↓" indicates OFF, Tqf↑3 indicates the time when the third switching state of parameter Tqf is ON
Parameter summation expression	Sum (enter expression, number expression) sum (P(Uz), 3) represents the value of the parameter Uz that is the sum of the first 3 points from the current point
Difference expression	Diff (input expression, compensation amount) diff (frameCount, 65536) calculates whether the difference between the adjacent values of frameCount is 1

4.1.3. Criterion Grammar. The basic syntax format of the criterion is shown in Figure 5, including the parameters to be judged, start time, end time, criterion formula, positive error, and negative error. In automatic interpretation, the calculation result of the criterion formula is compared with the value of the parameter to be judged and whether the difference between the two is within the range of positive and negative errors. Among them, if the start time is empty, it means interpretation starts from the first point, and if the end time is empty, it means interpretation ends at the last point.

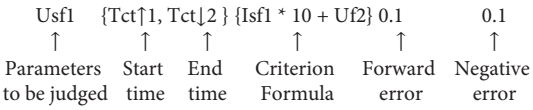


Figure 5: Diagram of criteria grammar.

4.2. Higher-Grade Criterion

4.2.1. Key Points Criterion. The key points criterion is used to interpret the value of the parameter at the specified time. In the interpretation algorithm part, several sets of keys (time, value) pairs are given. In the autointerpretation, the criterion will obtain the real data values of the parameters to be judged at several key points and make a difference with the theoretical value of the set key points to determine whether the difference is within the range of positive and negative errors.

4.2.2. Change Points Criterion. The operator CP is used to describe and interpret when the value of the parameter to be judged changes. In the interpretation algorithm part, several pairs of change points (value, time) are given. In the automatic interpretation, the criterion obtains the time when the real-time value of the parameter to be judged is equal to the value of the change point and makes a difference with the set theoretical time of the corresponding change point to determine whether the difference is within the range of positive and negative errors.

4.2.3. Consistency Criterion. The operator MP is used for description. When the interpretation is performed, the first parameter is used as the reference parameter, and the subsequent parameter values are different from the first parameter value point by point to determine whether the difference is within the allowed error range. Consistency criterion is usually used to judge those parameters that theoretically require consistent change trends. At the same time, in the diagnostic model, the interpretation results can be combined and processed to reduce the amount of model parameters and speed up the training speed.

4.2.4. Envelope Criterion. Take the data of multiple tests of the same parameter and take out the minimum and maximum values point by point and save them. All the minimum values form the lower envelope curve of the parameter, and the maximum value forms the upper envelope curve. By calculating whether the real-time point is over envelope or the degree of over envelope, abnormal situations such as outliers, data noise points, and data disturbance are identified.

4.2.5. Frame Dropping Analysis Criterion. It is specifically used for the frame count parameter to determine whether the difference between each frame count and the previous frame count is 1 and to confirm whether there is a frame loss situation. Generally, the default is full-time interpretation. For example, diff ("frameCount," 1, 65536) means that each data point is directly different from the previous data point. If the difference is less than 0, 65536 is added for calculation. If the theoretical value is 1, it means that no frame is dropped. If the difference is greater than 1, it means that there is a frame loss.

4.3. Variable Precision Rough Set. Variable precision rough set (VPRS), as a mathematical method to solve the nonlinear correspondence problem, can mine potential knowledge and rules from massive data without prior knowledge [16–18].

The core idea is to obtain the decision or classification rules for uncertain problems through knowledge reduction, while the classification ability is unchanged [19]. VPRS introduces the threshold parameter β, which indicates that the classification error rate can exist within a certain range. The general value range of β is $0.5 < \beta \leq 1$ [20]; when $\beta = 1$, VPRS is the classic rough set.

VPRS can be expressed as a quadruple $S = <U, A, V, f>$, where $U = \{y_1, y_2, \ldots, y_n\}$ is the universe, that is, by the sample object $y_i (i = 1, 2, \ldots, n)$ consisting of a finite set; $A = C \cup D$, $C \cap D = \varnothing$, $C = \{a_1, a_2, \ldots a_p\}$ is a finite set of conditional attributes, D is a set of decision attributes, V_a is the range of attribute a, f is the information function, f: $U \times A \longrightarrow V$ is a single mapping, that is, $\forall a \in A$, $y \in U$, $f(y, a) \in V_a$, and $f(y, a)$ is the information value of each attribute of each object in U [21].

By using the parameter interpretation results of a large number of single machine faults stored in HBase, all parameters that are abnormal when the fault is retrieved are retrieved to form the condition attribute set C, and the fault state of single machine is the decision attribute D.

4.3.1. Fault Dependency Calculation. The dependency of decision attribute D and condition attribute C in VPRS is defined as follows:

$$K = \gamma(C, D, \beta) = \frac{|pos(C, D, \beta)|}{|U|}, \quad (1)$$

where $pos(C, D, \beta)$ is the positive region of β.

Based on the attribute dependency calculation method of VPRS, the dependency degree between single machine fault and each interpretation parameter is analyzed quantitatively.

4.3.2. Key Parameters. In VPRS, the impact of the conditional attribute set C on the classification after removing the attribute r is

$$K(C, \{r\}, \beta) = |\gamma(C, D, \beta) - \gamma(C - \{r\}, D, \beta)|. \quad (2)$$

The impact of a single attribute r on classification is expressed as $\gamma(\{r\}, D, \beta)$; then the importance of the attribute r can be defined as

$$sig(C, \{r\}, \beta) = K(C, \{r\}, \beta) + \gamma(\{r\}, D, \beta). \quad (3)$$

The larger the value is, the more important the attribute is. It means that the abnormal parameter has a greater impact on the single machine state. With this method, the importance ranking of the interpretation parameters can be carried out, and the key parameters affecting the single machine fault state can be found [22, 23].

4.3.3. Parameter Reduction. The attribute reduction of the variable precision rough set is to reduce the condition attribute. If the dependency of a single attribute r, $\gamma(\{r\}, D, \beta)$ is equal to $\gamma(C, D, \beta)$, the condition attribute is considered as a redundant attribute.

Through this method, unnecessary parameters can be removed from many interpretation parameters, and then diagnosis rules can be extracted from the simplified parameters, which can better serve the real-time state diagnosis of a single machine.

4.3.4. Value Reduction. After attribute reduction, the decision table forms a decision rule for each sample. However, these rules are not the most streamlined and can be simplified by the following methods: For each rule in the decision rule set, if any attribute in the rule is removed and the rule does not conflict with other rules in the set, this attribute is deleted from the rule.

After value reduction, all diagnostic rules do not contain redundant condition attributes, which means that the most concise diagnostic rules are obtained from the diagnostic model.

5. Example Verification

In order to verify the effectiveness of the proposed "double-layer interpretation" architecture, taking the Long March series XX launch vehicle in Figure 6 as an example, the front end of the instrument cabin is selected as the single equipment for the research. This single machine is mainly responsible for testing the temperature and overload of the instrument cabin and judging the status of the instrument cabin. The main parameters collected by this single machine include the following: front cabin thermocouple verification (c_1), front cabin inner wall temperature (c_2), front cabin noise (c_3), front cabin end handle root inner wall temperature (c_4), front cabin end handle root cavity temperature (c_5), temperature of the inner wall near the front handle of the front cabin (c_6), front cabin vibration frequency (c_7), high accuracy overload of the front cabin (c_8), pitching attitude angle (c_9), rolling attitude angle (c_{10}), front end cold end compensation (c_{11}), tip handle head pressure (c_{12}), and tip handle tail pressure (c_{13}).

5.1. Implementation of Edge-Interpretation Layer. According to the requirements of the edge-interpretation layer in Section 2, upgrade the function of the data receiver on the arrow telemetry subsystem and add data storage, threshold comparison calculation, and data transmission hardware modules. When conducting sum-check of simulation flight at the range, the bus pushes the data of 13 monitoring parameters to the edge nodes at a speed of 10,000-bit rate. After the threshold interpretation, the local cache and screening are performed according to the flow of Figure 4 and then uploaded to the interpretation server at an average speed of 100 bits. Therefore, the total amount of data transmitted is reduced to 1%.

From the HBase database, search the fault history data of the single head end handle single machine, and select the eight sets of abnormal data shown in Table 2.

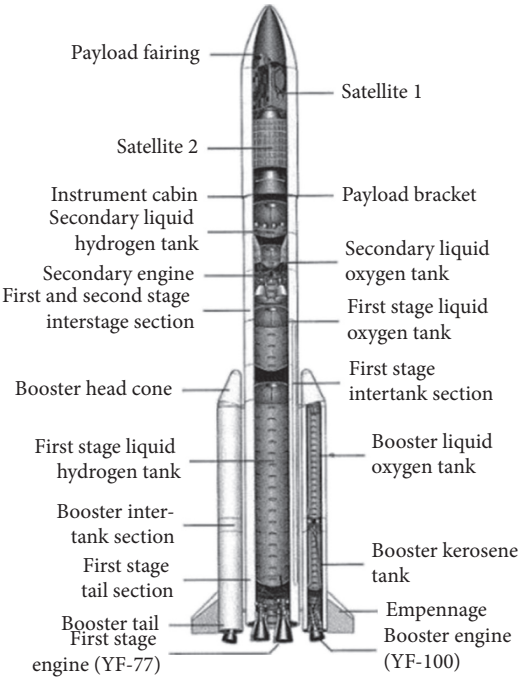


Figure 6: Long March XX Rocket.

Table 2: Original data of c10~c13 parameters.

Sample	Interpretation parameter													Time
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	
u_1	0.1	0.2	9.6	1.3	2.3	0.8	1.7	9.4	0.1	0.6	1.5	107	98	2
u_2	0.3	8.1	10	0.1	1.1	0.1	0.2	10	0.1	0.2	1.8	159	140	3
u_3	0.1	0.3	0.4	0.2	0.1	0.3	9.7	0.1	0.6	0.1	0.8	164	170	5
u_4	0.5	0.2	0.8	9.9	0.1	0.2	9.4	1	0.8	1	0.7	145	130	8
u_5	9.1	1.1	1	0.4	0.6	0.2	0.7	0.2	0.9	8.6	0.6	185	145	10
u_6	0.1	9.4	6.6	0.3	8.8	0.7	9.4	9.4	0.1	4.5	2.8	196	163	15
u_7	0.3	1.2	9.3	0.7	0.4	0.2	1.1	0.5	8	0.3	1.8	145	128	18
u_8	0.5	0.7	9.5	0.2	0.2	0.1	0.3	10	9.7	0.7	1.7	98	91	20

5.2. Implementation of Higher-Grade Criterion in Interpretation Server Layer. HBase column database and Qt cross-platform programming language are used to develop automatic interpretation software, which mainly implements functions such as data reception, criterion management, real-time automatic interpretation, and interpretation result sending. According to the specific interpretation requirements of the parameters, the syntax of the APIL language in Section 4.1, and the higher-grade criteria in Section 4.2, write the higher-grade criteria of the parameters. For the syntax of higher-grade criteria, the input .l and .y syntax files are compiled into .h and .cpp files that can be directly called by QT and C++ programs by using flex lexical compiler and bison syntax compiler.

The input of the .l and .y files and the output of the .h and .cpp files in this example verification are shown in Figure 7, and some high-level criteria for some parameters are shown in Figure 8. In the autointerpretation software, the real-time input interpretation data of the edge layer is read, the higher-grade criteria designed in Figure 8 are selected, and the specific interpretation process is performed, as shown in Figure 9.

By viewing the final interpretation report, the results of threshold and higher-grade interpretation of parameters c_{11}, c_{12}, and c_{13} are normal. Samples u_5, u_6, and u_7 are input into the existing diagnostic rule base, and it is determined that the stand-alone status is normal in these three cases. Therefore, these parameters and samples are not taken into consideration for subsequent real-time status diagnosis in the diagnostic model, and new diagnostic rules are extracted from the remaining abnormal parameters and fault samples.

5.3. Real-Time Diagnosis Based on VPRS. The diagnosis model is constructed and the diagnosis rules are extracted according to the following steps.

Step 1. Build the original decision table

Construct the universe set $U = \{u_1, u_2, u_3, u_4, u_8\}$, 5 sets of interpretation sample data as the original decision table, and

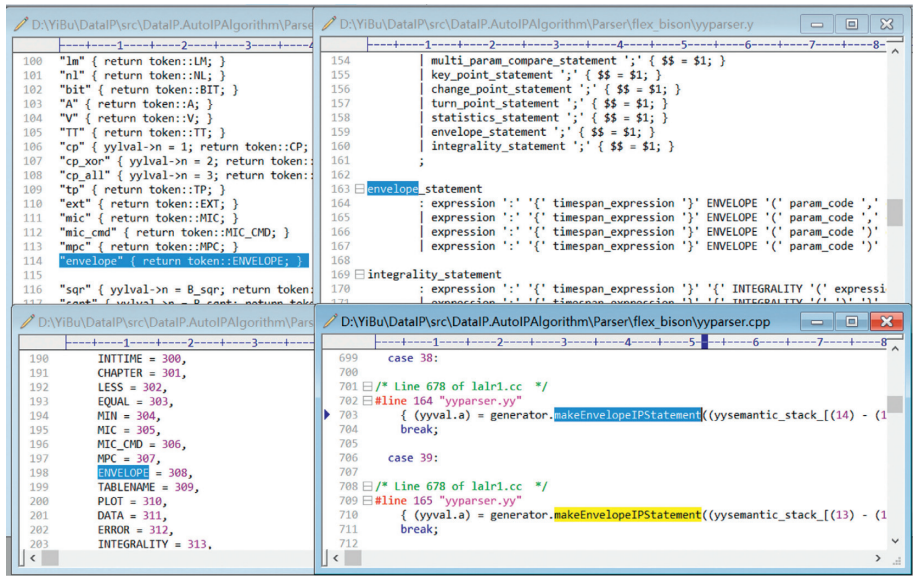


Figure 7: Input and output files of flex and bison.

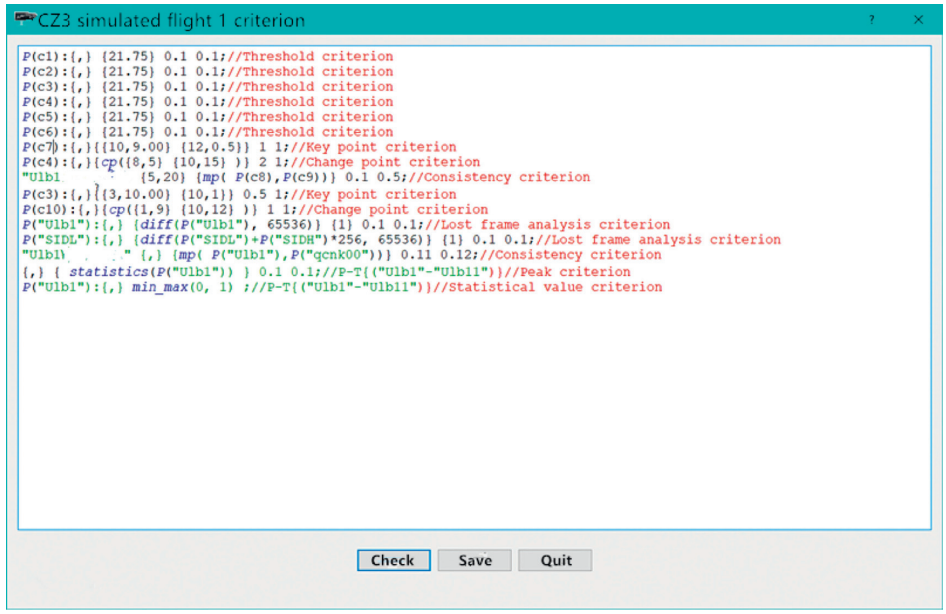


Figure 8: Higher-grade criterion of parameters.

10 interpretation parameters $C = \{c_1, c_2, \ldots, c_{10}\}$, as the set of conditional attributes in the decision table. According to the interpretation results of the parameters and the actual situation of the test, determine the type of single machine failure $D = \{d_1, d_2, \ldots, d_4\}$ as the decision attribute set of the decision table. Among them, d_1 represents the front cabin overload, d_2 represents the front cabin temperature abnormality, d_3 represents the front cabin vibration abnormality, and d_4 represents contact friction between the end handle and the inner wall.

Step 2. Data standardization

The condition attributes in the original decision table are standardized by Z-core normalization method, as shown in the following equation:

$$x_j = \frac{x_j - E(x_j)}{D(x_j)}, \quad j = 1, 2, \ldots m. \tag{4}$$

In the previous equation, $E(x_j)$ represents the mean value of the characteristic variable x_j, and $D(x_j)$ is the standard deviation corresponding to the characteristic variable x_j.

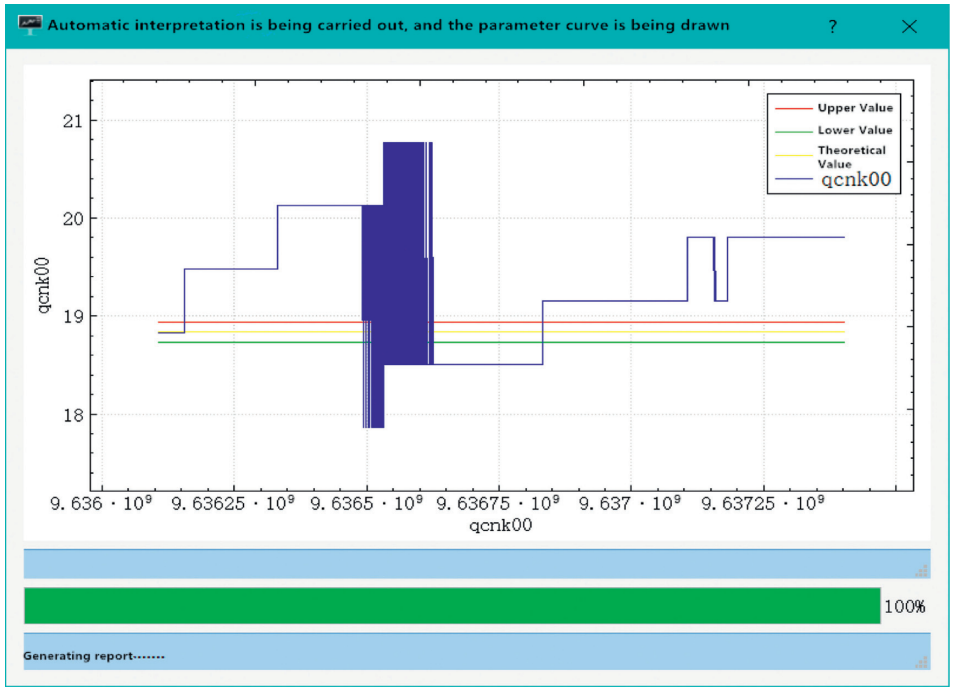


Figure 9: Perform process of autointerpretation.

Step 3. Data discretization

For the condition attributes and decision attributes in Table 2, discretization is needed after standardization according to Step 2. Take the interval value of [0, 5] as 0 and the interval value of [5, 10] as 1, and divide the range of continuous attributes into several subintervals to obtain the discrete decision table S', as shown in Table 3.

Step 4. Parameter dependency and importance calculation

According to formula (1), the dependence of fault state on each parameter is calculated; then, according to the calculation results of parameter dependence and according to formulas (2) and (3), the importance of parameters is calculated, and finally the order of parameter importance can be obtained as follows: $c_8 > c_4 > c_7 > c_2 > c_9 > c_5 > c_3 > c_6 > c_{10} > c_1$. It can be seen that the key parameter among the 10 parameters is c_8. In the subsequent manual interpretation and autointerpretation, according to the importance of the parameters, selectively increase the data upload amount of c_8 and c_4 and other parameters and speed up the sampling frequency of c_8 and c_4 and other parameters to focus on.

Step 5. Parameter reduction

According to the method in Section 4.3.3, set the threshold $\beta = 0.9$, use the VPRS knowledge reduction algorithm to reduce the 10 parameters, and the final reduced parameter set is $\{c_2, c_4, c_7, c_8, c_9\}$.

Step 6. Value reduction

For the decision table after attribute reduction, each sample forms a diagnostic rule. The value reduction method in Section 4.3.4 is used to remove redundant parameters in each diagnostic rule, and the minimum reduction rule is obtained as shown in Table 4.

Step 7. Extract diagnostic rules

The diagnostic rules are extracted for each fault sample after attribute reduction and value reduction in Table 4, and the results are as follows:

Rule 1: IF($c_2 \in [0,5] \cap c_8 \in [5,10] \cap c_9 \in [0,5]$ or $c_8 \in [5,10] \cap c_9 \in [5,10]$) THEN d_1(Front cabin overload)

Rule 2: IF(c2∈ [5,10]∩c7∈ [0,5]) THEN d2(Front cabin temperature abnormal)

Rule 3: IF($c_4 \in [0,5] \cap c_7 \in [5,10]$) THEN d_3(Front cabin vibration abnormal)

Rule 4: IF($c_4 \in [5,10]$) THEN d_4(Contact friction of inner wall of tip handle)

These rules are matched with the existing diagnostic rule library. If they are new rules, they are added to the rule base.

It can be seen from the range test that the real-time data of 13 parameters were initially collected at the edge, and the "double-layer interpretation" was used at the edge-interpretation layer to reduce the parameter data amount to 1/10 of the original. Then, in the interpretation server layer diagnosis model, the simplest diagnostic rules are obtained after reduction by VPRS theory. With two or three parameters, or even one parameter, the fault status of a single equipment can be determined, which greatly improves the efficiency of real-time diagnosis.

5.4. Discussion. Fault diagnosis is very important to the safety and reliability of equipment. There are many methods for equipment fault diagnosis. Some studies use machine

Table 3: Decision table of discretization.

Abnormal sample interpretation	Interpretation parameter										Single equipment fault
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	
u_1	0	0	1	0	0	0	0	1	0	0	d_1
u_2	0	1	1	0	0	0	0	1	0	0	d_2
u_3	0	0	0	0	0	0	1	0	0	0	d_3
u_4	0	0	0	1	0	0	1	0	0	0	d_4
u_8	0	0	1	0	0	0	0	1	1	0	d_1

Table 4: Minimal decision table.

Fault sample	Interpretation parameter					Fault condition
	c_2	c_4	c_7	c_8	c_9	
U_1	0	*	*	1	0	d_1
U_2	1	*	0	*	*	d_2
U_3	*	0	1	*	*	d_3
U_4	*	1	*	*	*	d_4
U_8	*	*	*	1	1	d_1

learning, artificial intelligence, and other methods, while others are based on current, acoustics, vibration, thermal analysis, and so on. In [10], an original method for feature extraction of thermal images BCAoID is proposed. The computed features were analyzed using the Nearest Neighbor classifier and the backpropagation neural network. It turns out to be the case that thermal images are a fast and noninvasive diagnostic method. It can be used for mechanical and electrical faults of the machine. This paper aims at the core problem that we have to use postinterpretation because of the large amount of interpretation data, so all the work focuses on how to reduce the amount of interpretation data and improve the speed of fault diagnosis, in order to achieve on-line interpretation and real-time diagnosis. Therefore, this paper uses VPRS method for fault diagnosis of rocket's single equipment. This method makes full use of large amounts of historical data and real-time data and uses big data mining technology to extract knowledge and intelligence. Finally, it can not only diagnose the current fault state of rocket's single equipment but also continuously simplify and optimize the diagnostic rules of equipment diagnosis. In the follow-up study, we will try to use thermal, acoustic, and other means for fault diagnosis.

6. Conclusion and Future Work

The main contributions in this paper are summarized as follows:

(1) By deploying the edge layer before the existing server, a "double-layer interpretation" architecture that includes the edge-interpretation layer and server interpretation layer is constructed.

(2) The specific process of edge interpretation is given. Through filtering the data according to the threshold interpretation, the amount of interpretation data is reduced and makes it possible to conduct real-time interpretation.

(3) In the interpretation server layer, AIPL criterion modeling language and high-grade criterion are designed to interpret the abnormal parameters further. VPRS is used to mine diagnostic rules from historical fault data, establish real-time diagnostic model, and realize the real-time diagnosis.

(4) The feasibility and effectiveness of the designed double-layer on-line interpretation and real-time diagnosis system are validated experimentally.

In addition to the VPRS method used in this paper, there are many other similar methods for status diagnosis at the interpretation server layer. Limited by the length of the article, we only adopt the VPRS method. In our next research, more methods will be tried to solve this problem, and the most suitable one will be selected finally through comparison of method results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Doctoral Innovation Fund of Xi'an University of Technology (no. 310-252072013) and the National Natural Science Foundation of China (Grant no. 52005404).

References

[1] L. P. Zhu, Y. S. Zhao, and L. M. Guo, "Real time automatic determine knowledge based research of telemeter parameter of carrier rocket," *Computer Engineering and Applications*, vol. 43, pp. 131–135, 2007.

[2] B. L. Liu and D. Jin, "Research simulation method of satellite temperature based on telemetry data," *Journal of Astronautics*, vol. 36, pp. 763–768, 2015.

[3] J. Guo and S. Jin, "Design and application of measurement data processing system," *Computer Measurement & Control*, vol. 27, pp. 250–252, 2019.

[4] P. Y. Guo, S. W. Ding, Z. H. Tian, H. L. Zhang, and X. H. Zhang, "System design and method research for optical measurement images real-time interpretation in test ranges," *Journal of National University of Defense Technology*, vol. 36, pp. 168–174, 2014.

[5] X. T. Shi, J. Y. Pang, X. Zhang, Y. Peng, and D. T. Liu, "Satellite big data analysis based on bagging extreme learning machine," *Chinese Journal of Scientific Instrument*, vol. 39, pp. 81–91, 2018.

[6] S. H. Dong, C. W. Cui, and P. J. He, "Automatic interpretation for telemetry parameters based on prediction method," *IOP Conference Series: Earth and Environmental Science*, vol. 514, pp. 6–22, 2020.

[7] B. Hu, Y. J. Zhang, and L. Li, "Research and design of comparative analysis platform for launch vehicle flight data," *Electronic Measurement Technology*, vol. 41, pp. 133–137, 2018.

[8] X. Li, J. Z. Gao, J. F. Cui, and H. Q. Xie, "A novel method of automatic interpretation for low-varying telemetry parameters," *Journal of Astronautics*, vol. 39, pp. 585–592, 2018.

[9] L. Gou, H. Li, H. Zheng, H. Li, and X. Pei, "Aeroengine control system sensor fault diagnosis based on CWT and CNN," *Mathematical Problems in Engineering*, vol. 2020, Article ID 5357146, 12 pages, 2020.

[10] A. Glowacz, "Fault diagnosis of electric impact drills using thermal imaging," *Measurement*, vol. 171, pp. 1–11, 2021.

[11] N. Shan, Y. Li, and X. Cui, "A multilevel optimization framework for computation offloading in mobile edge computing," *Mathematical Problems in Engineering*, vol. 2020, Article ID 4124791, 17 pages, 2020.

[12] W. X. Yin, Y. Z. Chen, and L. Chi, "Dynamic workload allocation for edge computing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 29, no. 3, pp. 1–11, 2021.

[13] J. L. D. Neto, S.-Y. Yu, D. F. Macedo et al., "ULOOF: a user level online offloading framework for mobile edge computing," *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2660–2674, 2018.

[14] A. Abouaomar, S. Cherkaoui, Z. Mlika, and A. Kobbane, "Resource provisioning in edge computing for latency sensitive applications," *IEEE Internet of Things Journal*, p. 1. In press, 2021.

[15] S. L. Huang, H. Z. Zhu, and X. L. Cheng, "Adaptive threshold setting method for layer-count in penetration fuse," *Acta Armamentarii*, vol. 41, pp. 1762–1771, 2020.

[16] W. Ziarko, "Variable precision rough set model," *Journal of Computer and System Sciences*, vol. 46, no. 1, pp. 39–59, 1993.

[17] S. Huang, "Limited and variable precision rough set model," *Journal of Information and Computational Science*, vol. 11, no. 10, pp. 3493–3501, 2014.

[18] H. Fu, J. Xu, H. Zhang, M. Zhang, and X. Xu, "Fault diagnosis of wireless sensor network based on optimized probabilistic neural network," *Journal Européen des Systèmes Automatisés*, vol. 51, no. 4-6, pp. 295–308, 2018.

[19] X. Chen, X. M. Wang, Y. Huang, and X. Y. Wang, "Fault diagnosis for tilt-rotor aircraft flight control system based on variable precision rough set-OMELM," *Control and Decision of Astronautics*, vol. 30, pp. 433–440, 2015.

[20] M. Beynon and K. L. Buchanan, "An illustration of variable precision rough set theory: the gender classification of the European barn swallow (Hirundo rustica)," *Bulletin of Mathematical Biology*, vol. 65, no. 5, pp. 835–858, 2003.

[21] X. Guan, X. Yi, and Y. He, "A new discretization algorithm for continuous interval-valued attributes," *Journal of Astronautics*, vol. 30, pp. 1164–1167, 2009.

[22] A. M. Isa, A. N. M. Rose, and M. M. Deris, "Dominance-based soft set approach in decision-making analysis," *Advanced Data Mining and Applications*, Springer, Berlin, Germany, pp. 299–310, 2011.

[23] Z. Y. Suo, S. Y. Cheng, X. J. Yuan, and Y. M. Li, "Rule acquisition method and application of dominance decision-making information system," *Acta Armamentarii*, vol. 36, pp. 539–544, 2015.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 7308454, 9 pages
https://doi.org/10.1155/2021/7308454

Research Article

Double-Acceptance Sampling Plan for Exponentiated Fréchet Distribution with Known Shape Parameters

M. Sridhar Babu,[1] G. Srinivasa Rao,[2] and K. Rosaiah[3]

[1]*Department of Sciences, St. Mary's College, Yousufguda, Hyderabad 500045, India*
[2]*Department of Mathematics and Statistics, The University of Dodoma, Dodoma, P.O.Box 259, Tanzania*
[3]*Department of Statistics, Acharya Nagarjuna University, Guntur 522007, India*

Correspondence should be addressed to G. Srinivasa Rao; gaddesrao@yahoo.com

Received 24 July 2020; Revised 1 October 2020; Accepted 18 February 2021; Published 28 February 2021

Academic Editor: Hussein Abulkasim

We suppose that a product's lifetime follow the exponentiated Fréchet distribution of defined shape parameters. Based on this assumption, a double-acceptance sampling plan is constructed. The zero and one failure framework is essentially thought of: if no errors are found from the first sample, then the lot is approved; also, if at least two failures occur, it is rejected. In the first sample, if one failure is observed, then the second sample is taken and decided for the same length as the first one. The cumulative sample sizes of the first and second samples are determined on the basis of the stated confidence level of the consumer to ensure that the actual median is longer than the given life. As indicated by the various ratios of the actual median life to specified median lifetime, the operating characteristics are calculated and placed in presented tables. To decrease the risk of the producer at the predefined level, the minimum ratios of this sort are additionally obtained. Lastly, examples are provided for representation reasons for the proposed model.

1. Introduction

Goods or products have lifetime differences, despite the fact that they are manufactured by the same producer and machine and under similar sates of assembling. Producers are exceptionally wary about the quality of their products, with the goal that when the customer shows up to buy them, they do not confront any challenges in getting them. It is not generally conceivable to test the full life of all goods from an enormous size assortment because of the expense and time of the test. The choice on the acceptance or rejection of the merchandise involving the manufacturer and the customer is therefore filled with vulnerabilities. The risk of the consumer is characterized as the probability of choosing a few items based on a sample from the lot that has the mean or median lifespan less than the predefined lifetime. The risk of the producer is the chance that a good quality product would be rejected as a defective product by the customer. An approach to limit the producer's risk and consumer's risk is to raise the size of the sample; however, this change will lead the producer to lose. To lessen such kind of risks, we should utilize a viable acceptance sampling plan. To lessen the trial period, truncated life tests are frequently implemented.

We use the distribution of statistical probabilities to calculate the variation in product lifetime. The underlying statistical distribution could be used to understand the sampling design parameters. Various studies have been done to develop single-sample plans under various statistical distributions dependent on life of the truncated tests. In a single sampling scheme, a lot is rejected and we stop the experiment based on the truncated life test if beyond c failures happen during the pre-fixed experiment time. In the event that the experiment time initiates c or fewer failures, the lot will be accepted.

Because of the simplicity of usage of the plan, much exertion has been made in the course of recent decades to explore the acceptance of single sampling plans for the testing and review of items for various sampling

circumstances. A number of authors have developed single-sampling plans for different distributions: Epstein [1]; Goode and Kao [2]; Gupta and Groll [3]; Gupta [4]; Kantam et al. [5]; Baklizi [6]; Baklizi and Masri [7]; Tsai and Wu [8]; Balakrishnan et al. [9]; Rosaiah and Kantam [10]; Rosaiah et al. [11]; Rao et al. [12, 13]; and Rosaiah et al. [14]. The above-listed plan has also recently been designed for various lifetime distributions.

In cases where normal distribution is often adopted, the double-sampling schemes were known to decrease the size of the sample or the risk of the producer in the area of quality control; e.g. Duncan [15]. The choice of double sampling decisions is based on the data collected from earlier decisions on the process. Aslam [16] developed Rayleigh distribution based on double-acceptance sampling depend on truncated life-tests. Aslam et al. [17, 18] proposed double-acceptance sampling plans based on truncated life tests for the Weibull distribution and general life distributions.

A double-acceptance sampling plan for generalized log-logistic distribution with known shape parameters have been developed by Aslam and Jun [19]. Rao [20, 21] considered double-acceptance sampling plans for the Marshall–Olkin extended exponential and Marshall–Olkin extended Lomax distributions depend on average life time of the truncated life testing data. Aslam et al. [22] discussed double-acceptance sampling plans for Burr type-XII distribution percentiles under the truncated life tests. In generalized exponential distribution, Ramaswamy and Anburajan [23] presented double acceptance sampling based on truncated life-tests. Gui [24] has developed a double-acceptance sampling plan for time-truncated life-tests based on Maxwell distribution. Malathi and Muthulakshmi [25] developed a zero-one double-acceptance sampling plan based on Marshall–Olkin extended exponential distribution for truncated life tests. Tripathi et al. [26] developed an application of time-truncated single-acceptance sampling plan based on generalized half-normal distribution. Tripathi et al. [27] studied double- and group-acceptance sampling plan for truncated life test based on inverse log-logistic distribution.

In the field of quality control, the normal distribution is often implemented; double sampling schemes have been reported to minimize the size of the sample or the producer's risk; see, e.g., Duncan [15] in the area of reliability. Nevertheless, a plan for double sampling has also not been developed. Based on the assumption that a product's lifetime follows the exponentiated Fréchet distribution, the main objective of this article is to recommend the double-acceptance sampling plans for truncated lifetime tests. In survival research, we especially use this distribution because of the monotonous nature of the hazard function. We essentially consider zero and one failure system where we accept a lot for no failures from the first sample and we reject if at least two failures are found. When there is only one defect, we will select the second sample and check with the same duration as the first sample. At the defined consumer's level of confidence, both the initial and second sample minimum sizes are calculated. The operational properties are calculated according to the ratio of true median life to the defined lifetime. The minimum ratios are often achieved in order to reduce the risk of the producer to the degree specified. Section 2 describes the exponential Fréchet distribution. Section 3 accounts for the proposed double sampling scheme, and Section 4 assesses its operating characteristics. Examples of the sampling technique are given in Section 5 and finally conclusions are given in Section 6.

2. The Exponentiated Fréchet Distribution (EFD)

Consider the product's lifetime following the exponentiated Fréchet distribution which was introduced and studied by Nadarjah and Kotz [28]; the probability density function and cumulative distribution function of EFD, respectively, are given by

$$f(t,\sigma,\lambda,\theta) = \sigma^{\lambda}\lambda\theta\left[1-e^{-(\sigma/t)^{\lambda}}\right]^{\theta-1} t^{-(1+\lambda)} e^{-(\sigma/t)^{\lambda}},$$
$$F(t,\sigma,\lambda,\theta) = 1-\left[1-e^{-(\sigma/t)^{\lambda}}\right]^{\theta}, \quad t>0,\sigma,\lambda,\theta>0. \tag{1}$$

Here, the scale parameter is σ and the shape parameters are λ and θ. Remember that this CDF represents the chance of failure of a parallel system with θ items having a Fréchet distributed lifetime. The exponentiated Fréchet distribution will then be used to check the reliability of the system. In constructing single-sampling plans, this exponentiated Fréchet distribution was considered in Rao et al. [13]. When $\theta = 1$, it is called the Fréchet distribution. For this analysis, the shape parameters λ and θ are assumed to be a priori known.

The parameters λ and θ can be predicted when data on failure are available. It is understood that for $\theta = 1$, the failure rate (hazard function) declines in the case of $\lambda \le 1$ while it rises and then declines in the case of $\lambda > 1$. This distribution is used as a model for lifetime in this study, as its failure rate pattern is very flexible.

Generally, the mean cannot necessarily be defined in a closed form, while exponential Fréchet distribution is possible in the closed form. However, the median life is the 50^{th} percentile of the exponentiated Fréchet distribution derived by

$$m = \sigma\left[-\ln\left(1-(0.5)^{1/\theta}\right)\right]^{-(1/\lambda)}. \tag{2}$$

When the parameters λ and θ are fixed, the median is directly proportional to the scale parameter. Note that the median reduces to $m = \sigma$ independently of λ for the Fréchet distribution ($\theta = 1$). Harlow [29] developed the Fréchet distribution function for applications. Nadarajah and Gupta [30] studied the beta Fréchet distribution. Abd-Elfattah and Omima [31] addressed the estimation of the unknown parameters of the generalized Fréchet distribution. Abd-Elfattah et al. [32] introduced the goodness of fit tests for generalized Fréchet distribution. Al-Nassar and Al-Omari [33] studied the acceptance sampling plan based on truncated life tests for exponentiated Fréchet distribution. Kotz and Nadarajah [34] studied theory and applications of extreme value distributions.

3. Design of Suggested Sampling Plan

Assume that the median (m) lifetime is used to determine the product's quality. We presumed that the lot of good quality is submitted if the statistics enables the below null hypothesis, H_0: $m \geq m_0$ towards the alternative hypothesis, H_1: $m < m_0$. The level of significance for the test, the risk of the consumer is used through $1 - p^*$, where p^* is the consumer's confidence level.

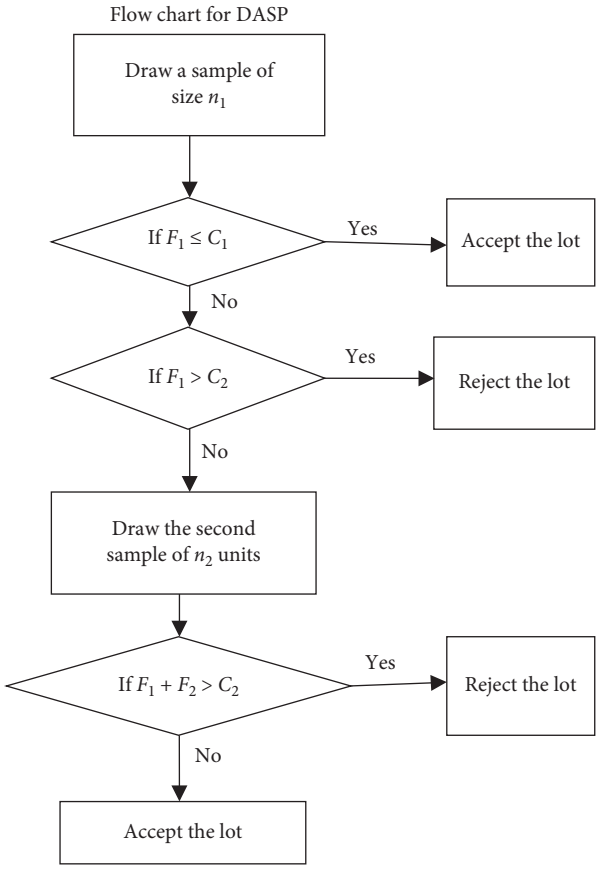
Based on the truncated life test, the following double-acceptance sampling plan is developed.

(i) Choose a sample of size n_1 at random and examine it. If C_1 or fewer failures occurred prior to t_0 the predefined experiment time, the lot is approved. The experiment has been truncated before time t_0 if $(C_2 + 1)$ failures are found, when the lot is rejected if $(C_1 < C_2)$.

(ii) If the failures in number lie between $C_1 + 1$ and C_2 by time t_0, then select the second sample of size n_2 and test them during time t_0. From both the samples, a maximum of C_2 failures are found; then the lot is approved. Failing that, the lot would be rejected.

For a given factor multiplier a, it is always feasible to fix the terminating time as a multiple of the defined lifetime m_0, in which case $t_0 = am_0$. The suggested sampling method is then described by the following five parameters (n_1, n_2, C_1, C_2, a), if $C_1 < C_2$.

The procedure for a double-acceptance sampling plan (DASP) for life test flow chart is given below.



The lot size is known as sufficiently large to calculate the acceptance probability of the lot in order to use the binomial distribution; see, for instance, Stephens [35], for further explanation of the application of binomial distribution. Under the suggested double-acceptance sampling plan, the lot acceptance probability is obtained by

$$P_a = \sum_{i=0}^{n_1} \binom{n_1}{i} p^i (1-P)^{n_1 - i} + \sum_{x=C_1+1}^{C_2} \binom{n_1}{x} p^x \cdot (1-P)^{n_1 - x} \left[\sum_{i=0}^{C_2 - x} \binom{n_2}{i} p^i (1-P)^{n_2 - i} \right]. \tag{3}$$

In the equation above, the probability that an item will fail before time t_0 is p, which is given by

$$p = 1 - \left[1 - e^{-\left(\sigma/t_0\right)^{\lambda}}\right]^{\theta} 1 - \left[1 - e^{-\left((m/m_0)/\eta_q a\right)^{\lambda}}\right]^{\theta}, \tag{4}$$

where $\eta_q = [-\ln(1 - (1-q)^{1/\theta})]^{-(1/\lambda)}$.

In general, we are focused in $C_1 = 0$ and $C_2 = 1$ and for the suggested double sampling plan, which would be regarded as zero and one failure scheme, as consumers prefer a sampling plan for acceptance with lower acceptance numbers. If the lot is accepted with several failed items from a test, this may not be understood by consumers though it may happen based on probabilities. In the zero and one failure plans, the probability of lot acceptance equation (3) reduces to

$$P_a = (1-p)^{n_1} \left[1 + n_1 p (1-p)^{n_2 - 1}\right]. \tag{5}$$

The minimum sizes of sample n_1 and n_2 guarantee $m \geq m_0$ at the level of confidence for consumers P^* can be achieved as a solution to the following inequality:

$$(1-p_0)^{n_1} \left[1 + n_1 P_0 (1-p_0)^{n_2 - 1}\right] \leq 1 - P^*, \tag{6}$$

where the probability p_0 in equation (3) obtained at $m = m_0$ as

$$p_0 = 1 - \left[1 - e^{-\left(1/\eta_q a\right)^{\lambda}}\right]^{\theta}, \quad \text{ASN} = n_1 p_1 + (n_1 + n_2)(1 - p_1). \tag{7}$$

The chance of rejection or acceptance is P_1 on the basis of first sample and is given by

$$P_1 = 1 - \sum_{i=c_1+1}^{c_2} \binom{n_1}{i} p^i (1-p)^{n_1 - i}. \tag{8}$$

For $C_1 = 0$ and $C_2 = 1$, we have

$$\text{ASN} = n_1 + n_1 n_2 p (1-p)^{n_1 - 1}. \tag{9}$$

The following optimization problem is then determined in terms of the minimum sample sizes for zero and one failure scheme in our double-acceptance sampling plan:

$$\text{ASN} = n_1 + n_1 n_2 p_0 (1-p_0)^{n_1 - 1}. \tag{10a}$$

Subject to

Table 1: ASN and minimum sample sizes under exponentiated Fréchet distribution.

a									
θ, λ	P^*	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9
(2.0, 2.0)	0.75	119, 93 (146.98)	11, 7 (12.85)	4, 3 (4.65)	2, 2 (2.39)	2, 1 (2.10)	2, 1 (2.05)	1, 1 (1.11)	1, 1 (1.08)
	0.90	176, 137 (202.26)	15, 14 (17.56)	6, 3 (6.36)	3, 3 (3.36)	2, 2 (2.21)	2, 1 (2.05)	2, 1 (2.03)	2, 1 (2.01)
	0.95	216, 183 (239.84)	19, 16 (20.87)	7, 5 (7.43)	4, 3 (4.19)	3, 2 (3.08)	2, 2 (2.10)	2, 1 (2.03)	2, 1 (2.01)
	0.99	315, 309 (328.60)	28, 20 (28.75)	10, 6 (10.16)	6, 3 (6.05)	4, 2 (4.03)	3, 2 (3.03)	3, 1 (3.00)	2, 2 (2.03)
(1.0, 2.0)	0.75	28, 20 (33.74)	7, 4 (7.97)	4, 2 (4.37)	3, 1 (3.14)	2, 2 (2.30)	2, 1 (2.10)	2, 1 (2.07)	2, 1 (2.05)
	0.90	40, 33 (46.24)	10, 6 (10.90)	5, 4 (5.53)	4, 2 (4.16)	3, 2 (3.15)	2, 2 (2.21)	2, 2 (2.14)	2, 1 (2.05)
	0.95	49, 45 (54.83)	12, 9 (12.93)	6, 5 (6.46)	4, 4 (4.33)	3, 3 (3.23)	3, 2 (3.08)	3, 1 (3.02)	2, 2 (2.10)
	0.99	73, 59 (75.42)	17, 14 (17.51)	9, 5 (9.13)	6, 4 (6.09)	5, 2 (5.03)	4, 2 (4.03)	4, 1 (4.01)	3, 2 (3.03)
(0.5, 2.0)	0.75	9, 9 (11.53)	5, 3 (5.66)	3, 3 (3.67)	3, 2 (3.32)	2, 2 (2.38)	2, 2 (2.31)	2, 1 (2.13)	2, 1 (2.11)
	0.90	14, 11 (15.86)	7, 4 (7.55)	5, 3 (5.33)	4, 2 (4.19)	3, 3 (3.34)	3, 2 (3.16)	3, 1 (3.06)	3, 1 (3.05)
	0.95	17, 14 (18.63)	8, 7 (8.73)	6, 3 (6.22)	5, 2 (5.11)	4, 2 (4.12)	4, 2 (4.08)	3, 2 (3.12)	3, 2 (3.09)
	0.99	25, 19 (25.71)	12, 8 (12.25)	8, 6 (8.17)	7, 3 (7.05)	6, 2 (6.03)	5, 3 (5.05)	4, 4 (4.10)	4, 3 (4.05)
(1.6684, 0.7130)	0.75	6, 5 (7.24)	4, 3 (4.68)	3, 3 (3.64)	3, 2 (3.33)	3, 1 (3.13)	2, 2 (2.35)	2, 2 (2.30)	2, 1 (2.13)
	0.90	9, 6 (9.87)	6, 4 (6.52)	5, 2 (5.20)	4, 2 (4.21)	3, 3 (3.38)	3, 2 (3.20)	3, 2 (3.15)	3, 1 (3.06)
	0.95	11, 7 (11.67)	7, 5 (7.47)	6, 3 (6.20)	5, 3 (5.18)	4, 3 (4.21)	4, 2 (4.10)	3, 3 (3.23)	3, 2 (3.12)
	0.99	15, 14 (15.52)	10, 8 (10.26)	8, 5 (8.13)	7, 3 (7.06)	6, 3 (6.05)	5, 4 (5.09)	5, 2 (5.03)	4, 4 (4.10)

$$\left(1-P_0\right)^{n_1}\left[1+n_1 p_0\left(1-P_0\right)^{n_2-1}\right] \le 1-P^*, \tag{10b}$$

$$1 \le n_2 \le n_1, \tag{10c}$$

$$n_1, n_2\text{: integers.} \tag{10d}$$

A quick search can solve this problem by modifying the values of n_1 and n_2. The size of the sample relating to the single-sampling method that estimates the initial values of n_1 and n_2:

$$\left(1-P_0\right)^{n} \le 1-P^*. \tag{11}$$

For example, Table 1 displays the lowest possible sample sizes for the first and second samples under the EFD according to various values of $P^* = (0.75, 0.90, 0.95, 0.99)$ and $a = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ three combinations of $(\theta = 0.5, 1.0, 2.0 \text{ and } \lambda = 2)$ were taken into consideration. It has been found that the sizes of sample rapidly increase as $(\lambda \text{ or } \theta)$ increases when the experiment time is comparatively less, but they remain about the same irrespective of $(\lambda \text{ or } \theta)$ when the experiment duration is longer.

4. Operating Characteristics (OC)

Therefore, we have to know the operating characteristics of the planned proposal on the basis of the ratio of actual median life to defined life t_q/t_q^0, i.e., (m/m_0) clearly, a scheme will become more appropriate if its OC increases very close to one. Tables 2–4 display the EFD OC values with of $P^* = (0.75, 0.90, 0.95, 0.99)$ and $a = (0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9)$ three combinations of $(\theta = 0.5, 1.0, 2.0 \text{ and } \lambda = 2)$. Table 5 display the OC values of EFD for the estimated parameters of $\widehat{\lambda} = 0.7130$ and $\widehat{\theta} = 1.6684$. When the actual median life increases more than the defined life, the chance of acceptance will increase. Therefore, we have to know the OC's for the suggested plan according to the ratio of the actual median life to the defined life, i.e. (m/m_0). Obviously, a method will be more suitable if its OC increases very close to one.

We may compare the DASP with the existing single sampling plan (when $c = 0$ and 1) in terms of OC values. From Table 4, the design parameters of DASP with $c_1 = 0, c_2 = 1$ when $\theta = 0.5, \lambda = 2, \beta = 0.25$ at $\delta_q = 0.5$ are $n_1 = 9$ and $n_2 = 9$. So, sample size for single sampling plan is 9. Therefore, we compare DASP with single sampling plan with $n = 9, c = 0$ and also for $n = 9, c = 1$. It is clear visible of these 3 plans in Figure 1. The OC values for DASP pass through middle of the two single sampling plans. It ultimately proves DASP provides better results than single stage group sampling plan.

5. Use of Tables and Example

The data obtained from the cleanup gradient test wells on vinyl chloride are considered. Vinyl chloride is an unstable chemical compound. This factor is of great interest as it is both anthropogenic and carcinogenic in environmental investigations. In addition, this component is found to be weak in a number of well-tracking history. Lower level identifications of this product are attributed to water or air cross-contamination, or to the testing method itself in safe historical wells. Bhaumik and Gibbons [36] and Krishnamoorthy et al. [37] discussed this principle in the development of predictive and tolerance intervals for gamma variables.

Data on vinyl chloride from safe groundwater observation wells (μg/L).

The validity of our model by plotting the superimposed data indicates that the EFD is a reasonable fit and also the goodness of fit is shown by the Q-Q plot seen in Figure 2. The maximum likelihood estimates of the EFD two-parameters for breaking carbon fiber stress are $\widehat{\lambda} = 0.7130$ and $\widehat{\theta} = 1.6684$, and using the Kolmogorov-Smirnov test, we found that the maximum distance between the data and the fitted of the EFD is 0.11804 with p value is 0.7306. The EFD therefore suits the vinyl chloride data perfectly.

Table 2: OC values for exponentiated Fréchet distribution with $\theta = 2$ and $\lambda = 2$.

Ratio m (m_0)								
P^*	n_1	n_2	a	2	4	6	8	10
0.75	119	93	0.5	0.9984	0.9989	0.9998	1.0000	1.0000
	11	7	0.7	0.9976	0.9992	0.9998	0.9999	1.0000
	4	3	0.9	0.9927	0.9982	0.9991	0.9995	1.0000
	2	2	1.1	0.9988	0.9991	0.9998	0.9999	1.0000
	2	1	1.3	0.9887	0.9956	0.9991	0.9989	0.9999
	2	1	1.5	0.9548	0.9789	0.9943	0.9991	0.9999
	1	1	1.7	0.8897	0.9823	0.9912	0.9992	0.9999
	1	1	1.9	0.8000	0.9899	0.9956	0.9990	0.9999
0.90	176	137	0.5	0.9900	0.9956	0.9976	0.9999	1.0000
	15	14	0.7	0.9891	0.9956	0.9991	0.9998	1.0000
	6	3	0.9	0.9919	0.9978	0.9991	0.9999	1.0000
	3	3	1.1	0.9966	0.9989	0.9994	0.9996	1.0000
	2	2	1.3	0.9685	0.9879	0.9923	0.9998	0.9999
	2	1	1.5	0.8835	0.9567	0.9964	0.9995	0.9999
	2	1	1.7	0.7424	0.9145	0.9989	0.9991	0.9998
	2	1	1.9	0.5789	0.9698	0.9987	0.9990	0.9999
0.95	216	183	0.5	0.9849	0.9878	0.9987	0.9997	1.0000
	19	16	0.7	0.9719	0.9919	0.9984	0.9996	0.9999
	7	5	0.9	0.9691	0.9899	0.9923	0.9988	0.9999
	4	3	1.1	0.9466	0.9879	0.9923	0.9998	0.9999
	3	2	1.3	0.9385	0.9567	0.9954	0.9985	0.9999
	2	2	1.5	0.8835	0.9145	0.9989	0.9995	1.0000
	2	1	1.7	0.7424	0.9098	0.9927	0.9939	0.9999
	2	1	1.9	0.5789	0.9198	0.9897	0.9919	0.9999
0.99	315	309	0.5	0.9987	0.9991	0.9991	0.9989	1.0000
	28	20	0.7	0.9912	0.9986	0.9991	0.9989	0.9999
	10	6	0.9	0.9799	0.9879	0.9923	0.9998	0.9999
	6	3	1.1	0.9433	0.9567	0.9974	0.9995	0.9999
	4	2	1.3	0.9114	0.9245	0.9969	0.9991	0.9998
	3	2	1.5	0.7994	0.9145	0.9929	0.9989	0.9998
	3	1	1.7	0.5948	0.9098	0.9917	0.9979	0.9999
	2	2	1.9	0.3955	0.9196	0.9789	0.9945	0.9987

Table 3: OC values for exponentiated Fréchet distribution with $\theta = 1$ and $\lambda = 2$.

Ratio m (m_0)								
P^*	n_1	n_2	a	2	4	6	8	10
0.75	28	20	0.5	0.9914	0.9979	0.9992	1.0000	1.0000
	7	4	0.7	0.9789	0.9982	0.9996	0.9999	1.0000
	4	2	0.9	0.9969	0.9982	0.9991	0.9995	1.0000
	3	1	1.1	0.9714	0.9991	0.9998	0.9999	1.0000
	2	2	1.3	0.9018	0.9856	0.9971	0.9988	0.9999
	2	1	1.5	0.7945	0.9689	0.9983	0.9996	0.9999
	2	1	1.7	0.6721	0.9823	0.9902	0.9982	0.9999
	2	1	1.9	0.5540	0.9899	0.9916	0.9929	0.9999
0.90	40	33	0.5	0.9678	0.9951	0.9978	0.9999	1.0000
	10	6	0.7	0.9599	0.9956	0.9991	0.9998	1.0000
	5	4	0.9	0.9449	0.9968	0.9991	0.9999	1.0000
	4	2	1.1	0.9349	0.9979	0.9991	0.9996	1.0000
	3	2	1.3	0.8530	0.9879	0.9923	0.9998	0.9999
	2	2	1.5	0.7091	0.9577	0.9955	0.9991	0.9999
	2	2	1.7	0.5604	0.9145	0.9989	0.9991	0.9998
	2	1	1.9	0.4303	0.9691	0.9937	0.9929	0.9999

Table 3: Continued.

Ratio m (m_0)								
P^*	n_1	n_2	a	2	4	6	8	10
0.95	49	45	0.5	0.9788	0.9878	0.9982	0.9996	1.0000
	12	9	0.7	0.9697	0.9908	0.9974	0.9995	0.9999
	6	5	0.9	0.9519	0.9868	0.9905	0.9989	0.9999
	4	4	1.1	0.9466	0.9879	0.9925	0.9988	0.9999
	3	3	1.3	0.8285	0.9567	0.9954	0.9986	0.9999
	3	2	1.5	0.6664	0.9145	0.9986	0.9993	1.0000
	3	1	1.7	0.5045	0.9091	0.9923	0.9935	0.9999
	2	2	1.9	0.3685	0.9198	0.9895	0.9917	0.9999
0.99	73	59	0.5	0.9916	0.9981	0.9997	0.9999	1.0000
	17	14	0.7	0.9899	0.9984	0.9992	0.9989	0.9999
	9	5	0.9	0.9769	0.9876	0.9917	0.9998	0.9999
	6	4	1.1	0.9169	0.9566	0.9973	0.9995	0.9999
	5	2	1.3	0.7498	0.9245	0.9969	0.9991	0.9998
	4	2	1.5	0.5455	0.9145	0.9935	0.9988	0.9999
	4	1	1.7	0.3667	0.9098	0.9917	0.9969	0.9999
	3	2	1.9	0.2358	0.9193	0.9789	0.9935	0.9977

Table 4: OC values for exponentiated Fréchet distribution with $\theta = 0.5$ and $\lambda = 2$.

Ratio m (m_0)								
P^*	n_1	n_2	a	2	4	6	8	10
0.75	9	9	0.5	0.9899	0.9965	0.9979	0.9992	0.9999
	5	3	0.7	0.9813	0.9959	0.9979	0.9992	0.9998
	3	3	0.9	0.9543	0.9913	0.9986	0.9994	0.9997
	3	2	1.1	0.8795	0.9976	0.9993	0.9997	0.9999
	2	2	1.3	0.7874	0.9967	0.9956	0.9991	0.9999
	2	2	1.5	0.6944	0.9675	0.9739	0.9945	0.9996
	2	1	1.7	0.6089	0.9679	0.9833	0.9914	0.9995
	2	1	1.9	0.5336	0.9384	0.9799	0.9944	0.9989
0.9	14	11	0.5	0.9898	0.9978	0.9996	0.9998	0.9999
	7	4	0.7	0.9865	0.9923	0.9946	0.9989	0.9999
	5	3	0.9	0.9164	0.9887	0.9967	0.9977	0.9999
	4	2	1.1	0.7931	0.9903	0.9981	0.9996	0.9996
	3	3	1.3	0.6565	0.9836	0.9879	0.9923	0.9998
	3	2	1.5	0.5324	0.9655	0.9767	0.9961	0.9996
	3	1	1.7	0.4293	0.9402	0.9215	0.9978	0.9994
	3	1	1.9	0.3469	0.8869	0.9668	0.9977	0.9998
0.95	17	14	0.5	0.9968	0.9991	0.9996	0.9998	0.9999
	8	7	0.7	0.9704	0.9945	0.9969	0.9994	0.9999
	6	3	0.9	0.8833	0.9776	0.9896	0.9933	0.9998
	5	2	1.1	0.7254	0.9689	0.9874	0.9921	0.9997
	4	2	1.3	0.5645	0.9506	0.9567	0.9954	0.9995
	4	2	1.5	0.4299	0.9046	0.9147	0.9987	0.9996
	3	2	1.7	0.3264	0.8956	0.9098	0.9924	0.9937
	3	2	1.9	0.2493	0.8471	0.9198	0.9896	0.9991
0.99	25	19	0.5	0.9895	0.9989	0.9993	0.9993	0.9999
	12	8	0.7	0.9559	0.9878	0.9985	0.9992	0.9999
	8	6	0.9	0.7715	0.9999	0.9879	0.9923	0.9998
	7	3	1.1	0.5334	0.9873	0.9542	0.9974	0.9995
	6	2	1.3	0.3445	0.9783	0.9845	0.9968	0.9998
	5	3	1.5	0.2199	0.9229	0.9545	0.9941	0.9999
	4	4	1.7	0.1426	0.8286	0.9398	0.9908	0.9989
	4	3	1.9	0.0948	0.7112	0.9296	0.9789	0.9925

Table 5: OC values for exponentiated Fréchet distribution with $\widehat{\theta} = 1.6684$ and $\widehat{\lambda} = 0.7130$.

Ratio m (m_0)								
P^*	n_1	n_2	a	2	4	6	8	10
0.75	6	5	0.5	0.9636	0.9990	1.0000	1.0000	1.0000
	4	3	0.7	0.8898	0.9926	0.9994	0.9999	1.0000
	3	3	0.9	0.7951	0.9761	0.9968	0.9995	0.9999
	3	2	1.1	0.6978	0.9484	0.9906	0.9981	0.9996
	3	1	1.3	0.6076	0.9111	0.9797	0.9950	0.9987
	2	2	1.5	0.5280	0.8672	0.9636	0.9896	0.9968
	2	2	1.7	0.4593	0.8196	0.9427	0.9813	0.9936
	2	1	1.9	0.4008	0.7705	0.9178	0.9702	0.9889
0.90	9	6	0.5	0.9385	0.9982	0.9999	1.0000	1.0000
	6	4	0.7	0.8235	0.9871	0.9989	0.9999	1.0000
	5	2	0.9	0.6884	0.9592	0.9944	0.9991	0.9998
	4	2	1.1	0.5616	0.9139	0.9837	0.9967	0.9993
	3	3	1.3	0.4539	0.8557	0.9651	0.9912	0.9977
	3	2	1.5	0.3664	0.7901	0.9385	0.9818	0.9944
	3	2	1.7	0.2968	0.7221	0.9049	0.9679	0.9889
	3	1	1.9	0.2418	0.6553	0.8660	0.9494	0.9806
0.95	11	7	0.5	0.9218	0.9976	0.9999	1.0000	1.0000
	7	5	0.7	0.7839	0.9831	0.9985	0.9998	1.0000
	6	3	0.9	0.6316	0.9476	0.9926	0.9988	0.9998
	5	3	1.1	0.4969	0.8915	0.9787	0.9956	0.9990
	4	3	1.3	0.3884	0.8216	0.9550	0.9885	0.9969
	4	2	1.5	0.3045	0.7454	0.9218	0.9763	0.9926
	3	3	1.7	0.2405	0.6687	0.8806	0.9585	0.9854
	3	2	1.9	0.1918	0.5956	0.8339	0.9353	0.9748
0.99	15	14	0.5	0.8792	0.9961	0.9998	1.0000	1.0000
	10	8	0.7	0.6865	0.9727	0.9975	0.9997	1.0000
	8	5	0.9	0.4965	0.9178	0.9880	0.9981	0.9997
	7	3	1.1	0.3473	0.8351	0.9659	0.9928	0.9984
	6	3	1.3	0.2406	0.7372	0.9291	0.9813	0.9949
	5	4	1.5	0.1671	0.6364	0.8792	0.9621	0.9880
	5	2	1.7	0.1171	0.5408	0.8194	0.9345	0.9764
	9	6	0.5	0.9385	0.9982	0.9999	1.0000	1.0000

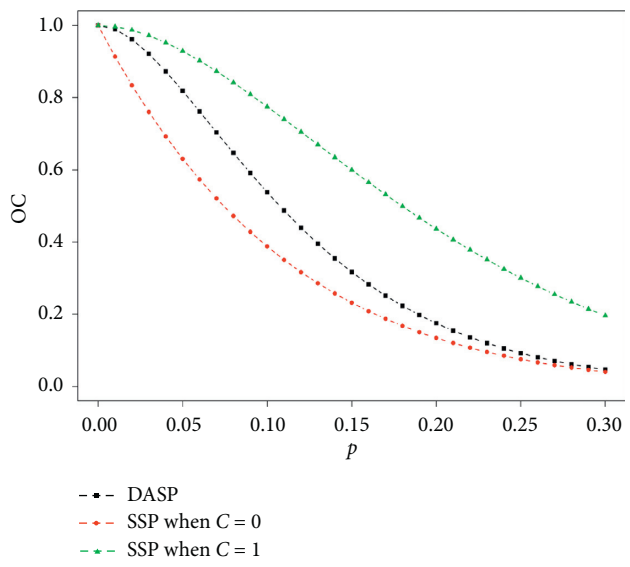


Figure 1: OC curve of double- and single-sampling plans.

Suppose the product's lifetime follows exponentiated Fréchet distribution with parameters $\theta = 2$ and $\lambda = 2$. It can also be known that the manufacturer would like to learn whether the median life of the product is above or equivalent to 1000 hours at a level of confidence 0.75. The researcher wants to end an experiment at 500 hours under the zero and one failure plan of the double sampling plan. It refers to the $a = 0.7$ terminator of the experiment. From Table 1, the required minimum sizes of sample are $n_1 = 11$ and $n_2 = 7$. The double sampling plan shall be described as follows. The very first thing to do here is to monitor eleven items for 500 hours and accept the lot if no failure occurs during the experiment. When the experiment produces at least two errors, the lot is rejected. The second sample of seven items is drawn and tested for 500 hours where only one failure was reported. The lot would be accepted if no failure exists in the second sample. When quality improves, the product can be correlated with the probability of acceptance, and they want to reduce the risk of the producer. Suppose the supplier knows what degree of quality leads to a risk of less than 0.05 from the product.

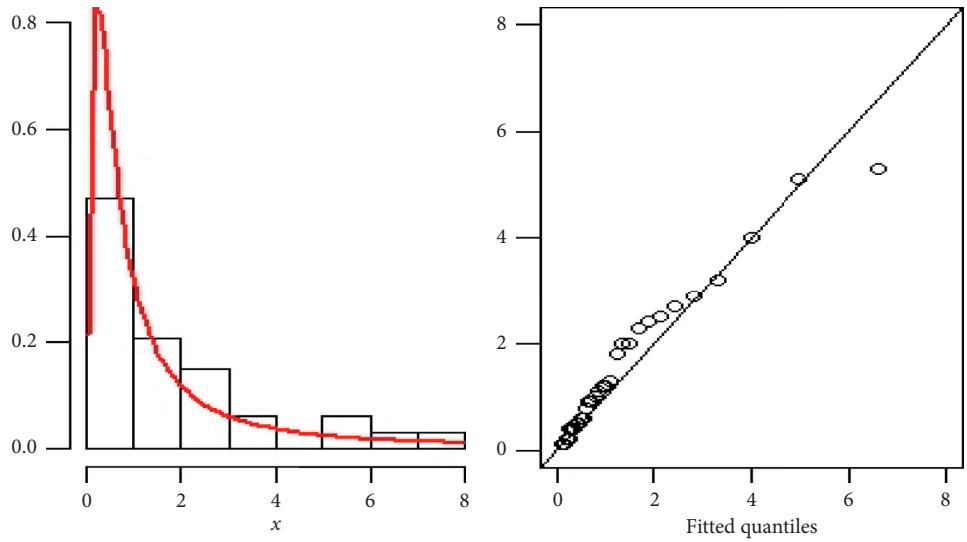


Figure 2: The empirical and theoretical cdfs and Q-Q plots for the vinyl chloride data. (a) Empirical and fitted PDFs, (b) Q-Q plot.

6. Conclusion

A double-sampling procedure for the decision to approve or reject the lot submitted was built on the basis of a truncated life test. The lifespan of the product is expected to follow exponentiated Fréchet distribution, which is useful in system reliability analysis because the failure rate is very flexible. It was observed that the necessary sample sizes declined steadily as the time of the experiment grew and that the size of the sample for the reasonable duration of the experiment was not very sensitive to the confidence level or the shape parameter. It has been revealed by examples that the double sampling plan would be more appropriate than a single sampling plan in terms of the OC values. A variable sampling plan will be preferred, because it utilizes all the details available. As a result, the potential research will establish a double acceptation sampling plan on variables.

Data Availability

The used data sets are given in the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Methodology and computation were carried out by GSR and writing and data collection done by MS and KR. Both authors read and approved the final manuscript.

Acknowledgments

Corresponding author from low income country got a 100% discount on article processing charge (APC) for the accepted articles.

References

[1] B. Epstein, "Truncated life tests in the exponential case," *The Annals of Mathematical Statistics*, vol. 25, no. 3, pp. 555–564, 1954.

[2] H. P. Goode and J. H. K. Kao, "Sampling plans based on the Weibull distribution," in *Proceeding of the Seventh National Symposium on Reliability and Quality Control*, pp. 24–40, Philadelphia, PA, USA, 1961.

[3] S. S. Gupta and S. S. Gupta, "Gamma distribution in acceptance sampling based on life tests," *Journal of the American Statistical Association*, vol. 56, no. 296, pp. 942–970, 1961.

[4] S. S. Gupta, "Life test sampling plans for normal and log-normal distributions," *Technometrics*, vol. 4, no. 2, pp. 151–175, 1962.

[5] R. R. L. Kantam, K. Rosaiah, and G. S. Rao, "Acceptance sampling based on life tests: log-logistic model," *Journal of Applied Statistics*, vol. 28, no. 1, pp. 121–128, 2001.

[6] A. Baklizi, "Acceptance sampling based on truncated life tests in the Pareto distribution of the second kind," *Advances and Applications in Statistics*, vol. 3, pp. 33–48, 2003.

[7] A. Baklizi and A. E. Q. El Masri, "Acceptance sampling based on truncated life tests in the Birnbaum Saunders model," *Risk Analysis*, vol. 24, no. 6, pp. 1453–1457, 2004.

[8] T.-R. Tsai and S.-J. Wu, "Acceptance sampling based on truncated life tests for generalized Rayleigh distribution," *Journal of Applied Statistics*, vol. 33, no. 6, pp. 595–600, 2006.

[9] N. Balakrishnan, V. Leiva, and J. López, "Acceptance sampling plans from truncated life tests based on the generalized Birnbaum-Saunders distribution," *Communications in Statistics-Simulation and Computation*, vol. 36, no. 3, pp. 643–656, 2007.

[10] K. Rosaiah and R. R. L. Kantam, "Acceptance sampling based on the inverse Rayleigh distribution," *Economic Quality Control*, vol. 20, no. 2, pp. 277–286, 2005.

[11] K. Rosaiah, R. R. L. Kantam, and C. Santosh Kumar, "Reliability of test plans for exponentiated log-logistic distribution," *Economic Quality Control*, vol. 21, no. 2, pp. 279–289, 2006.

[12] G. S. Rao, K. Rosaiah, K. Kalyani, and D. C. U. Sivakumar, "A new acceptance sampling plans based on percentiles for odds exponential log logistic distribution," *The Open Statistics & Probability Journal*, vol. 7, no. 1, pp. 45–52, 2016.

[13] G. S. Rao, K. Rosaiah, M. S. Babu, and D. C. U. SivaKumar, "A new acceptance sampling plans based on percentiles for exponentiated Fréchet distribution," *Economic Quality Control*, vol. 31, no. 1, pp. 37–44, 2016.

[14] K. Rosaiah, G. S. Rao, D. C. U. Sivakumar, and K. Kalyani, "The odd generalized exponential log logistic distribution: a new acceptance sampling plans based on percentiles," *International Journal of Advances in Applied Sciences*, vol. 8, no. 3, pp. 176–183, 2019.

[15] A. J. Duncan, *Quality Control and Industrial Statistics*, Irwin, Ed., Richard D. Irvin Inc., Homewood, IL, USA, 5th edition, 1986.

[16] M. Aslam, "Double acceptance sampling based on truncated life tests in Rayleigh distribution," *European Journal of Scientific Research*, vol. 17, no. 4, pp. 605–610, 2005.

[17] M. Aslam, C. H. Jun, and M. Ahmad, "A double acceptance sampling plan based on the truncated life tests in the Weibull model," *Journal of Statistical Theory and Applications*, vol. 8, no. 2, pp. 191–206, 2009.

[18] M. Aslam, C.-H. Jun, and M. Ahmad, "Design of a time-truncated double sampling plan for a general life distribution," *Journal of Applied Statistics*, vol. 37, no. 8, pp. 1369–1379, 2010.

[19] M. Aslam and C.-H. Jun, "A double acceptance sampling plan for generalized log-logistic distributions with known shape parameters," *Journal of Applied Statistics*, vol. 37, no. 3, pp. 405–414, 2010.

[20] G. S. Rao, "Double acceptance sampling plans based on truncated life tests for the Marshall-Olkin extended exponential distribution," *Austrian Journal of Statistics*, vol. 40, no. 3, pp. 169–176, 2011.

[21] G. S. Rao, "Double acceptance sampling plans based on truncated life tests for Marshall-Olkin extended Lomax distribution," *Journal of Modern Applied Statistical Methods*, vol. 10, no. 1, p. 12, 2011.

[22] M. Aslam, Y. Mahmood, Y. L. Lio, T.-R. Tsai, and M. A. Khan, "Double acceptance sampling plans for Burr type XII distribution percentiles under the truncated life test," *Journal of the Operational Research Society*, vol. 63, no. 7, pp. 1010–1017, 2012.

[23] A. S. Ramaswamy and P. Anburajan, "Double acceptance sampling based on truncated life tests in generalized exponential distribution," *Applied Mathematical Sciences*, vol. 6, no. 64, pp. 3199–3207, 2012.

[24] W. Gui, "Double acceptance sampling plan for time truncated life tests based on Maxwell distribution," *American Journal of Mathematical and Management Sciences*, vol. 33, no. 2, pp. 98–109, 2014.

[25] D. Malathi and S. Muthulakshmi, "Special double sampling plan for truncated life tests based on the Marshall-Olkin extended exponential distribution," *International Journal of Computational Engineering Research*, vol. 5, no. 1, pp. 56–62, 2015.

[26] H. Tripathi, M. Saha, and V. Alha, "An application of time truncated single acceptance sampling inspection plan based on generalized half-normal distribution," *Annals of Data Science*, 2020.

[27] H. Tripathi, S. Dey, and M. Saha, "Double and group acceptance sampling plan for truncated life test based on inverse log-logistic distribution," *Journal of Applied Statistics*, vol. 1, 2020.

[28] S. Nadarajah and S. Kotz, "The exponentiated Fréchet distribution, Inter Stat." *Electronics Journal*, 2006.

[29] D. G. Harlow, "Applications of the Fréchet distribution function," *International Journal of Materials and Product Technology*, vol. 17, no. 5/6, pp. 482–495, 2002.

[30] S. Nadarajah and A. K. Gupta, "The beta Fréchet distribution," *Far East Journal-Theory and Statistics*, vol. 14, no. 1, pp. 15–24, 2004.

[31] A. M. Abd-Elfattah and A. M. Omima, "Estimation of the unknown parameters of the generalized Fréchet distribution," *Journal of Applied Sciences Research*, vol. 5, no. 10, pp. 1398–1408, 2009.

[32] A. M. Abd-Elfattah, A. F. Hala, and A. M. Omima, "Goodness of fit tests for Generalized Fréchet distribution," *Australian Journal of Applied Sciences*, vol. 4, no. 2, pp. 286–301, 2010.

[33] A. D. Al-Nassar and A. I. Al-Omari, "Acceptance sampling plan based on truncated life tests for exponentiated Fréchet distribution," *Journal of Statistics*, vol. 25, no. 2, pp. 107–119, 2013.

[34] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, London, UK, 2000.

[35] K. S. Stephens, *The Handbook of Applied Acceptance Sampling: Plans, Principles and Procedures*, ASQ Quality Press, Milwaukee, WI, USA, 2001.

[36] D. K. Bhaumik and R. D. Gibbons, "One-sided approximate prediction Intervals for at LeastpofmObservations from a gamma population at each ofrLocations," *Technometrics*, vol. 48, no. 1, pp. 112–119, 2006.

[37] K. Krishnamoorthy, T. Mathew, and S. Mukherjee, "Normal-based methods for a gamma distribution," *Technometrics*, vol. 50, no. 1, pp. 69–78, 2008.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6640777, 8 pages
https://doi.org/10.1155/2021/6640777

Research Article

Calculation of Negative Frictional Resistance of Foundation Pile in Deep Fill Foundation

Ningyu Zhao,[1,2] **Hongjun Wu**,[2] **Yi Song**,[2] **and Shun Xiang**[2]

[1]*State Key Laboratory of Mountain Bridge and Tunnel Engineering, Chongqing Jiaotong University, Chongqing 400074, China*
[2]*School of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China*

Correspondence should be addressed to Ningyu Zhao; zhny@cqjtu.edu.cn

Received 4 November 2020; Revised 30 December 2020; Accepted 30 January 2021; Published 27 February 2021

Academic Editor: Hussein Abulkasim

In deep fill foundations, the pile foundation might suffer from negative frictional resistance (NFR) due to the consolidation and settlement of the soil. The NFR will cause the pile to settle excessively and reduce its bearing capacity. However, there are not yet many accurate methods to calculate the NFR of foundation piles in deep fill foundations. To make up for the gap, this paper carries out shear tests on the pile-soil interface and discusses the mechanism of pile-side frictional resistance. Considering the distribution law of pile-side frictional resistance with depth, the authors proposed a piecewise calculation model for pile-side frictional resistance, which couples the hyperbolic model and effective stress method. Then, the energy balance equation of the pile when the NFR occurs was established in the light of the energy transfer of the pile-soil system during the settlement of the soil around the pile. Furthermore, the calculation formulas of the axial force and displacement of the pile at different depths were derived, considering the pile-soil displacement and the potential energy change of the pile-soil system. The proposed method was applied to calculate the NFR of the foundation pile in a construction site, and the calculated results were compared with the measured data. The results show that the axial force-depth curve of the pile obtained through theoretical calculation agrees well with the measured data. Hence, our method can accurately reveal the mechanical features of the pile foundation in deep fill foundations.

1. Introduction

In mountainous and hilly areas, the foundation of construction sites often features undulating bedrock surfaces and deep fills. If the superstructure has a large load, the deep foundation structure needs to meet a high settlement requirement. In this case, pile foundation is one of the few deep foundation structures that can be built directly on the said foundation.

Prior to the construction of the pile foundation, the filling site usually undergoes years of foundation treatments, such as self-weight consolidation and dynamic compaction. Although most of these deep fills have completed consolidation and settlement, negative frictional resistance (NFR) will be induced by the compressive deformation and incomplete consolidation of the foundation soil.

The relative displacement between the pile and the soil has a great impact on the distribution and magnitude of the pile-side frictional resistance, especially at the neutral point. The calculation of pile-soil displacement is very important for accurately predicting the influence of the pull-down load caused by NFR on the bearing capacity design of the pile foundation.

Many scholars have attempted to calculate the NFR of pile foundation. For instance, Zhang et al. [1] combined the effective stress method with Terzaghi's theory of one-dimensional (1D) consolidation to identify the NFR features of precast concrete (PCC) piles. Chen et al. [2] proved that the pile-side NFR can be simplified into a linear model by the effective stress method. Poulos et al. [3–5] proposed the elastic theory method based on Mindlin's solution and obtained the size and distribution of pile-side NFR using the differential equation of the pile body. Drawing on Kezdi's double-broken-line model, Zhao et al. [6] designed a load transfer model based on the effect of the pile-soil displacement and deduced the piecewise analytical solution of

the pile-side frictional resistance on each soil layer. Considering the effects of effective stress and pile-soil displacement on pile-side frictional resistance, He et al. [7] established a calculation model for the friction on the side of the rock-socketed pile by the load transfer method and iteratively computed the analytical solution of the axial force and the solution of power series for pile displacement. Chen et al. [8] summarized the results of soil structure interaction (SSI) tests, improved the hyperbolic model considering the change of initial shear stiffness and the shear features of the pile-soil system, and applied the improved model to disclose the change law of NFR and maximum axial force.

To sum up, the above studies mainly adopt four methods, namely, elastic theory method, effective stress method, load transfer method, and shear displacement method. Among them, the elastic theory method has a large error except for end-bearing piles, for the neglection of the effects of pile-soil displacement and inability to acquire accurate parameters. Despite its simple and easy-to-use formula, the effective stress method only applies to the situation with small pile-soil displacement because of the relatively large calculation result. The model of the load transfer method is too complex to be solved easily. The shear displacement method ignores the stress state of the foundation soil, the layering features, and the pile-top settlement.

These methods seldom consider the essence of negative skin friction, that is, the potential energy loss of soil around the pile and the process of energy transfer to the pile body. Studies have shown that the NFR hinges on the gravitational potential lost during the settlement of the surrounding soil and its shear strength [9]. Therefore, the NFR calculation should consider how the change in the potential energy of the pile-soil system affects the pile-side NFR.

To disclose the influence of energy conversion on the evolution of pile-side NFR, this paper analyzes the change law of the shear stress/normal stress-shear displacement ($\tau/\sigma_n - u$) curve through shear tests on the pile-soil interface. Then, a piecewise calculation model was constructed for pile-side frictional resistance in the light of the pile-soil displacement. Based on the reciprocal theorem, the control equations of the axial force and displacement of the pile were derived under different pile-soil displacements and solved by the iterative method. The proposed method was proved feasible and accurate by comparing the calculation results with the measured data.

2. Calculation Models

2.1. Calculation Model of Pile-Side Frictional Resistance. The relative displacement on the pile-soil interface is the root cause of pile-side frictional resistance. To calculate pile-side frictional resistance, the key lies in setting up a suitable pile-soil displacement model. Moreover, the shear features of the pile-soil interface inevitably change with the normal stress, before and during the shearing of that interface [10]. Therefore, the normal stress must be considered in the calculation model of pile-side frictional resistance.

To study the load transfer function of pile-side frictional resistance in deep fill, shear tests were carried out on the pile-soil interface in a deep fill site under different normal stresses ($\sigma_n = 100$, 200, and 400 kPa). The shear stress-shear displacement curves of the pile-soil interface were normalized into the $\tau/\sigma_n - u$ curves (Figure 1).

As shown in Figure 1, the shear stress reached the limit state at the pile-soil displacement of 6 mm. Before that, the $\tau/\sigma_n - u$ curves were hyperbolic. Therefore, the stress state before the full exertion of pile-side frictional resistance can be simulated well by the hyperbolic model reflecting the pile-soil displacement and the lateral pressure of the surrounding soil. After the shear displacement of the pile-soil system arrived at the limit state, the shear stress/normal stress (τ/σ_n) remained constant.

In deep fill foundations, the settlement of soil around the pile decreases with the depth, while the settlement of the pile-bottom is a certain value under the load of the pile, and the compression deformation of the pile decreases with the depth. The displacement of the soil above the neutral point is greater than that of the pile, and the pile body above the depth is subject to NFR. When the relative displacement of pile-soil is greater than the limit relative displacement, the NFR changes greatly. The relative displacement of pile-soil decreases with the depth, and the relative displacement at the neutral point is zero. The displacement of the pile below the neutral point is greater than that of the soil, and the pile is subject to positive friction resistance (PFR), and the pile-soil relative displacement increases gradually with the decrease of soil settlement.

According to the above shear features of the pile-soil interface and the distribution law of pile-side frictional resistance with depth, the calculation of pile-side frictional resistance can be illustrated by the sketch map in Figure 2.

As shown in Figure 2, the pile-side frictional resistance can be calculated in three segments:

(1) The first segment is the NFR growth segment, where pile-soil displacement is greater than ultimate relative displacement ($s - s_p > u_m$), and the NFR increases with the vertical effective stress of the soil; in this segment, the pile-side frictional resistance is calculated by effective stress method

(2) The second segment is the transition segment from NFR to positive frictional resistance (PFR), where pile-soil displacement is smaller than ultimate relative displacement ($|s - s_p| \le u_m$); in this segment, the pile-side frictional resistance is calculated by the hyperbolic formula, which considers the lateral pressure of the surrounding soil

(3) The third segment is where the pile-soil displacement exceeds the ultimate relative displacement ($s_p - s > u_m$); in this segment, the PFR equals the ultimate frictional resistance τ_m

Under the above conditions, the calculation model of pile-side frictional resistance can be established as follows:

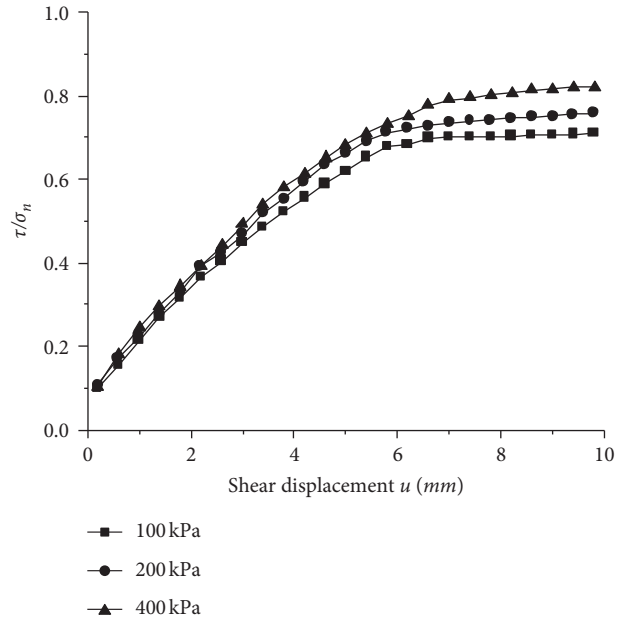


Figure 1: The $\tau/\sigma_n - u$curves.

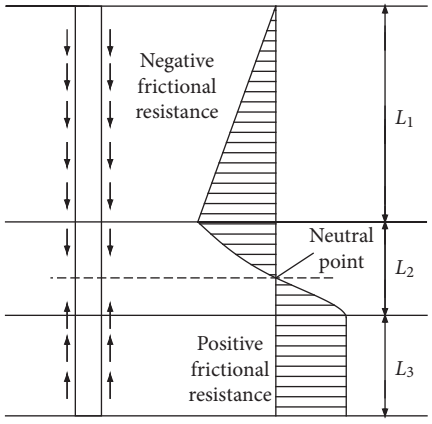


Figure 2: The calculation of pile-side frictional resistance.

$$\tau(z) = \begin{cases} \beta(q + \gamma' z), & s - s_p > u_m, \\ \dfrac{u\sigma_n}{A + Bu}, & |s - s_p| \le u_m, \\ \tau_m, & s_p - s > u_m, \end{cases} \quad (1)$$

where s_p and s are the settlements of the pile and the surrounding soil, respectively; β is the coefficient of frictional resistance; q is the overload acting on the ground;γ' is the weight of soil per unit volume; u_m is the ultimate relative displacement of the surrounding soil; $u$is the relative displacement;σ_n is the Earth pressure at rest; A and B are fitting parameters; τ_m is the ultimate frictional resistance.

The A and B values of the hyperbolic curve before the full exertion of the pile-side frictional resistance at the deep fill site were obtained by fitting the data of the $\tau/\sigma_n - u$ curves prior to the ultimate shear displacement with the hyperbolic formula in formula (1). The fitting results are presented in Figure 3.

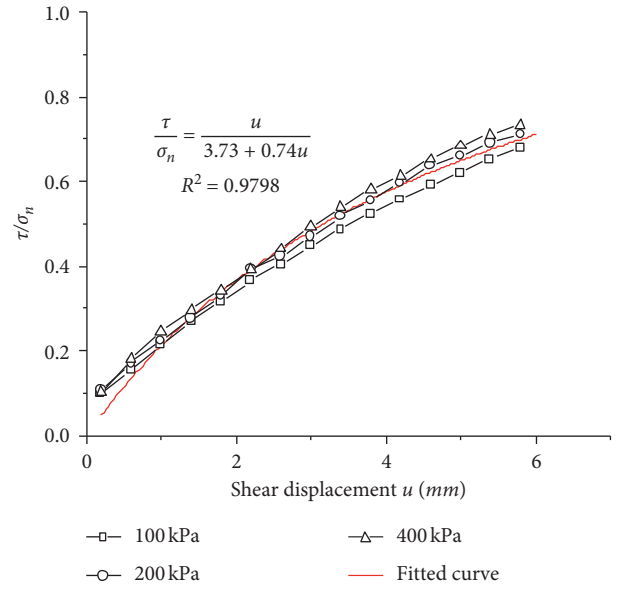


Figure 3: The fitted curve for pile-soil load transfer function.

From the $\tau/\sigma_n - u$ curves in Figure 3, it can be obtained that the fitting parameters $A = 3.73$, $B = 0.74$, and the coefficient of determination $R^2 = 0.9798$. Substituting A and B into formula (1), the transfer function of pile-side frictional resistance in the elastic state can be obtained.

2.2. Calculation of Pile-Bottom Resistance. Within the elastic range, the pile-bottom settlement basically has a linear relationship with pile-bottom resistance [11–14], and the pile-bottom settlement increases with pile-top settlement. Since the current code [15] forbids excessive pile-top settlement, it is assumed that the pile-bottom bearing layer does not have plastic deformation. Considering the pile-bottom as a rigid pressure block, it is realistic to calculate pile-bottom settlement by the Boussinesq equation [16]:

$$s_b = \frac{(1 - \mu_b)P_b}{4r_0 G_b}, \quad (2)$$

where r_0 is the radius of the pile; μ_b is Poisson's ratio of the soil at the pile-top; P_b is the pile-bottom resistance; and $G_b$is the shear modulus.

3. Building and Solving Energy Balance Equations

According to the NFR transfer law of piles in deep fill sites [17–20], and considering the engineering practicability of the mechanical model, the following assumptions were put forward before setting up the energy balance equations for the pile-soil system by the energy method:

(1) Under the pile-top load, the pile-side frictional resistance is exerted before the pile-top resistance

(2) The pile only suffers from elastic deformation rather than plastic deformation

(3) The pile elements are only affected by the vertical load.

3.1. Energy Balance Equations of the Pile-Soil System. The total potential energy of the pile-soil system in the foundation consists of the deformation energy of the pile, the potential energy increment caused by settlement, and the potential energy of the external forces [21]. The potential energy of the pile can be expressed as follows:

$$\Pi_p = \frac{1}{2}\int_0^L E_p A_p \left(\frac{\mathrm{d}\Delta}{\mathrm{d}z}\right)^2 dz - mg\left(s_b + \frac{\Delta}{2}\right), \tag{3}$$

where $L$is the pile length; $E_p$is the elastic modulus of the pile; $A_p$is the cross-sectional area of the pile; Δ is the deformation of the pile; $s_b$is pile-bottom displacement; m is the mass of the pile; and g is the acceleration of gravity.

The potential energy of the external force mainly encompasses the work of the pile-top load, the work of the pile-top resistance, and the work of the pile-side frictional resistance. Hence, the potential energy of the external forces can be expressed as follows:

$$\Pi_w = P_0 s_0 + \iint \tau(z)\left(s_b + \frac{\Delta}{2}\right)\mathrm{d}S - P_b s_b, \tag{4}$$

where P_0, P_b are pile-top load and pile-bottom load, respectively, and s_0 is pile-top displacement.

Based on the reciprocal theorem, the potential energy of the pile equals the total potential energy of the external forces acting on the pile. Combining formulas (3) and (4),

$$\begin{aligned}&\frac{1}{2}\int_0^L E_p A_p \left(\frac{d\Delta}{dz}\right)^2 \mathrm{d}z - mg\left(s_b + \frac{\Delta}{2}\right)\\ &\quad = P_0 s_0 + \iint \tau(z)\left(s_b + \frac{\Delta}{2}\right)\mathrm{d}S - P_b s_b.\end{aligned} \tag{5}$$

From the pile-top, the pile was discretized into n units (Figure 4). The nodes are denoted in turn as 0, 1, 2, . . ., $i-1$, i, $i+1$, . . ., n. The length of each element is $h = l/n$. The ith unit can be analyzed as follows:

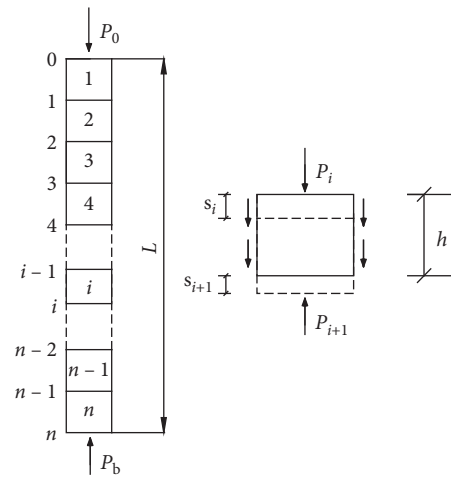


Figure 4: The discretization of the foundation pile.

$$\Delta_i = s_{pi} - s_{p(i+1)}, \tag{6}$$

$$\frac{\mathrm{d}\Delta_i}{\mathrm{d}z} = \frac{s_{pi} - s_{p(i+1)}}{h}, \tag{7}$$

where h is the length from the bottom to the top of the ith unit; Δ_i is the deformation of the ith unit; s_{pi} is the pile-top displacement of the ith unit; and $s_{p(i+1)}$is the pile-bottom displacement of the ith unit.

For the ith unit, substituting formulas (6) and (7) into formula (5):

$$\begin{aligned}P_i s_{pi} - P_{i+1} s_{p(i+1)} = \frac{1}{2}\int_{l_i}^{l_{i+1}} \cdot \Bigg\{ & E_p A_p \left(\frac{s_{pi} - s_{p(i+1)}}{h}\right)^2 - 2\pi r_0 \\ & \times \left(\tau_i s_{pi} + \tau_{i+1} s_{p(i+1)}\right)\Bigg\}\mathrm{d}z \\ & - \frac{mg}{2}\left(s_{pi} + s_{p(i+1)}\right).\end{aligned} \tag{8}$$

Integrating the right part of formula (8):

$$P_i s_{pi} - P_{i+1} s_{p(i+1)} = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right)^2 - \pi r_0 h\left(\tau_i s_{pi} + \tau_{i+1} s_{p(i+1)}\right) - \frac{mg}{2}\left(s_{pi} + s_{p(i+1)}\right). \tag{9}$$

Expanding formula (9) and merging similar terms:

$$\begin{aligned}P_i s_{pi} - P_{i+1} s_{p(i+1)} = & \left\{\frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) - \pi r_0 h\tau_i - \frac{mg}{2}\right\} \\ & \cdot s_{pi} - \left\{\frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) + \pi r_0 h\tau_{i+1} + \frac{mg}{2}\right\} s_{p(i+1)}.\end{aligned} \tag{10}$$

Judging by the structure on both sides, formula (10) holds only if

$$P_i = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) - \pi r_0 h \tau_i - \frac{mg}{2}, \tag{11}$$

$$P_{i+1} = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) + \pi r_0 h \tau_{i+1} + \frac{mg}{2}. \tag{12}$$

The top pressure and bottom axial force in each unit can be solved by formulas (11) and (12).

The pile-bottom deformation of each unit can be obtained by transforming formula (11):

$$s_{p(i+1)} = s_{pi} - \frac{2h}{E_p A_p}\left(P_i + \pi r_0 h \tau_i + \frac{mg}{2}\right). \tag{13}$$

Substituting formula (1) into formulas (11) and (13), the axial force and displacement at each position of the pile can be obtained.

If $s - s_p > u_m$, then

$$P_i = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) - \pi r_0 h \beta\left(q + \gamma' z\right) - \frac{mg}{2}, \tag{14}$$

$$s_{p(i+1)} = s_{pi} - \frac{2h}{E_p A_p}\left(P_i + \pi r_0 h \beta\left(q + \gamma' z\right) + \frac{mg}{2}\right). \tag{15}$$

If $|s - s_p| \le u_m$, then

$$P_i = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) - \frac{\pi r_0 h u \sigma_n}{A + Bu} - \frac{mg}{2},$$
$$s_{p(i+1)} = s_{pi} - \frac{2h}{E_p A_p}\left(P_i + \frac{\pi r_0 h u \sigma_n}{A + Bu} + \frac{mg}{2}\right). \tag{16}$$

If $s_p - s > u_m$, then

$$P_i = \frac{E_p A_p}{2h}\left(s_{pi} - s_{p(i+1)}\right) - \pi r_0 h \tau_m - \frac{mg}{2}, \tag{17}$$

$$s_{p(i+1)} = s_{pi} - \frac{2h}{E_p A_p}\left(P_i + \pi r_0 h \tau_m + \frac{mg}{2}\right). \tag{18}$$

3.2. Numerical Algorithm. Under external forces, the total potential energy II of the pile-soil system equals the deformation energy of the pile and the soil plus the potential energy of the external forces. The numerical method can be introduced for actual calculation: each pile is discretized into n units according to pile length and soil thickness, each of which satisfies the conditions of static balance and continuity; the physical parameters of the target foundation pile are known, while the pile-top load and displacement are unknown. The numerical calculation can be implemented in the following steps (Figure 5):

(1) According to the calculation accuracy and pile length, the pile is divided into n units of the length $h = l/n$, and the mass of each unit is solved.

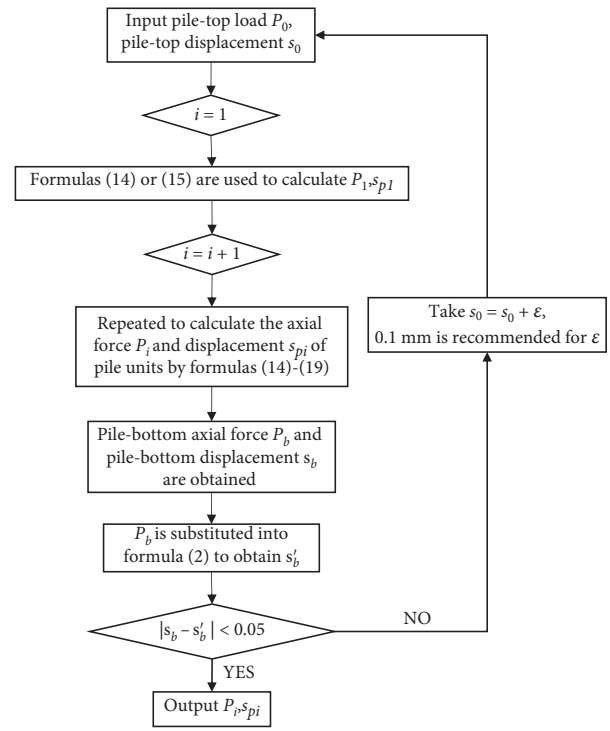


Figure 5: The calculation flow chart.

(2) Let P_0 and s_0 be pile-top load and pile-top displacement, respectively. Then, pile-soil displacement $u = s_i - s_{pi}$ is calculated, while P_1 and s_{p1} are solved by formulas (14) and (15), respectively.

(3) According to pile-soil displacement, Step (2) is repeated to calculate the axial force$P_i$and displacement s_{pi} of pile units by formulas (14)–(18) from top to bottom. If $s_i - s_{pi} = 0$, the axial force P_i of the unit is the maximum axial force of the pile, and the NFR is zero; that is, the node depth is the depth of the neutral point.

(4) Pile-bottom axial force P_b and pile-bottom s_b are obtained through further iterations. Then, P_b is substituted into formula (2) to obtain s_b'. If $|s_b - s_b'|/s_b > 0.05$, then the calculated pile-bottom displacement s_b will be increased by 0.1 mm, and a new iteration will be carried out from Step 1; if $|s_b - s_b'|/s_b < 0.05$, then the calculated pile-bottom displacement s_b is true, and Step 5 will be executed.

(5) The axial force P_i and displacement s_{pi} of each unit are calculated.

4. Comparative Analysis

To verify its engineering applicability, our method was used to calculate the NFR of foundation pile in a construction site, and the calculated results were compared with the data measured from pile-side frictional resistance tests [22]. The iterative calculation program was compiled on MATLAB.

The construction site is formed through land preparation and backfilling over two years. The backfilled soil is 3.8–29.8 m

Table 1: The physical-mechanical parameters of different soil layers.

Soil layer	Compression modulus (MPa)	Poisson's ratio	Cohesion (KPa)	Internal friction angle (°)	Unit weight (kN/m^3)
Fill	9.23	0.35	12.83	14.7	19
Slightly dense pebble	22.15	0.3	0	26	22
Completely weathered mudstone	15.22	0.3	20	25	19

thick. Prior to the construction of the pile foundation, the site was dynamically compacted at the maximum energy level of 18,000 kN·m, with the replacement piers reaching 8–12 m in depth. Some parts of the site were dynamically compacted at the energy level of 3,000 kN·m.

The test piles are cast-in-situ bored piles (diameter: 0.8 m; material: C30 concrete; elastic modulus: 3e7 KPa; weight per unit volume: 24 kN/m^3; Poisson's ratio: 0.16). Among them, 1# and 2# are friction piles, which reach 2 m into the bearing layer of slightly dense pebbles; 3# is the end-bearing pile, which reaches 3 m into the bearing layer of strongly weathered argillaceous siltstone.

During the calculation, the ultimate relative displacement was set to 6 mm. The ultimate frictional resistance was empirically set as per the properties of the soil layers. The physical-mechanical parameters of the surrounding soil are listed in Table 1.

The settlement of the surrounding soil can be calculated by different methods, depending on the actual situation. To pinpoint the transfer law of the pile-side frictional resistance at the site, the measured settlement of the surrounding soil was linearly fitted, and the pile-soil displacement was analyzed to verify the proposed method.

Figure 6 shows the variation of the settlement of the surrounding soil with depths obtained by the linear fitting of the measured settlement. The fitting formula can be expressed as follows:

$$s(z) = 43.789 - 1.353z. \quad (19)$$

The standard error of the intercept of the linear fitting formula was 1.1011, the standard error of the slope was 0.0601, and R^2 was 0.99215. These results basically agree with the measured data. The calculation shows the maximum settlement (43.789 mm) occurs on the surface.

Figure 7 compares the calculated axial force with the measured data. It can be seen that the axial force first increased and then decreased along the pile depth, which is consistent with the distribution law of measured axial force with depth. The maximum axial force appeared at the neutral point.

The calculated axial forces of 1#-3# peaked at 673 kN, 619 kN, and 680 kN, respectively, while their measured axial forces were 655 kN, 586 kN, and 646 kN, respectively. The calculated axial forces were all greater than the measured values. The relative errors of 1#-3# were 2.7%, 5.6%, and 5.0%, respectively.

In addition, the calculated depths of neutral point of 1#-3# were 14.3 m, 14.6 m, and 17.9 m, respectively, while the measured depths of the neutral point were 15 m, 15 m, and

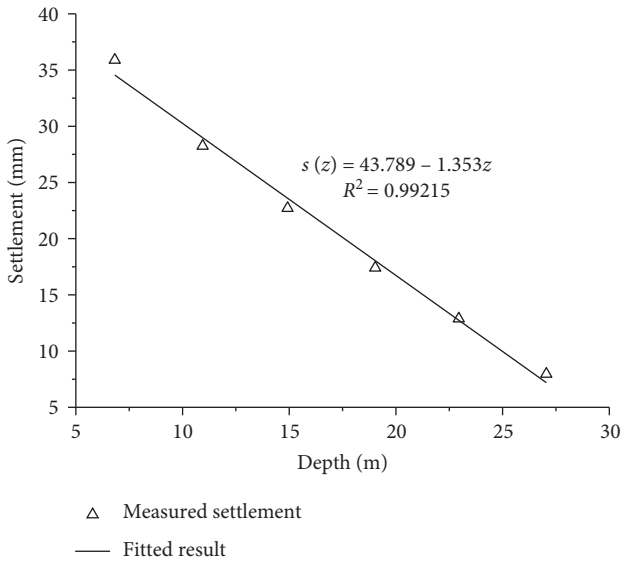


Figure 6: The calculation flow chart.

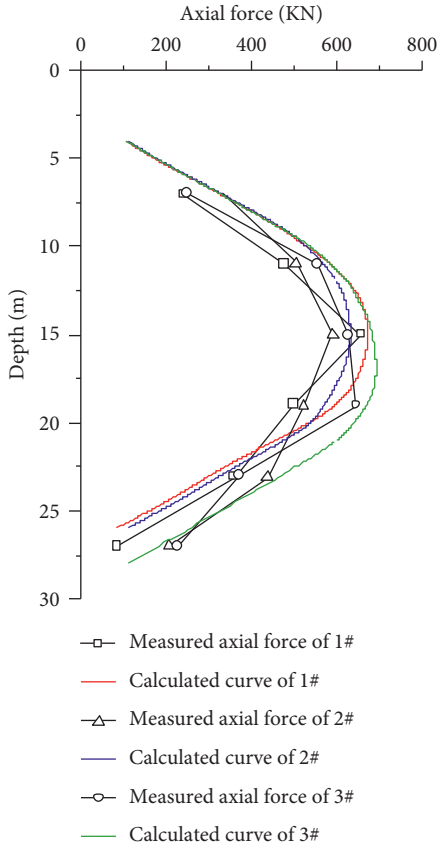


Figure 7: The comparison between calculated axial force and measured axial force.

19 m, respectively. The calculated depths were all shallower than the measured values. The relative errors of 1#-3# were 4.6%, 2.6%, and 5.7%, respectively.

The above data analysis shows that the theoretical distribution of axial force with depth is in line with the test results, while the calculated axial force is slightly larger than the measured value. A possible reason is that the NFR exertion is suppressed effectively within the range of dynamically compacted soil. The calculated value is safer than the measured value.

5. Conclusions

This paper probes deep into the mechanism and transfer law of pile-side frictional resistance in deep fill foundations and then establishes a piecewise calculation model for pile-side frictional resistance. In addition, the theoretical formulas of pile foundation were derived by the energy method. The main conclusions are as follows:

(1) The hyperbolic model was adopted to calculate the pile-side frictional resistance before the pile-soil shear displacement reaches the limit state, and the piecewise calculation model for pile-side frictional resistance was established by combining the hyperbolic model with an effective stress method.

(2) Based on the piecewise calculation model and the reciprocal theorem, the calculation formulas were theoretically derived for the axial force and displacement of the pile in deep fill foundations, in the light of the pile-soil displacement and the potential energy change of the pile-soil system.

(3) Taking a construction site, for example, the distribution of the axial force of the piles along the depth obtained by our theoretical formula basically agrees with that measured in tests. For the three test piles, the maximum error between calculated and measured peak axial forces was 5.6%, and the maximum error between calculated and measured depths of the neutral point was 5.7%. The results indicate that our method has high accuracy and provides a good reference for related engineering practices.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Project no. 51608081) and the Scientific and Technological Research Program of the Chongqing Municipal Education Commission (nos. KJQN201800743 and KJ1600532).

References

[1] X. J. Zhang, H. L. Liu, K. Fei, and Y. F. Gao, "Study on mechanism for negative skin friction of cast-in-situ concrete thin-wall pipe pile," *Rock and Soil Mechanics*, vol. 26, pp. 91–94, 2005.

[2] J. Chen, J. J. Zheng, B. G. Chen, and Y. E. Lu, "Analysis of working behavior of rigid pile composite ground considering influence of negative skin friction," *Rock and Soil Mechanics*, vol. 29, no. 7, pp. 1955–1959, 2008.

[3] H. G. Poulos and N. S. Mattes, "The analysis of down-drag in end-bearing piles," in *Proceedings of the 7th ICSMFE*, pp. 203–209, Sociedad Mexicana de Mecanica de Suelos, Mexico City, September 1969.

[4] H. G. Poulos, "Analysis of the settlement of pile groups," *Géotechnique*, vol. 18, no. 4, pp. 449–471, 1968.

[5] H. G. Poulos and E. H. Davis, "The development of negative friction with time in end-bearing piles," *Australian Geomechanics*, vol. 35, no. 1, pp. 11–20, 1972.

[6] M. H. Zhao, W. He, and W. G. Cao, "Study on calculation of negative skin friction resistance on piles," *Rock and Soil Mechanics*, vol. 25, no. 9, pp. 1442–1446, 2004.

[7] C. B. He, M. H. Zhao, and Y. Lei, "Negative friction computational research of rock-socketed pile based on load transfer method," *Engineering Mechanics*, vol. 31, no. 11, pp. 110–115, 2014.

[8] R. P. Chen, W. H. Zhou, and W. P. Cao, "Improved hyperbolic model of load-transfer for pile-soil interface and its application in study of negative friction of single piles considering time effect," *Chinese Journal of Geotechnical Engineering*, vol. 29, no. 6, pp. 824–830, 2007.

[9] D. Sahoo, B. Mohanty, and A. Maalika Veetil, "Evaluation of bond strength and flash mass on friction surfaced deposition of aluminium 6063 over IS 2062 low carbon steel using different mechtrode face," *Annales de Chimie - Science des Matériaux*, vol. 44, no. 2, pp. 109–119, 2020.

[10] C. Zhao, H. Gong, C. Zhao, and Q. Miu, "Elastoplastic analysis of interface between clay and concrete considering effect of normal stress history," *Chinese Journal of Rock Mechanics and Engineering*, vol. 31, no. 4, pp. 848–855, 2012.

[11] B. H. Fellenius, "Results from long-term measurement in piles of drag load and downdrag," *Canadian Geotechnical Journal*, vol. 43, no. 4, pp. 409–430, 2006.

[12] S. Yamato and M. B. Karkee, "Reliability based load transfer characteristics of bored precast piles equipped with grouted bulb in the pile toe region," *Soils and Foundations*, vol. 44, no. 3, pp. 57–68, 2004.

[13] J. An and C. Sun, "Safety assessment of the impacts of foundation pit construction in metro station on nearby buildings," *International Journal of Safety and Security Engineering*, vol. 10, no. 3, pp. 423–429, 2020.

[14] T. W. Lai, H. Lei, Z. Y. Ji, and Y. Liang, "Effects of cement-modified soil as blocking cushion of saline soil subgrade," *Revue des Composites et des Matériaux Avancés-Journal of Composite and Advanced Materials*, vol. 30, no. 1, pp. 49–53, 2019.

[15] China Academy of Building Sciences, *JGJ94–2008 Technical Code for Building Pile Foundation*, China Architecture and Building Press, Beijing, China, 1994.

[16] B. Zhu, *Computational Soil Mechanics*, Shanghai Scientific and Technical Publishers, Shanghai, China, 1990.

[17] Y. Zheng, J. Mao, and S. Liang, "Negative skin friction of pile foundation considering soil consolidation in high fill site,"

Journal of Jilin University (Engineering and Technology Edition), vol. 47, no. 4, pp. 1075–1081, 2017.

[18] J. Xu, *The Negative Skin Friction Characteristics and Research on Design Method of Rock-Socketed Piles in Deep Recent Filling Soil Areas*, Southwest Jiaotong University, Chengdu, China, 2013.

[19] M. Liu, *"Research on the Calculation and Characteristic of Negative Friction of Pile Foundation in Newly Thick fill*, Hunan University of Science and Technology, Xiangtan, China, 2014.

[20] G. Ye, W. Zheng, and Z. Zhang, "Investigation on distribution of negative friction of frictional piles in large filling sites," *Rock and Soil Mechanics*, vol. 40, no. 1, pp. 440–448, 2019.

[21] M. H. Zhao and S. S. Liu, "Numerical simulation of negative skin friction on single pile in multiple layer deposits," *Chinese Journal of Geotechnical Engineering*, vol. 30, no. 3, p. 336, 2008.

[22] J. K. Lv, *"Study on the Frictional Resistance of Friction Piles after Deep-Filling and Strong ramming*, Southwest University of Science and Technology, Mianyang, China, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6651615, 14 pages
https://doi.org/10.1155/2021/6651615

Research Article

A Novel Scheduling Algorithm for Common Rail Dual Automatic Guided Vehicles Particle Filtering Algorithm for Industrial Process Control

Yanghua Gao, Weidong Lou, and Hailiang Lu

Information Center, China Tobacco Zhejiang Industrial Co., Ltd., Hangzhou 310008, China

Correspondence should be addressed to Yanghua Gao; yhgao@zju.edu.cn

Received 3 November 2020; Revised 20 November 2020; Accepted 27 November 2020; Published 6 January 2021

Academic Editor: Hussein Abulkasim

This paper explores deep into the collaborative scheduling of common rail dual automatic guided vehicles (AGVs). Firstly, a dual AGV scheduling model was constructed to minimize the overall time of material distribution. Then, a novel scheduling algorithm was developed to dynamically plan the orders based on time windows. To effectively minimize the distribution time, heuristic algorithms were adopted to initialize the distribution order of materials. On this basis, the collaboration between the two AGVs was innovatively designed based on dynamic planning and time windows, making up for the defects of traditional methods in AGV cooperation. This greatly shortens the running time of the entire system in material distribution. The computing results fully demonstrate the efficiency and rationality of our algorithm. Finally, our algorithm was proved to be superior to the benchmark method through experiments on actual industrial instances.

1. Introduction

With economic growth and industrial development, there is a rising demand for automation and intelligent transport. This gives birth to a series of automated logistics equipment, such as automatic identification system, automatic guided vehicle (AGV) system, and cargo tracking system. The AGV system boasts greater application potential than the other systems, owing to its advantages in automation and efficiency.

The operations in many factories are faced with a nondeterministic polynomial-time (NP) hard problem: two AGVs at the opposite ends of a rail need to deliver materials to tanks along the rail, and the material distribution should be planned reasonably to minimize the distribution time. This problem was defined by Erdoğan in 2013 as the twin robots scheduling problem (TRSP) [1].

The TRSP applies widely in the robotic industry, where gantry robots are employed by job-shops for packaging, assembly, palletizing, and depalletizing [2]. Similar scheduling problems have arisen in the context of automated parking garages [3] and automated libraries [4]. In addition, the TRSP can be extended to multidimensional problems, namely, robots working on shelves, and multirobot problems like multirobot cooperation system [5, 6].

The dispatching of common rail dual AGVs is a typical TRSP in factories to quickly dispatch materials between tanks or automatically schedule express delivery in warehouses. Traditionally, this novel and specific problem is solved by simple heuristic algorithms, such as genetic algorithms (GA) and [7–10] particle swarm optimization (PSO) [11–13]. Due to the sheer scale of the problem, it is difficult for these algorithms to converge to the global optimal solution. Thus, it is urgent an efficient and feasible scheduling algorithm for common rail dual AGVs.

This paper establishes a model of the common rail dual AGVs and presents an efficient solution to the model. The remainder of this paper is organized as follows: Section 2 sets up a dual AGV scheduling model to minimize the running time of the AGV system; Section 3 presents a novel heuristic scheduling algorithm that dynamically plans material distribution tasks based on time windows; Section 4 compares our method with the benchmark method through experiments on a set of benchmark instances; Section 5 wraps up this work.

2. Model Construction

To model the common rail dual AGV system, the actual factory scheduling problem was abstracted and redefined as follows (Table 1).

The common rail dual AGV system has two AGVs, one rail, and two independent hangars. During the operation, the two AGVs automatically drive in opposite directions along the same rail. Each end of the rail leads to a hangar. The tanks are evenly arranged along one or both sides of the rail for material stacking. Multiple materials can be stacked in the same tank but can only be picked up from top to bottom.

Each AGV is preallocated by the system. The operation of each AGV includes three steps: picking, transporting, and placing along the rail. The AGV can only handle one material at a time. After completing all tasks, the AGV must return to its hangar to be reset for the next operation.

To avoid collisions, a safe distance must be maintained between the two AGVs. Suppose that the two AGVs operate at the same speed on the rail, which can be calculated at time t when passing by a tank. Then, the two AGVs take the same amount of time (on average), T, to pick up and put down the material (Figure 1).

After abstracting the actual problem, the objective is to design a dual AGV scheduling strategy that minimizes the times for the two AGVs to complete the assigned material tasks and return to their respective hangars.

Table 1: The symbol description.

Symbol	Meaning
$l1$	The length of tank interval
$l2$	The length of each tank
L	The slot length
W_i	Material, where i is the material number
CT_i	Tank, where i is the tank number
M	The total number of tanks
z_{Ai}	The current loading states of each AGV
O_{Ai}	The current directions of each AGV
m_{Ai}	The motion states of each AGV
C_i	The current states of each tank
CT_k^F	Serial number of the target tank
T_{temp}	Material W_i is temporarily stored in the left neighbor of the target tank
T_{clr}	The time it takes to empty the target tank
$\varphi(A_i)$	Whether AGV Ai carries compatible materials
d_{k1k2}	The distance between the two AGVs
t	The time to pass through a tank
T	The time of picking up and putting down materials
T_{total}	The running time of the entire AGV system
D	The length of each AGV journey

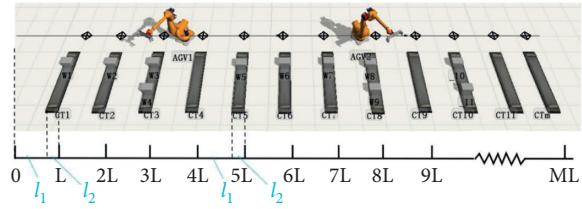


Figure 1: The abstract diagram of the dual AGV system.

2.1. Dual AGV Scheduling Model with Minimized Time

2.1.1. Motion States and Directions of AGVs. To clearly illustrate the coordinate change of AGVs, the rail was abstracted into a numerical axis with the leftmost end as the origin. It is assumed that the tanks are arranged at equal intervals. The length of each interval and the length of each tank are denoted as $l1$ and $l2$, respectively. Then, the total length of each slot, that is, each tank and its left interval, can be calculated by

$$L = l1 + l2. \tag{1}$$

The current loading states of the two AGVs can be described by the state variable z_{Ai} $(i = 1, 2)$:

$$z_{Ai} = \begin{cases} 0, & \text{AGV } Ai \text{ is unloaded,} \\ 1, & \text{AGV } Ai \text{ is loaded,} \\ 2, & \text{AGV } Ai \text{ is picking up materials,} \\ 3, & \text{AGV } Ai \text{ is putting down materials.} \end{cases} \tag{2}$$

The current directions of the two AGVs can be represented by a 0–1 decision variable O_{Ai} $(i = 1, 2)$:

$$O_{Ai} = \begin{cases} 0, & \text{AGV } Ai \text{ moves to the left,} \\ 1, & \text{AGV } Ai \text{ moves to the right.} \end{cases} \tag{3}$$

Taking the picking and placing processes as static state (through which the two AGVs avoid collision), the current motion states of the two AGVs can be illustrated by another 0–1 decision variable m_{Ai} $(i = 1, 2)$:

$$m_{Ai} = \begin{cases} 0, & \text{AGV } Ai \text{ is at rest,} \\ 1, & \text{AGV } Ai \text{ is in motion.} \end{cases} \tag{4}$$

2.1.2. States of Tanks. If a material is placed in its target tank, the material is compatible with the tank. Otherwise, the material is not compatible with the tank. Let C_i be the current state of the tank. If the tank belongs to the state $C_i = -1$, all materials have been delivered to the target tank and the two AGVs have returned to the hangars:

$$C_i = \begin{cases} 0, & \text{the tank is empty,} \\ 1, & \text{there are incompatible materials in the tank,} \\ -1, & \text{the tank is full of compatible materials.} \end{cases} \tag{5}$$

2.1.3. Tank Constraints. The materials are handled by the last-in, first-out principle. That is, each material is transported to the current tank to the target tank CT_k^F, where k is the serial number of the target tank. The destination slot must be empty or only compatible before the material is placed:

$$\text{CT}_k^F = 0 \text{ or } -1. \tag{6}$$

If $CT_k^F = 1$, there are incompatible materials. The number of noncompatible materials is denoted as j. Then, it is necessary to judge the direction O_{Ai} of the AGV. If $O_{Ai} = 0$, the

AGV temporarily places the target material W_{inow} in the right neighbor of the target tank CT_{k+1}^F. If $O_{Ai} = 1$, the material W_{inow} is temporarily stored in the left neighbor of the target tank CT_{k-1}^F, which can be recorded as $T_{\text{temp}} = T$.

Next, the AGV drives to CT_k^F and transports j incompatible materials to the left neighbor CT_{k-1}^F or the right neighbor CT_{k+1}^F without temporary storage of W_{inow} until $CT_k^F = 0$ or -1. The time it takes to empty the target tank is denoted as $T_{\text{clr}} = (2T + 2t)j$. Figure 2 illustrates the rules for material delivery.

2.1.4. Distribution Constraint. When being delivered to its destination tank, each material must be distributed by the specified AGV. Hence, an index $\varphi(A_i)$ was introduced to judge whether AGV A_i carries materials that are compatible with the target tank:

$$\varphi(A_i) = \begin{cases} 0, & \text{AGV } Ai \text{ carries compatible materials,} \\ 1, & \text{AGV } Ai \text{ carries incompatible materials,} \end{cases} \tag{7}$$

and, then, at any moment, make both $\varphi(A_i)$ values of the two AGVs zero:

$$\sum \varphi(A_i) = 0, \tag{8}$$

and this constraint ensures that each AGV is delivering the materials allocated to it at any time.

2.1.5. Safe Distance Constraint. The positions of AGV1 and AGV2 on the rail, denoted as X_{k1}^{A1} and X_{k2}^{A2}, can be, respectively, derived as

$$X_{k1}^{A1} = k1L, X_{k2}^{A2} = k2L, \tag{9}$$

where $k1$ and $k2$ are tank numbers of AGV1 and AGV2, respectively.

Then, the distance between A_1 and A_2 can be represented by d_{k1k2}. There is a phased conflict between the schedules of the two AGVs. To avoid directional conflict between A_1 and A_2, at least one slot length L must be maintained between them as the safe distance [14]:

$$d_{k1k2} = X_{k1}^{A1} - X_{k2}^{A2} = bL, \quad (b \in Z^+, b \geq 2). \tag{10}$$

Thus, the relationship between the locations of the two AGVs can be described as

$$k2 = k1 + b, \quad (b \in Z^+, b \geq 2). \tag{11}$$

2.2. Dual AGV Scheduling Model with Minimal Time. Under the above constraints, a dual AGV scheduling model can be established to minimize the time of material distribution. The paths of the AGVs can be described by an undirected graph $G = (\text{CT}, E)$, where $\text{CT} = \{\text{CT1}, \text{CT2}, \ldots, \text{CT}m\}$ is the set of tanks, and $E = \{E1, E2, \ldots, Ei\}$ is the set of intervals between tanks CT_i and CT_j. Then, the length of the journey of each AGV can be described as an integer multiple of slot length L:

$$E_{ij} = |i - j|L. \tag{12}$$

Suppose that an AGV changes direction for b times on the way to deliver materials. Let CTi and CTj be the start and end points of the AGV at the second change of direction B ($0 \leq \text{BB} \leq \text{bb}$), respectively. Then, the route of the AGV is denoted as E_{ij}^B. The travel length of the AGV can be expressed as

$$D = \sum_{b=0}^{B} E_{ij}^B. \tag{13}$$

Obviously, each AGV picks up the same amount of materials as those it delivers. Let n be the times of pick-ups (and also the times of placements) of the AGV, let t be the time to pass through a tank, and let T be the time of picking up and putting down materials. Then, the time for each AGV to complete its tasks and return to its hangar can be determined. The total time for AGV1 to complete its tasks and return to its hangar can be calculated by

$$T_{\text{total}\,1} = D_1 t + n_1 T, \tag{14}$$

where D_1 is the length of the journey of AGV1 and n_1 is the number of pick-ups or placements of AGV1.

The total time for AGV2 to complete its tasks and return to its hangar can be calculated by

$$T_{\text{total}\,2} = D_2 t + n_2 T, \tag{15}$$

where D_2 is the length of the journey of AGV2 and n_2 is the number of pick-ups or placements of AGV2.

Then, the total running time of the entire AGV system equals the longer running time between the two AGVs:

$$T_{\text{total}} = \max\{T_{\text{total}\,1}, T_{\text{total}\,2}\}. \tag{16}$$

In this way, a dual AGV scheduling model can be established aiming to let the two AGVs complete their tasks and return to their respective hangar as early as possible, that is, to minimize the running time of the entire AGV system:

$$\min T_{\text{total}}$$
$$\text{s.t.} \begin{cases} \text{CT}_k^F = 0 \text{ or } -1, \\ \sum \varphi(A_i) = 0, \\ T_{\text{temp}} = T, \\ T_{\text{clr}} = (2T + 2t)j, \\ d_{k1k2} = X_{k1}^{A1} - X_{k2}^{A2} = bL, \ (b \in Z^+, b \geq 2), \\ k2 = k1 + b, \ (b \in Z^+, b \geq 2), \\ E_{ij} = |i - j|L, \\ D = \sum_{b=0}^{B} E_{ij}^B, \\ T_{\text{total}\,1} = D_1 t + n_1 T, \ T_{\text{total}\,2} = D_2 t + n_2 T, \\ T_{\text{total}} = \max\{T_{\text{total}\,1}, T_{\text{total}\,2} \end{cases} \tag{17}$$

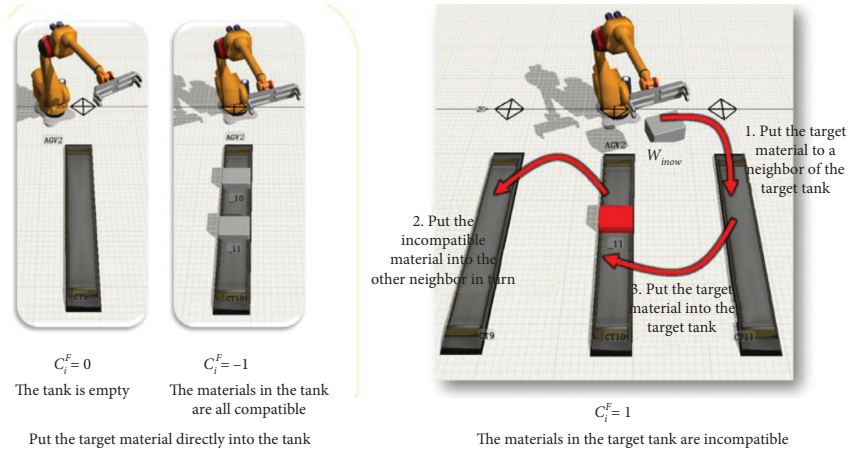


Figure 2: The rules for material delivery.

3. Scheduling Algorithm

3.1. Time Windows Model. Let r_i^j be the j-th occupied time window on the i-th tank, and let f_m^n be the n-th idle time window on the m-th tank. At each moment t, AGV1 and AGV2 occupy the time window where a tank is located. Then, the time $T_{\text{arr}}(f_{k0}^p, f_{kF}^q)$ to move from the current tank $k0$ of the material to the target tank kF can be calculated by

$$T_{\text{arr}}\left(f_{k0}^p, f_{kF}^q\right) = L\left(f_{k0}^p\right) + t_{k0kF} + t_\Delta, \tag{18}$$

where $L(f_{k0}^p)$ is the time to reach the starting tank $K0$ in a free time; t_{k0kF} is the time to move from the starting tank $K0$ to the end tank KF without considering safe distance; t_Δ is the additional time to move from $K0$ to KF incurred by the location and motion states of the other AGV [15].

3.2. Adding Safe Distance. Under the safe distance, the tank r_i^j where AGV is located and its left neighbor r_{i-1}^j and right neighbor r_{i+1}^j are occupied simultaneously; moreover, the other AGV cannot appear in this section with this AGV at the same time. Then, the current state v_i of the current slot can be depicted as [16]

$$v_i = \begin{cases} 0, & \text{the tank is occupied and the AGV is in the } i^{\text{th}} \text{ or } i \pm 1^{\text{th}} \text{ tank,} \\ 1, & \text{the tank is free and the AGV is not in the } i^{\text{th}} \text{ or } i \pm 1^{\text{th}} \text{ tank.} \end{cases} \tag{19}$$

For every time t, the state of each slot needs to be changed with the location of the AGV. Figure 3 explains the occupation state of the time window.

3.3. Heuristic Scheduling Algorithm for Dynamic Planning of Material Distribution Tasks Based on Time Windows (DPTW)

3.3.1. Workflow

Step 1. Find the shortest time delivery sequence for the independent delivery state of each AGV by the genetic algorithm (GA)

(1) Parameter initialization
Initialize the selection, crossover, and mutation probabilities according to the performance of the AGVs.

(2) Encoding
Set up a coding string of the length n_{Ai} and values between 1 and n_{Ai}, where n_{Ai} is the total number of materials assigned to A_i.

(3) Generating the initial solution
Randomly generate a coded string of length n_{Ai}; that is, randomly arrange the order of material delivery, as the initial solution.

(4) Fitness calculation
Take the penultimate time required to complete the local materials in the current order ($1/T_{\text{total}}$) as the fitness function. The reciprocal value of the fitness is negatively correlated with the time cost and positively with the fitness of each individual.

(5) Population selection
Use the Tournament Selection Method (TSM) [17, 18] to remove some individuals from the population, and select one of the best fit individuals to enter the next population. Repeat this process until the new population equals the original population in size.

(6) Crossover and mutation
Implement the reverse mutation operator [19] and exchange the order of material distribution.

(7) Evolution

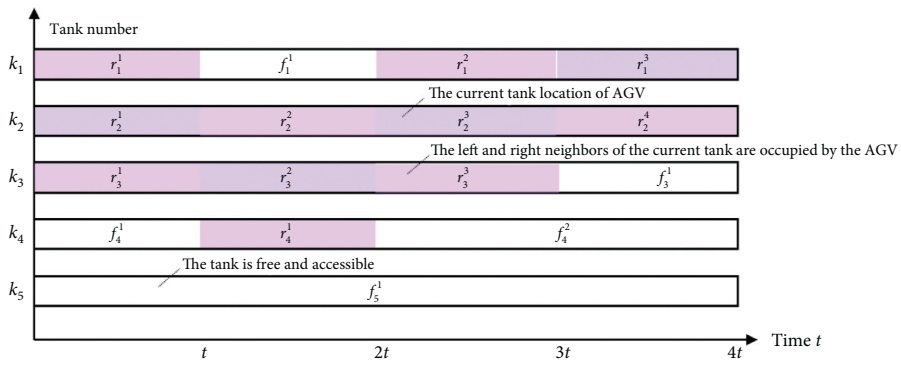


Figure 3: The occupation state of the time window.

Obtain the new population and repeat Step (4) until the 10,000-th iteration to get the best fit individuals.

(8) Decoding

Decode the genetic sequence of the order of local material distribution for each AGV. Store the results of AGV1 and AGV2 in one-dimensional (1D) arrays p_0^1 and p_0^2, respectively, which reflect the minimum running time in the case of independent interaction.

Step 2. Order distribution based on the initial minimum running time

As shown in Figure 4, distribute the first task to AGV1 and AGV2 according to the initial scheduling plan p_0^1, obtain the time window of the first task of each AGV, and judge whether there is an overlapping section RP(S) between the routes of the two AGVs.

If yes, let the AGV closer to the destination tank continue with its delivery, and halt the other AGV until the overlapping section is cleared.

If not, repeat Step 2 after the delivery to assign the second task as per the initial plan until an overlapping section arises. Once an overlapping section occurs, go to Step 3 for dynamic planning of the next task.

Step 3. Dynamic planning of the next task based on time windows

Assign the next task $\mathrm{RW}_{i+1}(K0, \mathrm{KF}, t_{i+1})$ to an AGV after it has placed the current task material and entered the unloaded state $z_{Ai} = 0$.

(1) Case 1: the two AGVs do not end the current task in a synchronized manner.

If $A1$ terminates the current task but $A2$ continues with the current task, dynamically plan the next task of $A1$ in combination with the time window of $A2$; conversely, dynamically plan the next task of $A2$ in combination with the time window of $A1$.

Firstly, search for the nearest tanks on the left and right of the current slot of $A1$, and denote them as optional task points RW1 and RW2, following the greedy strategy. Take the number of the next materials in the original plan p_0^1 as alternative task point RW3.

Next, obtain the time window of the alternative tank separately, and plan the route and time to reach the starting and target tanks.

Finally, compare $A2$ with the time windows of the current target tank, measure the set of overlapping sections RP1(SS), RP2(SS), and RP3(SS), and take the RW corresponding to min{RP1(SS), RP 2(SS), RP3(SS)} as the next task.

(2) Case 2: the two AGVs end the current task in a synchronized manner.

If $A1$ and $A2$ terminate the current task at the same time and if the distribution of the next task to one of them affects that of the other, then $A1$ and $A2$ search for the nearest tanks on the left and right of their current slots and denote them as optional task points $(\mathrm{RW}_1^{A1}, \mathrm{RW}_2^{A1})$ and $(\mathrm{RW}_1^{A2}, \mathrm{RW}_2^{A2})$ following the greedy strategy and also take the number of the next materials in the original plan p_0^1 as alternative task points RW_3^{A1} and RW_3^{A2}.

Next, obtain the time windows of the three alternative task points of $A1$ and $A2$, respectively, and plan the route and time for each AGV to reach the starting and target tanks.

Finally, establish the set of overlapping sections, and take the least overlapping combinations as the next task of the two AGVs.

Step 4. Determine the motion state for the next task

(1) Case 1: the two AGVs do not end the current task in a synchronized manner.

Let an AGV receive the next task as a collaborator, compare the time windows for transporting the current materials, and choose to stop or stay away for cooperation with the AGV still engaged in material delivery at the end of the overlapping section, according to its location k_i and whether k_{i-2}/k_{i+2} is busy at the next moment:

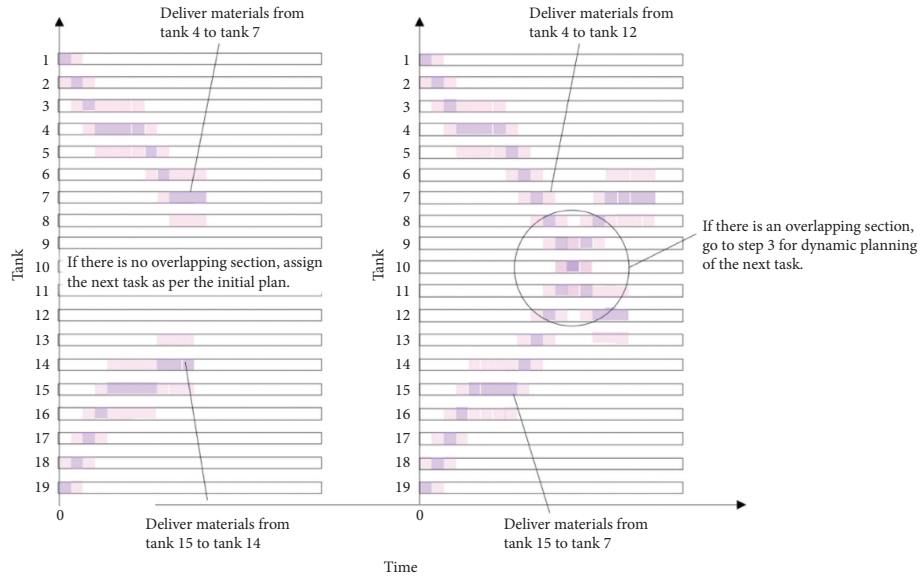


Figure 4: The handling of overlapping section.

$$\begin{cases} v_{ki} = 1, v_{ki-2} \| v_{ki+2} = 1, & \text{move left/right,} \\ v_{ki} = 1, v_{ki+2} = 0, & \text{move to the/end tank,} \\ v_{ki} = 0, v_{ki-2} \| v_{ki+2} = 1, & \text{Stop.} \end{cases} \quad (20)$$

(2) Case 2: the two AGVs end the current task in a synchronized manner.

Since there is an overlapping section between the two AGVs, choose the collaborator based on their distances from the starting tank to the target tank. Make the AGV farther away from its target tank as the collaborator, and choose to stop or stay away. The determination of the motion state for the next task is illustrated in Figure 5.

Step 5. Repeating Steps 3 and 4 until the end of material distribution

Repeat Steps 3 and 4 until all material tasks are completed, return to the hangar, count the transport time of each task, and obtain the shortest time $\min T_{\text{total}}$.

(2) Running time optimization

To minimize the overall time cost of placing the target materials, the two AGVs can cooperate with each other to empty their target tank under the following conditions.

One AGV can issue an emptying request to the other AGV if their distance d_{k1k2} satisfies

$$d_{k1k2} \geq \left(\frac{2T}{t} + 2\right)L. \quad (21)$$

Suppose that AGV1 is the one issuing the request. Let k_2 and k_{A2} be the current position and starting tank of AGV2, respectively; let k_{A1} be the target tank of AGV1 ($k_2 < k_{A1} < k_{A2}$). If AGV2 is unloaded currently and moves in the same direction (right) as AGV1, then AGV2 will accept the request under the following conditions:

$$\begin{cases} O_{A1} = O_{A2} = 1, \\ z_{A2} = 0, \\ k_2 < k_{A1} < k_{A2}. \end{cases} \quad (22)$$

Suppose that AGV2 is the one issuing the request. Let k_1 and k_{A1} be the current position and starting tank of AGV1, respectively; let k_{A2} be the target tank of AGV2 ($k_1 < k_{A2} < k_{A1}$). If AGV1 is unloaded currently and moves in the same direction (right) as AGV2, then AGV1 will accept the request under the following conditions:

$$\begin{cases} O_{A1} = O_{A2} = 1, \\ z_{A1} = 0, \\ k_1 < k_{A2} < k_{A1}. \end{cases} \quad (23)$$

Figure 6 explains the process of dual AGVs collaborating with each other.

3.3.2. Algorithm Flow. Figure 7 presents the workflow of the proposed algorithm.

4. Computational Results in Industrial Implementation

To verify its performance, the proposed algorithm was applied to minimize the overall time of the common rail dual AGVs in handling 16 actual orders from a factory. For comparison, the classical GA was also tested to test whether our algorithm can outshine the common baseline search algorithm. Tables 2 to 17 provide the details of the 16 orders.

As shown in Table 18, our algorithm greatly shortened the overall distribution time of the GA, an evidence to its rationality and high performance. Take order 7 as an

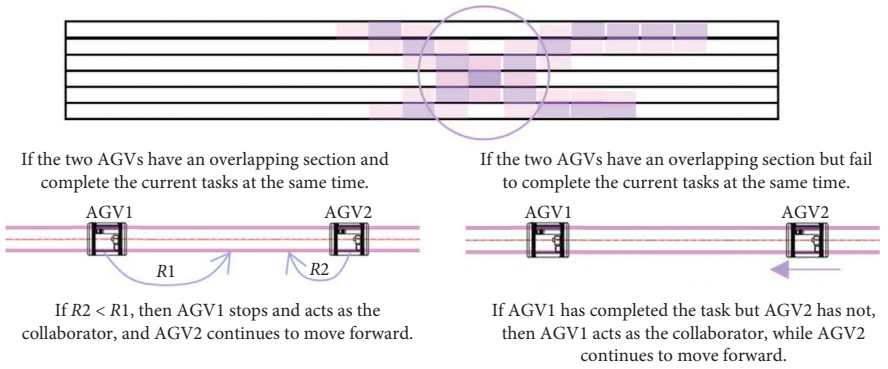


Figure 5: The determination of the motion state for the next task.

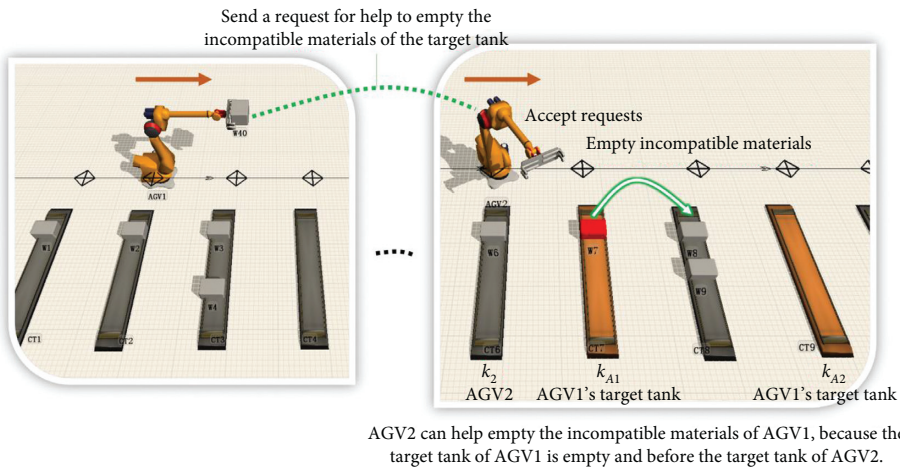


Figure 6: The collaboration between the two AGVs.

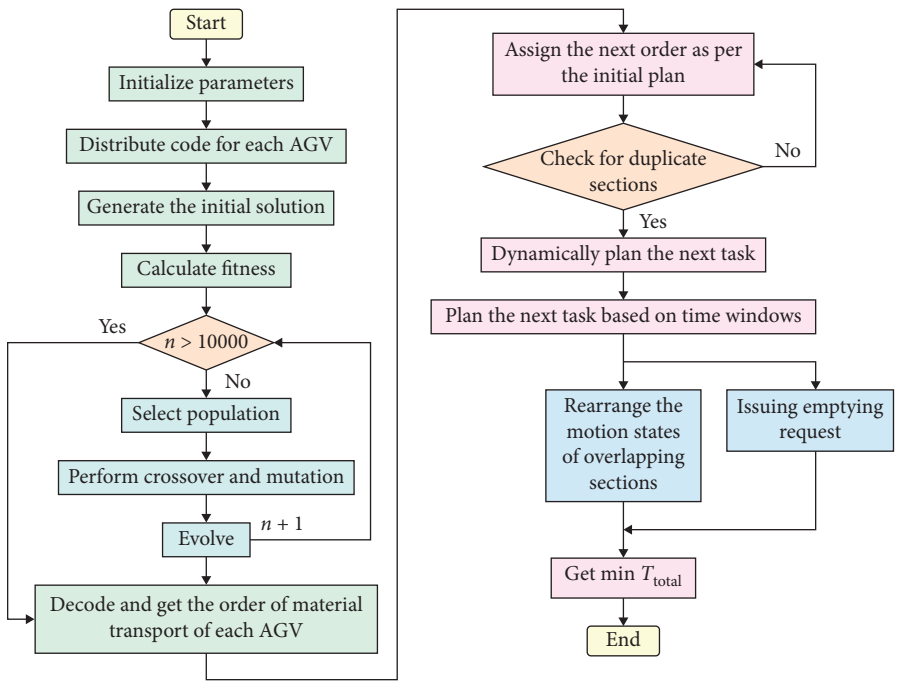


Figure 7: The workflow of the proposed algorithm.

Table 2: Instance 1 (number of tanks: 20).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	20
2	1	2	3
3	2	4	11
4	1	7	8
5	1	9	7
6	1	10	9
7	2	11	12
8	1	14	7
9	2	16	5
10	2	17	8
11	1	18	1
12	2	19	12
13	1	20	13

Table 3: Instance 2 (number of tanks: 20).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	10
2	2	3	9
3	1	4	4
4	2	5	1
5	1	6	8
6	2	7	2
7	2	10	18
8	1	11	16
9	1	12	5
10	1	13	17
11	2	15	15
12	2	16	14
13	1	17	8
14	2	18	18
15	1	19	8
16	2	20	8

Table 4: Instance 3 (number of tanks: 20).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	1	8
2	2	2	14
3	1	3	6
4	1	3	18
5	1	5	15
6	2	7	1
7	2	9	11
8	2	9	19
9	2	10	2
10	2	10	4
11	1	10	5
12	2	12	20
13	1	13	10
14	2	18	12
15	1	19	7

Table 5: Instance 4 (number of tanks: 20).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	13
2	2	3	6
3	2	5	13
4	1	7	2
5	1	9	4
6	2	10	5
7	1	11	12
8	2	11	4
9	2	11	4
10	1	12	16
11	2	12	5
12	1	16	11
13	1	16	17
14	1	17	17
15	2	17	16
16	1	18	10
17	1	19	19

Table 6: Instance 5 (number of tanks: 20).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	1	18
2	2	2	2
3	1	2	14
4	1	3	17
5	2	4	20
6	2	7	8
7	2	8	13
8	2	8	16
9	1	8	9
10	1	9	12
11	2	10	9
12	2	10	5
13	2	11	12
14	1	13	11
15	1	13	1
16	1	13	12
17	2	15	20
18	1	18	12
19	1	20	7

Table 7: Instance 6 (number of tanks: 25).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	1	2
2	2	2	11
3	2	3	10
4	2	7	10
5	1	9	5
6	1	10	16
7	1	11	20
8	2	12	14
9	1	13	17
10	2	14	21

Table 7: Continued.

Material serial number	Allocated AGV	Current tank	Target tank
11	1	15	24
12	2	16	14
13	2	17	16
14	2	18	21
15	1	19	16
16	2	20	6
17	2	22	19
18	2	23	6
19	1	25	13

Table 8: Instance 7 (number of tanks: 25).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	1	18
2	2	3	4
3	1	4	3
4	1	4	10
5	1	7	6
6	2	7	10
7	2	7	16
8	1	8	13
9	2	9	15
10	2	10	2
11	1	11	10
12	2	12	2
13	1	12	17
14	1	13	17
15	2	14	15
16	1	15	14
17	2	18	18
18	2	18	3
19	1	21	24
20	2	22	20

Table 9: Instance 8 (number of tanks: 25).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	3
2	1	2	3
3	2	3	2
4	1	4	7
5	1	5	19
6	2	7	17
7	2	8	12
8	2	9	4
9	1	10	3
10	1	12	20
11	2	13	24
12	1	15	4
13	2	16	4
14	2	17	2
15	1	18	11

Table 9: Continued.

Material serial number	Allocated AGV	Current tank	Target tank
16	1	19	1
17	2	20	7
18	2	21	15
19	2	22	23
20	1	23	23
21	1	24	16
22	1	25	25

Table 10: Instance 9 (number of tanks: 25).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	3	19
2	1	5	4
3	2	6	23
4	2	7	5
5	1	9	7
6	1	10	20
7	1	11	2
8	2	14	22
9	1	14	9
10	2	15	4
11	1	15	18
12	2	17	7
13	2	20	20
14	2	20	7
15	1	21	21
16	2	22	25
17	1	22	16
18	1	24	5

example to illustrate the rationality and effectiveness of our algorithm. Figure 8 explains the next routes of the two AGVs in the event of a conflict [20].

By our algorithm, the final results of AGV operations and order of material distribution were obtained. Then, the distribution order was randomly changed to obtain new solutions to AGV running time. After the changes, the running time of the AGVs and the relevant solutions were observed to see if our algorithm could lead to the best solution.

Figure 9 presents the randomly changed order and the new running time of the 30 orders in Tables 2 to 17. It can be seen that all the new solutions were greater than the original running time, indicating that the scheduling plan of our algorithm is the best possible solution.

The next step is to verify whether any solution in the 16 solutions is the optimal solution. Taking order 7 as the object, the order of material distributions was changed randomly twice, three times, and four times, respectively. The running time of each change is shown in Figure 10.

Table 11: Instance 10 (number of tanks: 25).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	3	14
2	2	5	14
3	1	8	10
4	1	10	18
5	1	11	14
6	2	12	7
7	2	12	19
8	2	14	24
9	2	15	5
10	2	16	4
11	1	16	24
12	2	17	14
13	1	17	18
14	1	19	11
15	1	19	14
16	1	19	2
17	2	19	12
18	1	20	17
19	1	20	7
20	1	21	16
21	2	21	3
22	2	22	17
23	1	24	25

Table 12: Instance 11 (number of tanks: 30).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	1
2	1	2	27
3	1	2	1
4	2	5	13
5	1	6	28
6	2	6	30
7	1	8	10
8	2	9	29
9	1	10	19
10	1	11	9
11	2	11	7
12	1	13	12
13	1	13	25
14	2	13	24
15	1	15	26
16	2	16	7
17	2	17	23
18	2	17	27
19	2	18	10
20	2	18	4
21	2	24	14
22	2	27	5
23	1	28	23
24	1	29	12

Table 13: Instance 12 (number of tanks: 30).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	5
2	2	2	2
3	2	3	12
4	1	7	28
5	1	8	13
6	1	9	21
7	2	9	26
8	2	10	16
9	1	11	20
10	1	11	9
11	1	14	1
12	2	16	1
13	2	18	6
14	2	24	25
15	2	24	13
16	2	25	20
17	1	26	3
18	1	27	15
19	2	27	21
20	2	29	11
21	1	30	28
22	2	30	25

Table 14: Instance 13 (number of tanks: 30).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	1	14
2	2	2	6
3	2	3	25
4	1	4	28
5	2	5	15
6	1	6	24
7	1	7	3
8	2	7	17
9	2	8	9
10	1	12	11
11	1	12	7
12	1	12	5
13	1	14	20
14	1	15	15
15	1	15	1
16	2	16	15
17	2	17	9
18	2	18	28
19	2	18	12
20	2	19	6
21	1	19	7
22	1	21	26
23	2	24	7
24	2	25	14
25	2	26	3
26	2	26	3
27	1	27	29

Table 15: Instance 14 (number of tanks: 30).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	1	27
2	1	1	8
3	1	1	28
4	1	3	11
5	2	3	20
6	1	4	18
7	1	4	20
8	2	6	9
9	2	8	16
10	2	8	27
11	1	9	25
12	1	10	22
13	2	10	13
14	2	11	18
15	1	11	10
16	1	11	1
17	1	13	16
18	2	13	20
19	1	14	9
20	2	14	1
21	1	16	22
22	1	16	20
23	1	16	10
24	2	17	14
25	1	19	28
26	1	19	6
27	2	19	8
28	1	20	23
29	2	24	11
30	2	24	25
31	2	26	30
32	2	27	19
33	2	29	27

Table 16: Instance 15 (number of tanks: 30).

Material serial number	Allocated AGV	Current tank	Target tank
1	2	3	22
2	1	4	29
3	2	4	1
4	1	4	29
5	2	6	21
6	2	7	16
7	1	7	5
8	1	8	5
9	1	8	25
10	2	11	1
11	1	11	1
12	2	12	24
13	1	13	30
14	2	14	4
15	1	14	19
16	2	15	24
17	1	16	9
18	1	16	2
19	1	17	24
20	1	20	23
21	1	20	12
22	1	22	21
23	2	22	8
24	2	23	25
25	2	26	9
26	1	26	5
27	2	29	16

Table 17: Instance 16 (number of tanks: 40).

Material serial number	Allocated AGV	Current tank	Target tank
1	1	4	15
2	2	4	18
3	1	7	19
4	1	7	18
5	2	11	28
6	2	11	27
7	2	12	2
8	2	12	32
9	2	12	24
10	2	15	26
11	1	15	24
12	2	19	26
13	1	21	28
14	2	22	28
15	1	23	4
16	1	23	25
17	2	24	16
18	2	24	28
19	1	25	29
20	2	25	15
21	1	26	25
22	2	26	32
23	2	26	36
24	1	28	16
25	1	28	16
26	1	29	16
27	1	30	28
28	1	30	37
29	1	31	23
30	2	32	15
31	2	36	8
32	1	37	13
33	1	38	20
34	2	38	28
35	1	39	20
36	2	39	22

It can be seen that, after the changes in material distribution order, the final new solutions were greater than the benchmark value of 785. Hence, the proposed algorithm can always obtain the optimal solution.

Table 18: The computing results (unit: s).

Instance	DPTW	GA
Instance 1	550	755
Instance 2	620	805
Instance 3	595	780
Instance 4	650	855
Instance 5	920	1365
Instance 6	665	915
Instance 7	785	1190
Instance 8	995	1375
Instance 9	910	1120
Instance 10	975	1235
Instance 11	1355	1860
Instance 12	1075	1665
Instance 13	1485	1910
Instance 14	1875	2350
Instance 15	1595	2130
Instance 16	2295	3345

Figure 8: The illustration of conflict avoidance in order 7.

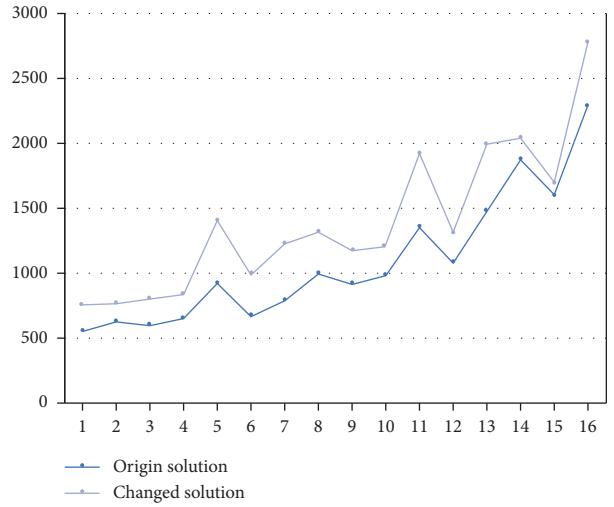


Figure 9: The running time after random changes of material distribution order.

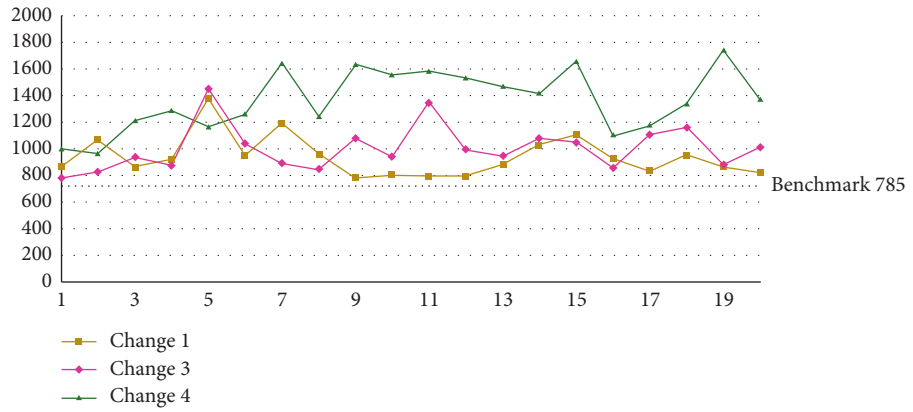


Figure 10: The running time after three types of changes in material distribution order.

5. Conclusions

This paper models and solves the problem of common rail dual AGVs scheduling in a novel way. The proposed solving algorithm for the problem consists of five steps, each of which is described in detail, and provides the scheduling plans considering the various situations of AGV motions. The performance of our algorithm was fully verified through experiments. The results show that our algorithm can solve the problem more efficiently and is more effective than the GA in all cases. Even if the solution structure was partly changed by random, our algorithm could still find the optimal solution very quickly. The future research will extend proposed method to multi-AGV cooperative scheduling or develop even more efficient online dynamic programming algorithms for this problem.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] G. Erdoğan, M. Battarra, and G. Laporte, "Scheduling twin robots on a line," *Naval Research Logistics (NRL)*, vol. 61, no. 2, pp. 119–130, 2014.

[2] P. G. Ranky, "Collaborative, synchronous robots serving machines and cells," *Industrial Robot: An International Journal*, vol. 30, no. 3, 2003.

[3] A. Mathijssen and A. J. Pretorius, "Verified design of an automated parking garage," in *Proceedings of the International Workshop on Parallel and Distributed Methods in Verification*, pp. 165–180, Bonn, Germany, August 2006.

[4] K. E. Dimitri, F. D. Gallo, J. E. Kulakowski, R. J. Means, J. L. Thrall, and D. J. Winarski, "Automated data storage library dual picker interference avoidance," U.S. Patent 6,038,490, 2000.

[5] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 376–386, 2005.

[6] T. Arai, E. Pagello, and L. E. Parker, "Advances in multi-robot systems," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 655–661, 2002.
[7] M. Gemeinder and M. Gerke, "GA-based path planning for mobile robot systems employing an active search algorithm," *Applied Soft Computing*, vol. 3, no. 2, pp. 149–158, 2003.
[8] H. Özdemir, R. Sever, and Ö. Polat, "GA-based optimization of SURF algorithm and realization based on Vivado-HLS," *Traitement du Signal*, vol. 36, no. 5, pp. 377–382, 2019.
[9] Y. Lu, X. Pei, C. Zhang, H. Luo, B. Liu, and Z. Ma, "Design of multimodal transport path optimization model and dual pheromone hybrid algorithm," *Journal Européen des Systèmes Automatisés*, vol. 52, no. 5, pp. 477–484, 2019.
[10] A. P. Gopi, V. Lakshman Narayana, and N. Ashok Kumar, "Dynamic load balancing for client server assignment in distributed system using genetical algorithm," *Ingénierie des Systèmes d'Information*, vol. 23, no. 6, pp. 87–98, 2018.
[11] W. Zhang, X. L. Zhang, and Y. Li, "Path planning for intelligent robots based on improved particle swarm optimization algorithm," *Journal of Computer Applications*, vol. 34, no. 2, pp. 510–513, 2014.
[12] D. Wu, "Multi-objective Decision-making of new retailing terminals based on particle swarm optimization and genetic algorithm," *Journal Européen des Systèmes Automatisés*, vol. 52, no. 6, pp. 607–615, 2019.
[13] Z. Li, Y. Zhou, and R. Bao, "An image classification method based on optimized fuzzy bag-of-words model," *Traitement du Signal*, vol. 36, no. 3, pp. 239–244, 2019.
[14] K. L. Peng, "Safe following distance warning system and method for a vehicle," U.S. Patent Application 12/057,725, 2009.
[15] N. Smolic-Rocak, S. Bogdan, Z. Kovacic, and T. Petrovic, "Time windows based dynamic routing in multi-AGV systems," *IEEE Transactions on Automation Science and Engineering*, vol. 7, no. 1, pp. 151–155, 2009.
[16] D. I. Ferguson and D. Dolgov, "Safely navigating on roads through maintaining safe distance from other vehicles," U.S. Patent 8,725,342, 2014.
[17] O. Andrzej and K. Stanislaw, "A new constraint tournament selection method for multicriteria optimization using genetic algorithm," *Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No. 00TH8512)*, vol. 1, pp. 501–508, 2000.
[18] M. Vard, M. Neyestan, and A. Ghorbanian, "Supplier selection and order allocation problem modeling with the aim of comparing incremental discounts versus wholesale discounts by using GA and NSGA algorithms," *Journal Europeen des Systemes Automatises*, vol. 52, no. 1, pp. 23–34, 2019.
[19] T. P. Hong and H. S. Wang, "A dynamic mutation genetic algorithm," *IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems (Cat. No. 96CH35929)*, vol. 3, pp. 2000–2005, 1996.
[20] R. H. Möhring, E. Köhler, E. Gawrilow, and B. Stenzel, "Conflict-free real-time AGV routing," in *Proceedings of the Operations Research Proceedings*, pp. 18–24, Tilburg, Netherlands, July 2005.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6656222, 13 pages
https://doi.org/10.1155/2020/6656222

Research Article

Transfer Learning and Identification Method of Cross-View Target Trajectory Utilizing HMM

Long Liu, Le Yang, and Jie Ding

The Faculty of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China

Correspondence should be addressed to Long Liu; liulong@xaut.edu.cn

Received 2 October 2020; Revised 16 November 2020; Accepted 9 December 2020; Published 24 December 2020

Academic Editor: Hussein Abulkasim

The behavior identification of the target trajectory is one of the important issues in space behavior analysis. Since the target trajectory model obtained from a fixed view cannot be adapted to the change of the observation perspective, it needs to be retrained when being faced with a new view, which leads to a great amount of increment in application cost. This study proposes a hidden Markov model (HMM) based on the cross-view transfer learning and the recognition method that firstly constructs a linear mapping relationship between the observation matrices of the source and target view utilizing the domain trajectory of the HMMs and obtains the observation matrix parameters of the target domain through the mapping system. Secondly, the transfer probability of the source domain is further optimized to obtain the target domain of the HMM and to identify the behavior of the target domain trajectory utilizing a small number of samples from the view of the target domain. The experimental results denote that the proposed method could effectively realize the identification of the trajectory behavior utilizing a small sample size in the target domain and would greatly reduce the application cost of the identification of the cross-view target trajectory.

1. Introduction

Target trajectories reveal important information on target behavior, and the identification of the trajectory is an important part pertinent to understanding an event and behavior analysis. In real applications, a video surveillance system often composes of multiple cameras located at different angles. The motion trajectory of the same target is displayed differently from diverse perspectives. The training of the trajectory model relies on samples from a fixed perspective. When the perspective alters, new samples need to be collected and retrained, leading to the excessively high training cost for the trajectory model, which would greatly raise the application cost of the identification of the trajectory behavior. This is not conducive to its promotion. Pointing out that it has a high application value and practical importance to research on the behavior identification method of the low-cost cross-view target trajectory.

Hidden Markov model (HMM) [1–8] could effectively model time-series signals and is a powerful tool to model the target tracking. Qiming and Cheng propose that trajectory modeling is executed according to the trajectory coordinate sequence [1]. In this method, the trajectory sequence is firstly clustered, and then, the corresponding HMM trajectory model is attained utilizing the training of the diverse categories. Dapeng et al. suggest that the HMM is utilized to model the trajectory of the enemy submarine and to predict the trajectory of the target to provide operational information support for warship operations [2]. Hervieu et al. propose that the curvature and velocity of the coordinate points are utilized for the HMM trajectory modeling, and the trajectory characteristics are invariant to translation, rotation, and scaling [3]. Qian and Lau suggest that a layered HMM is proposed to model the continuous target trajectory in a multilayer scene that leads to both a low-level HMM to model the single-layer trajectory and a high-level HMM connecting the multilayer trajectory to form the target trajectory model [4].

On the other hand, transfer learning has gradually become an important research area in machine learning in recent years since it mainly has resolved the problem

related to the learning of the cross-domain knowledge and could realize the model migration when the data of the target domain is unlabeled and the data of the source domain is called a multisource. Moreover, it has a high theoretical value in many practical problems. Transfer learning, which relaxes two basic assumptions in traditional machine learning, is a new method in the field of machine learning. Its main characteristics could be summarized as follows: (1) there must be enough samples to learn for a good classification model. (2) Both the training samples were utilized for learning, and the new test samples meet the conditions of independence and identical distribution properties. Some scholars have proposed various HMM transfer methods for some cross-domain applications. For instance, Van Kasteren et al. propose that a parameter transfer learning method utilizes the HMM that constructs the sensor mapping relationship according to the location information of the spatial structure and then finds out the parameters of the HMM utilizing the EM algorithm [9]. Zheng et al. propose that, it takes the received real-time signal intensity of mobile devices as the observed value to construct the HMM positioning model and realizes the transfer of it at different times in the same position [10]. Bingtao et al. suggest that giving weight to the sample data of the source domain by calculating the similarity between the source domain and the target domain sample data, then finds out the HMM parameters in the weighted data set, and improves the learning algorithm to realize the instance-based HMM transfer learning [11]. Similar research studies conducted by Kim et al. propose that an HMM transfer method utilizing the maximum a posterior (MAP) and maximum likelihood linear regression (MLLR) is applied to speech and text recognition [12–15]. Some empirical research studies related to HMM are Fei et al. in [16, 17].

As a contribution to this research area, to solve the low-cost application problem of the HMM trying to determine the trajectory behavior in different perspectives, this manuscript proposes an HMM-based transfer learning and recognition method for the behavior of the target trajectory where the cases consist of the perspective samples of the sufficient source domain, and the perspective samples of few marked target domain exist. This method achieves the purpose of transfer learning and recognition by transferring the trajectory behavior model in the source perspective domain and optimizing a small number of samples in the target domain, which provides an effective way to resolve the mentioned problem.

2. The Proposed Model: Trajectory Modeling and Simulation

2.1. Target Trajectory Model Based on the HMM. Hidden Markov model (HMM) is a probabilistic model describing the process of time series, which can be described by five components denoted by a quintuple $\lambda = (N, M, A, B, \text{and } \pi)$. Moreover, its simplified form is represented by $\lambda = (A, B, \text{and } \pi)$:

(1) N indicates the number of hidden states in the model. All states in the model are interconnected, and any state can be reached from other states.

(2) M represents the observation symbol of each hidden state in the model, i.e., the number of observation states.

(3) A is called the probability distribution of state transition. $A = \{a_{ij}\}$, where, $a_{ij} = p(Q_{t+1} = j | Q_t = i)$, $0 \le a_{ij} \le 1$, and $\sum_{i=1}^{N} a_{ij} = 1, 1 \le i, j \le M$. Q_t is called the hidden state at time t.

(4) B is called the probability distribution of the observable state in the hidden state at moment j, $B = \{b_{ik}\}$, where, $b_{ik} = p(Q_t = k | Q_t = i)$ and $1 \le i \le N, 0 \le k \le M$. Q_t is called the observable state at time t.

(5) π is called the distribution of the initial state $\pi = \{\pi_i\}$, where $\pi_i = p(Q_0 = i), 1 \le i \le N$.

According to the basic definition of the HMM, this manuscript firstly conducts the HMM modeling for the target trajectory. The direction angle of the trajectory of the target in a unit of time is utilized as the observation characteristic. The sequence of the angle representing the trajectory information is called the observable state vector. The hidden state is called the transfer characteristic of the angle change of the target trajectory. The angular direction of the target trajectory is computed by

$$\varphi_t = \tan^{-1}\left(\frac{y_t - y_{t-1}}{x_t - x_{t-1}}\right), \tag{1}$$

where (x_t, y_t) and (x_{t-1}, y_{t-1}) are called the target positions at time t and $t-1$, respectively. According to φ_t, the observed value Q_t is attained by discretizing the 24-direction freeman code depicted in Figure 1. The parameters of the trajectory of the HMM are obtained by training.

The trajectory presented in Figure 2 is utilized for modeling. The coordinate sequence of the target trajectory is presented in Table 1. The angle characteristic is computed by (1), and its discretization is presented in Table 2. While the number of the hidden states called N is set to 4, that of observation called M is set to 24. The probability vector of the initial state denoted by π is defined as uniform distribution. The transfer probability and the probability distributions of the observation are initialized randomly. The trajectory samples above are trained by the Baum–Welch algorithm to attain the parameters of the initial distribution π, transfer probability matrix A, and observation probability matrix B, respectively, that are presented in Table 3.

2.2. Simulation of the Target Trajectory. There are two important assumptions in the HMM, which are expressed as follows: (1) homogeneous Markov Chain hypothesis, that is, the hidden state at any time only depends on its previous hidden state. (2) Observation independence hypothesis, that is, the observed state at any time only depends on the hidden state at the current time.

As presented in Figure 3, the hidden Markov Chain based on the above hypothesis is determined by the initial

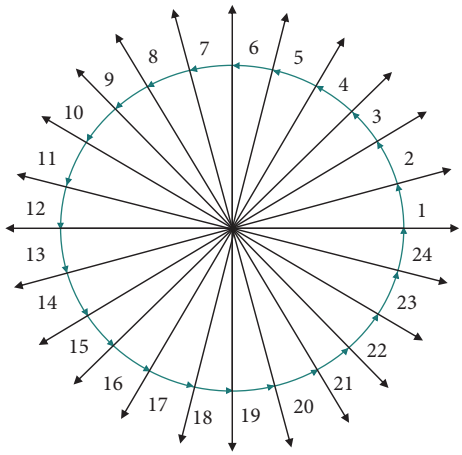


Figure 1: 24-direction freeman code.

(a) (b)

Figure 2: Target behavior trajectory. (a) The tracked target trajectory is overprinted on the background image. (b) A sequence of the tracked target trajectories.

Table 1: Centroid coordinate sequence of the target trajectory.

	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$	$P9$	$P10$
X	280	278	275	272	270	269	266	262	258	252
Y	155	157	160	169	169	170	173	177	182	189

state probability vector and the state transfer probability matrix A. Hence, the result of which is to generate an unobservable hidden state sequence. The observation probability matrix B and the hidden state sequence are combined to determine the way of generating an observable sequence.

Given the HMM model in the form of $\lambda = (N, M, A, B, \text{and } \pi)$, the observation sequence $O = O_1, O_2, \ldots, O_T$ could be generated by the algorithm steps presented in Table 4. According to the initial coordinates of the trajectory and the sequence of the observed values, the coordinates of the trajectory sequence could be uniquely determined.

The trajectory of the HMM is trained and simulated under four different circumstances, and the results are presented in Figure 4.

2.3. Statistical Analysis of the Trajectory Characteristics from Different Perspectives. To further analyze the relationship between the characteristics of the target trajectory from diverse perspectives, this manuscript utilizes both a straight line and curve trajectories as illustrative examples to examine their statistical characteristics of the observation characteristics from diverse perspectives. When the target linear trajectory is a concern, Figure 5 denotes the observation of the linear motion of the same target from two different perspectives. The statistical features of the characteristics of the target trajectory observation from two different perspectives are presented in Figure 6 suggesting that the statistical envelope of the trajectory characteristics between the two perspectives is very similar, while the center of gravity is different.

The trajectory of the curve motion is presented in Figure 7. Statistical analysis is conducted on the observation characteristics from different perspectives, and the characteristic of the statistical curve is presented in Figure 8 suggesting that the statistical characteristic of the curved trajectory from different perspectives have a certain similarity.

To sum up, although there exist differences in the perspective, a certain correlation exists in the performance of the feature sequence of the target trajectory. If there were more tracks along a certain direction, they would accumulate near the characteristic coding of that direction. Hence, it can be considered that the characteristic frequency curves of the target trajectory from different perspectives denote a certain linear translation.

3. Transfer Learning Based on the Trajectory of the HMM

For the recognition of the cross-view target trajectory, small sample data are employed in this research utilizing the transfer learning strategy to obtain the trajectory of the HMM model for the target view and to realize the low-cost modeling of the behavior model of the cross-view trajectory. The basic idea of the HMM model based on the transfer learning in this manuscript is presented as follows: firstly, according to the statistical translation of the trajectory characteristics in the different perspectives in Section 2.3, the linear regression model is constructed to attain the mapping between the source domain and the target domain concerning the HMM observation probability matrices of the characteristics, which leads to obtaining the observation probability. Secondly, the transfer probability matrix is randomly initialized, and the simulation data are generated. The objective optimization function is constructed according to the similarity between the simulation and the domain data of the target. Finally, the parameters of the transfer probability matrix for the trajectory HMM utilizing the target domain are obtained through the iterative optimization, and the transfer learning process of the cross-view HMM is realized.

Table 2: Direction freeman code of the target trajectory.

i	2	3	4	5	6	7	8	9	10
φ_i 0°	135°	135.6°	161.6°	90.0°	135.0°	135.0°	135.0°	141.3°	139.4°
Code	10	10	11	7	10	10	10	10	10

Table 3: The parameters of the HMM trajectory.

$\pi_{1:4,1}$	0.0000	0.0000	1.0000	0.0000		
A1, 1 : 4	0.0024	0.9976	0.0000	0.0000		
A2, 1 : 4	0.0000	0.0000	0.0536	0.9464		
A3, 1 : 4	1.0000	0.0000	0.0000	0.0000		
A4, 1 : 4	0.9044	0.0000	0.0956	0.0000		
B1, 7 : 12	0.0000	0.0000	0.0000	0.9261	0.0000	0.0739
B2, 7 : 12	0.0000	0.0000	0.0000	0.7153	0.2831	0.0000
B3, 7 : 12	0.0000	0.0000	0.0000	0.7169	0.2831	0.0000
B4, 7 : 12	0.0835	0.1670	0.4174	0.0015	0.3306	0.0000

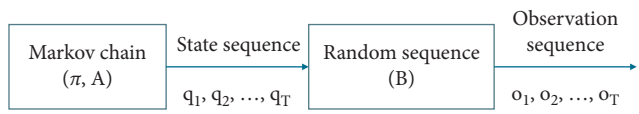
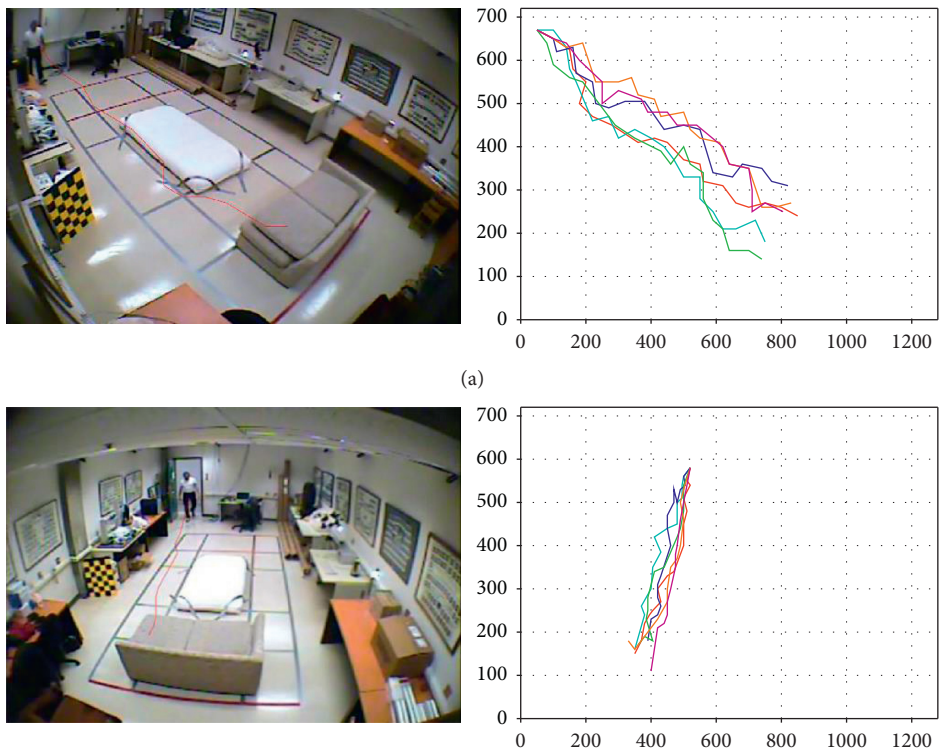


Figure 3: The simulation process of the HMM.

Table 4: Algorithm of the simulation of the HMM to generate data.

Algorithm steps of the simulation of the HMM to generate data.
(1) According to the initial state probability distribution $\pi = \pi_i$, select an initial state $Q_1 = i$
(2) For $t = 1$, 2, 3, K, and T
According to the output probability distribution b_{jk} of state i, output $Q_t = k$;
END For



(a)

(b)

Figure 4: Continued.

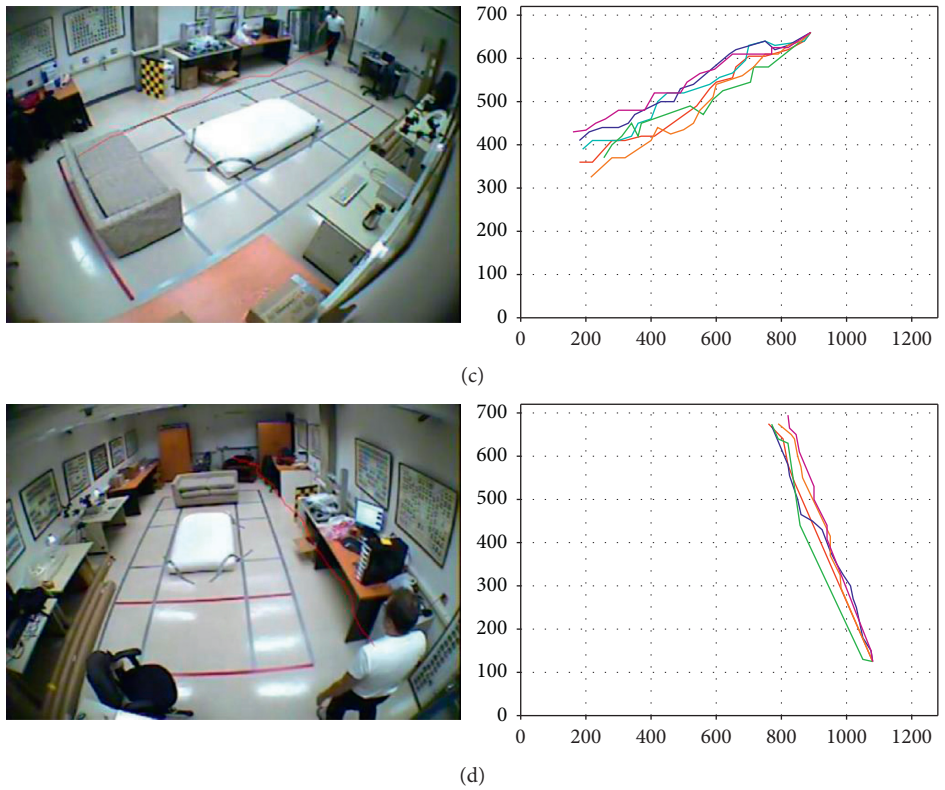


Figure 4: Simulation data generated by the training model of the target trajectory from different views (from top to bottom are different views, from left to right are the target trajectory samples, and the results of the simulation generated by the training model).

Figure 5: The position change of the linear trajectory of the same target in different perspectives (from top to bottom are the perspectives 1 and 2 and from left to right are the frames 22, 33, 44, 55, and 66).

3.1. Transfer of the Observation Probability Matrix B. The key to the transfer of the observation probability matrix is to determine the mapping relationship between the feature spaces of the target domain and the source domain. By utilizing this mapping relationship to transfer the HMM parameters of the source domain, a parameter model is formed that can reflect the target trajectory characteristics of the target domain.

In this manuscript, the least square method is adopted to construct the mapping model, and the trajectory characteristic of the mapping model under two different perspectives is assumed to be represented by

$$\overline{O_t} = wO_s + b, \tag{2}$$

where w and b are called the coefficients of a curve equation fitting by feature mapping, O_s is called a coding sample of the source domain, and $\overline{O_t}$ is called the coding data of the target domain after the mapping. The objective function is defined by

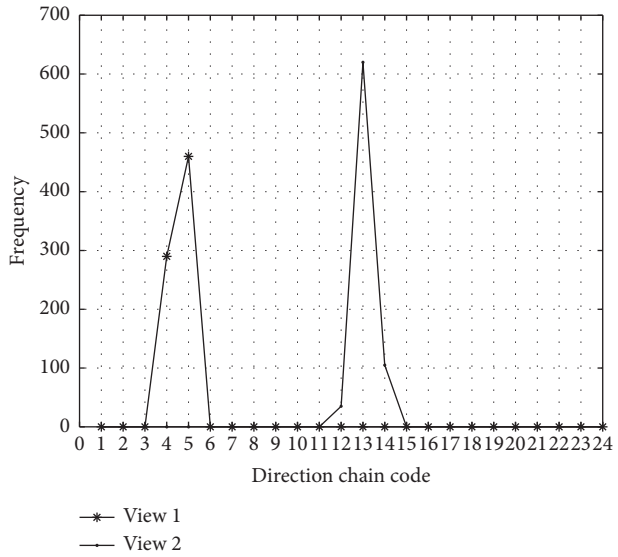


Figure 6: The statistical frequency curves of the characteristics of the linear trajectory from different perspectives.

(a) (b)

Figure 7: Position change of the curved trajectory of the same target under different perspectives (from top to bottom are called the perspectives 1 and 2, from left to right are called the positioning of the target under different perspectives, and the trajectory of the target under different perspectives).

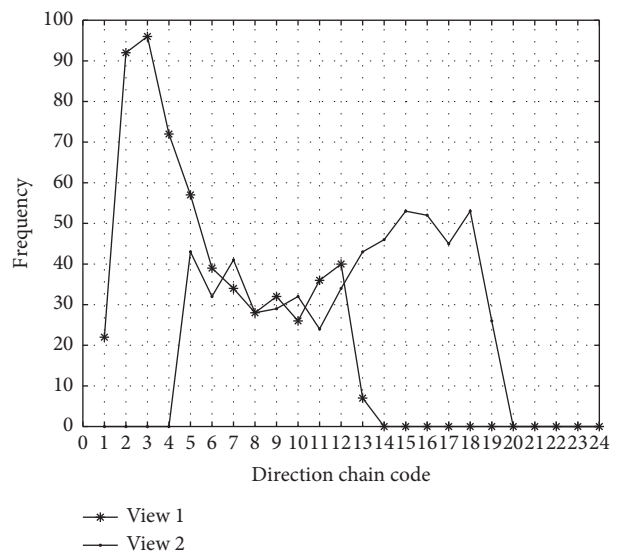


Figure 8: Statistical frequency curves of the characteristics of the curve trajectory from different perspectives.

$$\sum \left(O_t - \overline{O_t}\right)^2 = \sum \left(O_t - (wO_t + b)\right)^2, \quad (3)$$

where O_t is called the encoding data of the real target domain.

According to the statistical analysis of the trajectory characteristics in Section 2.3, the transfer of the parameter of the observation probability mainly includes three steps. First, both perspective trajectory samples of the source domain and target domain are collected. Each group of samples contains the source domain that is labeled as the samples of the same behavior and as a small number of target domains that are labeled as the samples. The source domain of the HMM trajectory model denoted by $\lambda_S = (A_S, B_S, \pi_S)$ is obtained according to the method of the training model in Section 2.1. Then, the linear regression mapping model about source domain and target domain features is constructed. Finally, the observation probability denoted by B_T of the target domain is attained utilizing the relation transfer matrix B_S of the linear regression.

3.2. Transfer Learning Algorithm of the Transfer Probability Matrix A. Ignoring the effect of the initial distribution, the initial distribution π_S of the source domain is directly migrated to the target domain to form π_T, and the transition probability of the source domain acts as the transition probability of the initial value $\widehat{A}_T$ of the target domain. Then, the transferred initial HMM is denoted by $\widehat{\lambda}_T = (\widehat{A}_T, B_T, \pi_T)$. In this study, the optimization algorithm is utilized to adaptively optimize the transfer of the model parameters to attain better performance of the target domain model. The parameters of the optimized target source model are more suitable for the target domain data, that is, the more similar the simulation data of the optimized model to the target domain data would be, the better it would converge. The objective function is defined by

$$\min_{\Delta A} \operatorname{sim}\left[\overline{O_T}, tO_T\right] = \operatorname{sim}\left[g\left(\widehat{A}_T + \Delta A, B_T, \pi_T\right), O_T\right], \quad (4)$$

where $g(\cdot)$ is the mean value of the data generated by the simulation $\lambda(\widehat{A}_T + \Delta A, B_T, \pi_T)$ and ΔA is called a variable in the optimization problem. The trajectory data is simulated according to the algorithm steps in Section 2.2. The inherent characteristics of the HMM require that the transition matrix is non-negative and the sum of row elements is equal to 1, so the constraint conditions of this optimization problem are that the elements of the transition probability matrix $\widehat{A}_T$ and $\widehat{A}_T + \Delta A$ are non-negative and the sum of row elements is equal to 1. The measurement of similarity is determined by the Euclidean distance, which is defined by

$$d\left(\overline{O_T}, tO_T\right) = \sqrt{\left(\sum \left(\overline{O_T} - tO_T\right)^2\right)}, \quad (5)$$

where $\overline{O}_T$ and $O_T$are the mean values of the feature sequence set of the simulated trajectory model represented by $\widehat{\lambda}_T = (\widehat{A}_T, \widehat{B}_T, \pi_T)$ and target domain, respectively. Hence, both of which belong to the same trajectory category. The similarity computation is defined by

$$\operatorname{sim}\left(\overline{O}_T, O_T\right) = \frac{1}{1 + d\left(\overline{O}_T, O_T\right)}. \quad (6)$$

Utilizing the constraints and objective function presented above, the interior point method is undoubtedly one of the most suitable methods to resolve the optimal value ΔA. Table 5 presents the steps of this solution procedure leading to the optimality of ΔA employing the interior point method. The similarity between the simulated data of the target domain model and O_T is calculated. If the similarity is greater than or equal to the similarity threshold denoted by δ, the obtained ΔA from the previous optimization step is reentered into the iteration of the interior point method as the initial value until the computed similarity is less than the threshold δ, and then, the HMM $\lambda_T = (\widehat{A}_T + \Delta A, B_T, \pi_T) = (A_T, B_T, \pi_T)$ of the target domain is attained.

3.3. HMM Transfer Learning and Recognition Framework. To resolve the problem of the recognition of the cross-view target trajectory, this manuscript proposes a cross-view HMM transfer method. The algorithm framework is presented in Figure 9. Utilizing the objective function analysis, the statistics of the characteristics of the target trajectory under the two perspectives suggest a linear relationship. Therefore, the mapping function between the characteristic spaces of the source domain and the target domain is first obtained by resolving an optimization problem based on the linear regression model. Utilizing the mapping function, the parameters of the observation probability of the source domain are transferred. Then, the similarity between the generated data of the simulated HMM and the labeled data obtained from the target domain perspective is computed to further optimize the transition probability of the source domain. Hence, the target domain HMM is obtained. Finally, the forward algorithm is utilized to assess the match, classification, and identification of the cross-view target trajectory.

4. Experimental Results and Analysis

4.1. Experimental Evaluation. The data sets employed in this experiment are all self-collected data, which come from the simultaneous shooting of two cameras installed in the $10 \times 12\,\mathrm{m}^2$ room. The 100 samples of three kinds of trajectories (300 in total) from view A of the first camera can better reflect the real movement behavior of the moving target. The 300 samples are taken as the source domain data set, while the corresponding 300 samples of the view B of the second camera are taken as the target domain sample. All collected samples are manually marked. In this manuscript, the performance of the model is assessed by the accuracy measurement defined by

$$\mathrm{Accu} = \frac{n_{\mathrm{accu}}}{\mathrm{Num}}, \quad (7)$$

where n_{accu} is called the number of the correct trajectory classification and Num is called the total number of the test data.

Table 5: Steps of the algorithm of the transition probability optimization.

Steps of the algorithm of the transition probability optimization.
(1) Simulate the trajectory samples by model $\lambda(\widehat{A}_T, B_T, \text{and } \pi_T)$
(2) Calculate the similarity between simulation set and target domain label sample set: sim
(3) While sim>δ
Solve ΔA according to the interior point method, update $\widehat{A}_T$to $\widehat{A}_T + \Delta A$
Simulate trajectory samples via the model $\lambda(\widehat{A}_T, B_T, \text{and } \pi_T)$
Calculate the similarity between simulation set and target domain label sample set: sim
END While

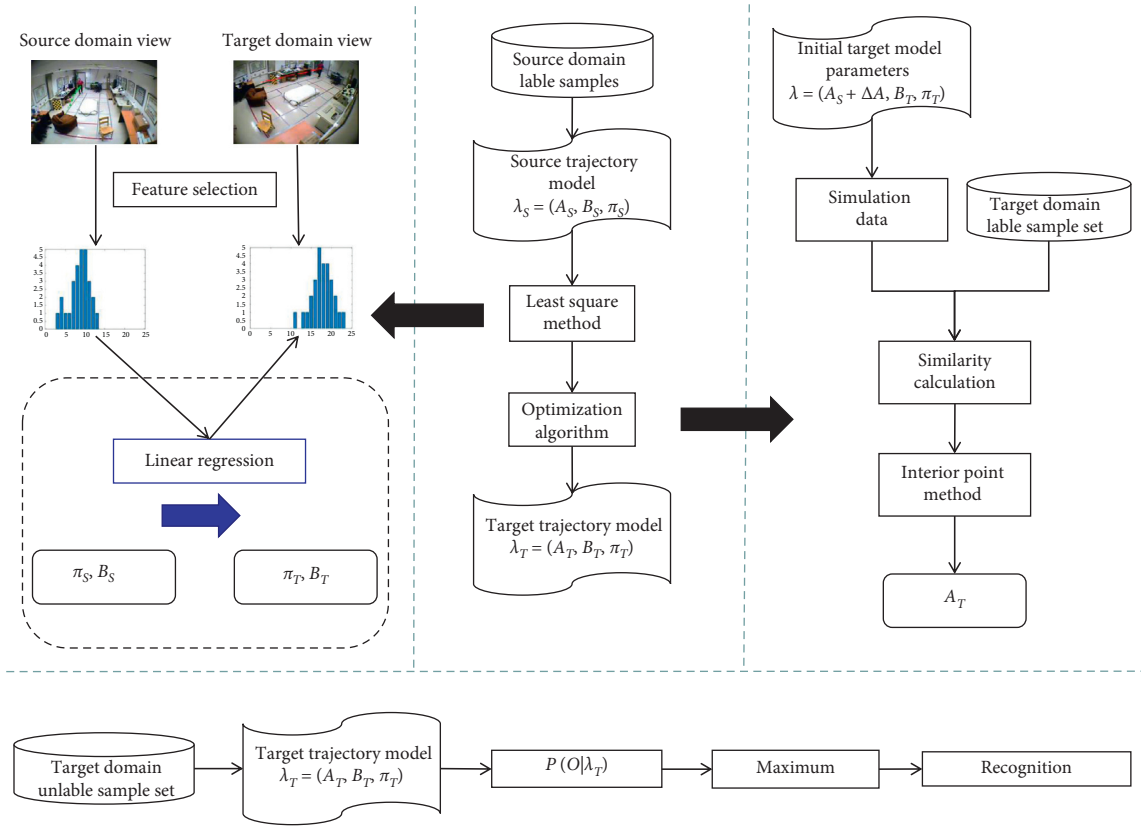


Figure 9: The HMM transfer method framework. $\lambda_S = (A_S, B_S, \text{and } \pi_S)$ represents the source domain of the HMM; $\lambda_T = (A_T, B_T, \pi_T)$ represents the target domain of the HMM; $P(O|\lambda_T)$ is called the probability of the occurrence of the observation value sequence O in the model λ_T.

4.2. Analysis of Experimental Results. The first experiment aims to qualitatively compare the performance of the HMM transfer method described in this manuscript with the one that existed in the literature [9]. In this experiment, k is set to some sample values such as ($k = 10$, 15, 20, 25, 30, 40, and 50) representing each type of the trajectory in the sample set of the target domain that is randomly selected and combined with the sample set of the source domain as the training data. In addition to k number of the selected samples, all other data in the sample set of the target domain are employed as the test data. Figures 10–12 are the comparisons of the accuracies of the test data between the source and the target domains when the number k of each type of trajectory sample changes with the angle of the view $\theta = 30^\circ$, $\theta = 45^\circ$, and $\theta = 60^\circ$, respectively. The experimental results suggest that the accuracy of the method in the experiment is improved with the increment of the number of the target domain samples, k. As the number of the target domain samples increases, the learning algorithm could extract more knowledge from the target domain. When only $\theta = 60^\circ$, the accuracy of trajectory 2 does not improve, or even decreases, indicating that overfitting occurs in this case. It is worth noting that the accuracy of this algorithm increases faster than other algorithms with the increment of samples. When the included angle varies from 30° to 45° and 60°, and k is kept constant, the corresponding accuracy of most trajectories decreases, indicating that the increment of the angle of the view leads to more obvious deformation of the trajectory at different included angles (Figure 12). The experimental results also suggest that when compared with the method researched in [9], this method could achieve a higher recognition rate of the behavior trajectory of the cross-view

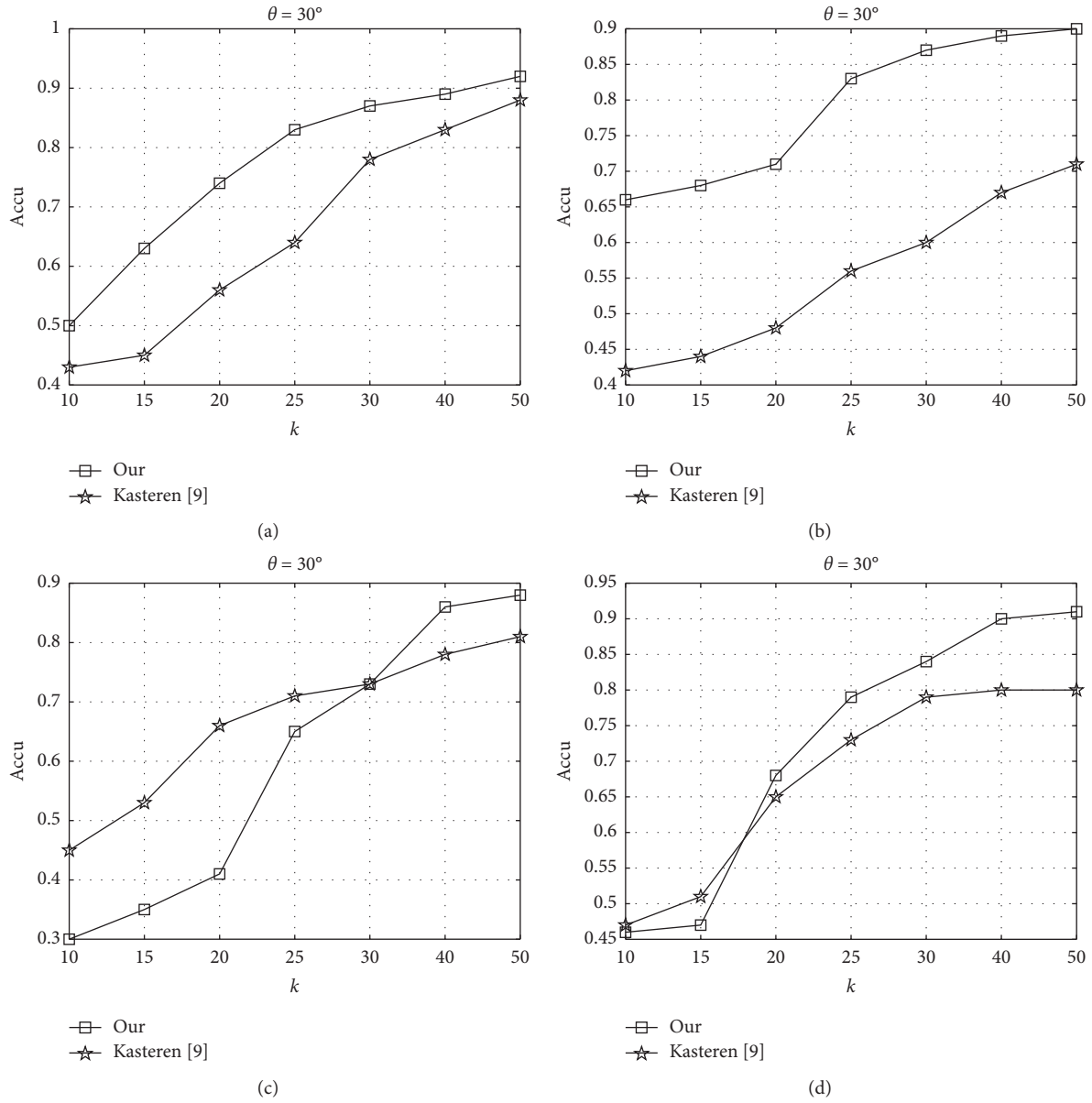


Figure 10: The comparison between the proposed method and the method in [9] on the change of the accuracy in testing data utilizing the number of the target domain markers ($\theta = 30^{\circ}$). (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4.

target utilizing only a small amount of the marked data from the target domain perspective, which would greatly reduce the annotation work of the target domain perspective.

In the second experiment, according to the results of the previous experiment, the number of training data of the target domain perspective is fixed to 50. To verify the effectiveness of this method, the proposed method is compared with both Method 1 and Method 2, and the results are recorded. In Method 1, the marked training data from view B are utilized to find out the parameters of the model. The marked data in the number of view B are small, so the leave-one-out cross-validation technique is implemented. In this technique, a group of 20 samples is taken from the target domain data set as the test set, the rest are left as the training data until all samples have been tested. A total of M times (the size of the data set) are conducted and calculated, and finally, the accuracy is averaged. In Method 2, only the marked training data in view A are utilized to conduct the Baum–Welch algorithm to attain the HMM, and the determined parameters of the model are utilized to attain the accuracy of the test data set of view B.

Tables 6–8 present the accuracy rates of the different methods at the different θ values. The purpose of the comparison with Method 1 is to decide whether a trained model of the source domain could be utilized directly across visual angles. The results imply that the effect is poor, especially from different views since the knowledge extracted from the source domain cannot be directly applied to the target domain, especially when the two

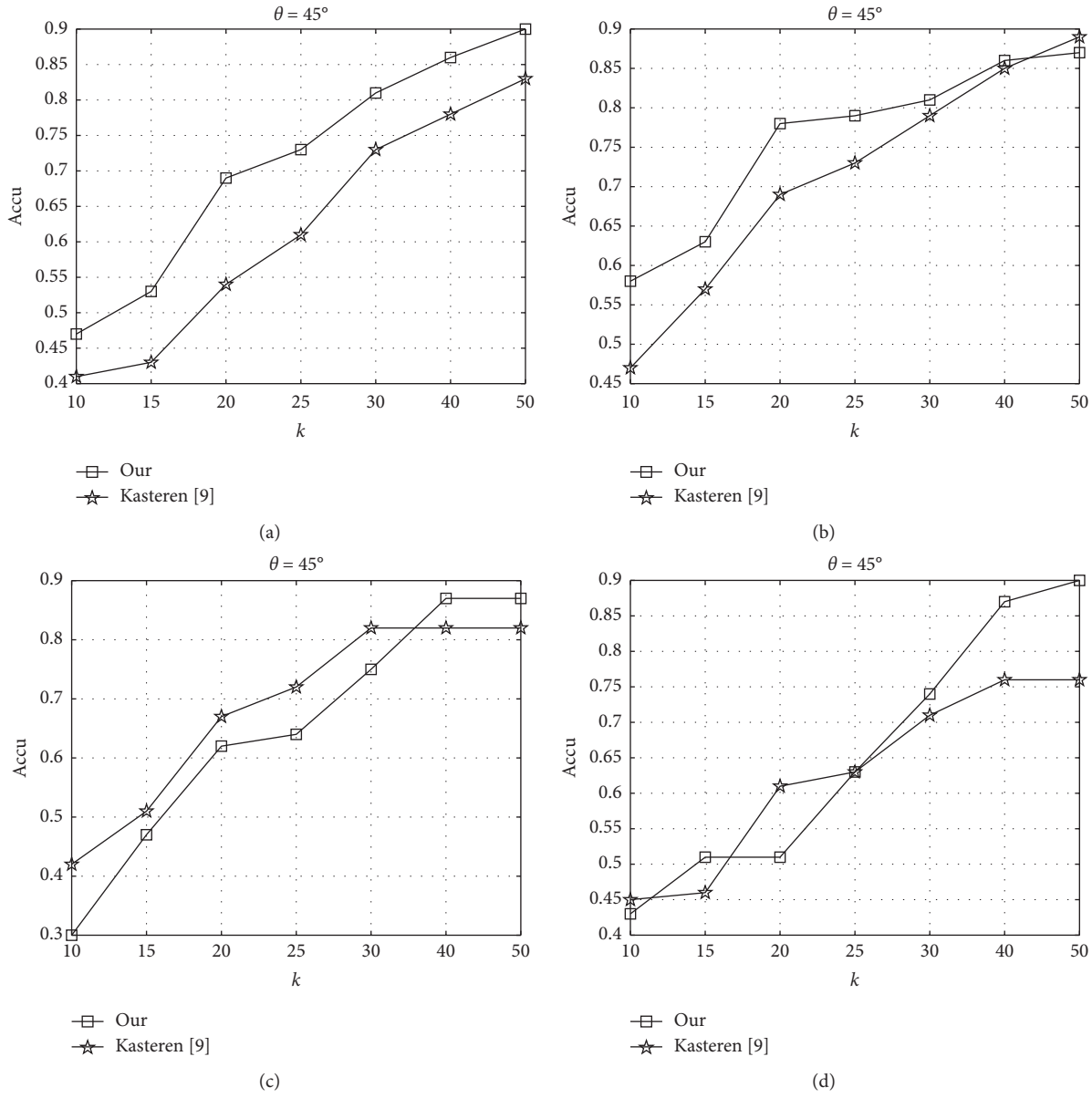


Figure 11: The comparison between the proposed method and the method in [9] on the change of the accuracy in testing data utilizing the number of the target domain markers ($\theta = 45^{\circ}$). (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4.

domains differ much. Compared with Method 1, the method in this manuscript greatly improves the outcomes of the recognition. This indicates that although the HMM trajectory modeling is very successful when implemented from a single perspective, its performance would significantly decrease when both cross views and crossing view have a great impact on trajectory recognition. Compared with Method 2, the recognition effect is similar, indicating that when a small data set of the target domain are provided, the performance of the HMM transfer model is almost the same as that of the training model with a large data set of the target domain. In conclusion, the transfer learning method in this manuscript makes use of the prior knowledge that has been already extracted in the source domain model and hence, could effectively identify the cross-view target trajectory. Besides, it does not require a high number of target domain samples. Since there are enough samples in the source domain, which contains a large amount of characteristic information, the transfer learning could utilize only a small target data set to combine with these prior pieces of knowledge to attain better outcomes.

In the third experiment, according to the outcomes of the first experiment, the training data of the target domain perspective are fixed to 50, and the performance of the proposed method is compared with the HMM transfer method described in the literature [10, 11].

Seen from Tables 9–11 that the HMM transfer method in this manuscript is more effective than other HMM transfer methods under the same experimental settings.

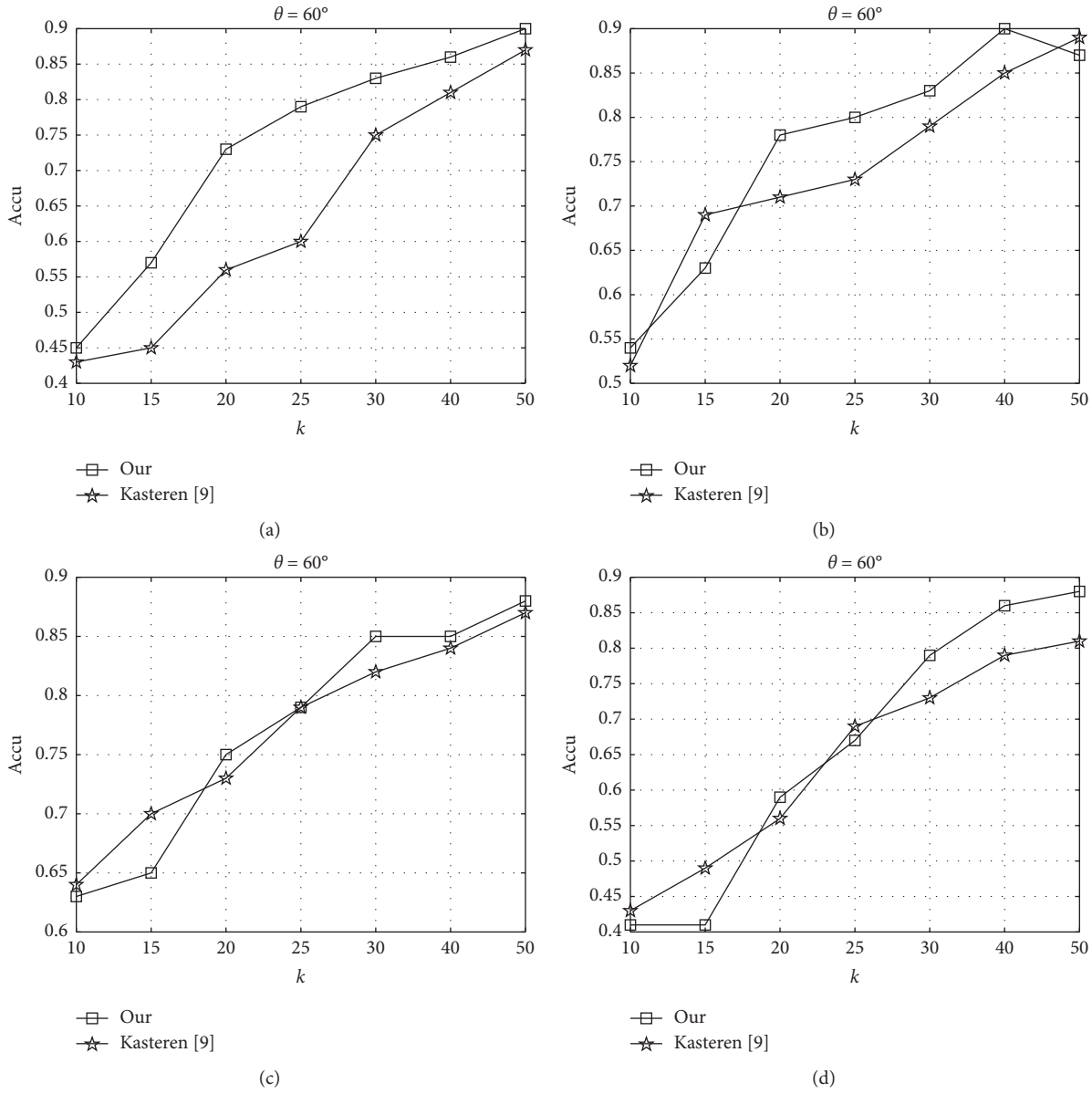


Figure 12: The comparison between the proposed method and the method in [9] on the change of the accuracy in testing data utilizing the number of the target domain markers ($\theta = 60^{\circ}$). (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4.

Table 6: The accuracies of Method 1, Method 2, and the proposed method ($\theta = 30^{\circ}$).

Methods	Method 1	Method 2	The proposed method
Trajectory 1	0.95	0.63	0.92
Trajectory 2	0.85	0.61	0.90
Trajectory 3	0.90	0.37	0.88
Trajectory 4	1.00	0.41	0.91
Average	0.93	0.48	0.90

Table 7: The accuracies of Method 1, Method 2, and the proposed method ($\theta = 45^{\circ}$).

Methods	Method 1	Method 2	The proposed method
Trajectory 1	0.95	0.59	0.90
Trajectory 2	0.90	0.54	0.87
Trajectory 3	0.95	0.39	0.87
Trajectory 4	0.85	0.34	0.90
Average	0.91	0.47	0.89

Table 8: The accuracies of Method 1, Method 2, and the proposed method ($\theta = 60^{\circ}$).

Methods	Method 1	Method 2	The proposed method
Trajectory 1	1.00	0.39	0.90
Trajectory 2	0.90	0.50	0.87
Trajectory 3	0.85	0.59	0.87
Trajectory 4	0.90	0.37	0.88
Average	0.91	0.46	0.88

Table 9: The accuracies of the methods in the [10, 11] and the proposed method ($\theta = 30^{\circ}$).

Methods	The method in [10]	The method in [11]	The proposed method
Trajectory 1	0.81	0.75	0.92
Trajectory 2	0.79	0.86	0.90
Trajectory 3	0.84	0.88	0.88
Trajectory 4	0.91	0.78	0.91
Average	0.84	0.82	0.90

Table 10: The accuracies of the methods in [10, 11] and the proposed method ($\theta = 45^{\circ}$).

Method s	The method in [10]	The method in [11]	The proposed method
Trajectory 1	0.90	0.68	0.90
Trajectory 2	0.74	0.80	0.87
Trajectory 3	0.82	0.88	0.87
Trajectory 4	0.87	0.79	0.90
Average	0.83	0.79	0.89

Table 11: The accuracies of the methods in [10, 11] and the proposed method ($\theta = 60^{\circ}$).

Methods	The method in [10]	The method in [11]	The proposed method
Trajectory 1	0.93	0.77	0.90
Trajectory 2	0.81	0.85	0.87
Trajectory 3	0.71	0.84	0.87
Trajectory 4	0.92	0.81	0.88
Average	0.84	0.82	0.88

In the transfer recognition of the cross-view HMM trajectory model, the accuracy rate of the proposed method reaches over 85% in most cases, and the half could be over 90%.

The experiments present that the linear regression model based on the least square method can be successfully applied to the transfer learning method of target behavior recognition. The characteristic representation of both the source domain and target domain differs due to the variation of a visual point. To overcome this issue, we need to find statistical patterns between the characteristics of the source domain and the target domain. Then, the curve needs to be fitted. By doing so, transferring the probability distribution of the emission of the source domain adaptively is transferred to the target domain. Hence, it leads to transferring the correlation knowledge between different perspectives and behaviors from one domain to another. The existence of a small number of the labeled samples in the target domain provides a reference for the optimization of the state transition probability of the source domain. The improved and optimized HMM performs well on the data set with higher accuracy and better robustness.

5. Conclusion

This manuscript mainly deals with identifying the behavior categories of the trajectory data of the target behavior when the cross perspective is a concern. In the case of having insufficient labeling data of the target domain, having difficulty in labeling data and dealing with a higher cost, the problem is transformed into the identification problem of the feature sequence of the trajectory utilizing the HMM transfer learning. Through knowledge transfer of relevant data from different perspectives, the classification model is constructed, and then, the behavior trajectory category of the data to be classified in the target domain is identified. In this manuscript, first, according to the data that is labeled from the perspective of the source domain, an HMM is constructed to train the parameters of the triples; second, the frequency of the coding sequence features utilizing the same type of data that is labeled from the perspectives of the

source domain and target domain is computed. Then, the mapping curves of the two perspectives are fitted by the least square. Hence, the observed probability matrix is adaptively transferred by mapping relationships. Afterward, the transition probability based on a small labeled data of the target domain is transferred to the target domain through an optimization algorithm. Therefore, the target domain classification model is constructed. The experimental results on the data set denote that the transfer learning method based on the HMM could construct the classification model, and it has a better performance in the recognition of the cross-view target trajectory when there exist a small number of labeled data in the target domain perspective.

Data Availability

All the data are included in the manuscript, and further data can be requested from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (no. 61673318, target trace analysis and abnormal behavior warning in structural space), the Key Industrial Chain Project of Shaanxi Province (no. 2020ZDLGY04-04), and the Science and Technology Planning Project of Xi'an City (no. GXYD14.1, intelligent identification system of residue state for Subway construction).

References

[1] P. Qiming and Y. Cheng, "Trajectory recognition of moving objects based on the hidden markov model," *Computer Application Research*, vol. 25, no. 7, pp. 1988–1991, 2008.

[2] B. Dapeng, Y. Shanshan, Z. Shi, Y. Minghui, and W. Yun, "The method of fleet call and search based on hidden markov model," 2020.

[3] A. Hervieu, P. Bouthemy, and J. P. L. Cadre, "An HMM-based method for recognizing dynamic video contents from trajectories," in *Proceedings of the IEEE International Conference on IEEE ICIP 2007*, San Antonio, TX, USA, 2007.

[4] L. Qian and H. C. Lau, "A layered hidden markov model for predicting human trajectories in a multi-floor building," in *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT) ACM*, Singapore, 2015.

[5] G. Lin, B. Yun, and Z. Weigong, "Recognition of abnormal interactions based on coupled hidden Markov models," *Journal of Southeast University*, vol. 43, pp. 1217–1221, 2003.

[6] R. Vasanthan, G. Ver Steeg, A. Galstyan, and A. Tartakovsky, "Coupled hidden Markov models for user activity in social networks," in *Proceedings of the IEEE International Conference on Multimedia & Expo Workshops IEEE*, San Jose, CA, USA, 2013.

[7] P. Kumar, G. Himaanshu, P. Roy, and D. P. Dogra, "Coupled HMM-based multi-sensor data fusion for sign language recognition," *Pattern Recognition Letters*, vol. 86, 2017.

[8] L. Qian and H. C. Lau, *A Layered Hidden Markov Model for Predicting Human Trajectories in a Multi-Floor Building*, IEEE Computer Society, Washington, DC, USA, 2015.

[9] T. L. M. Van Kasteren, G. Englebienne, and B. J. A. Kröse, "Recognizing activities in multiple contexts using transfer learning," in *Proceedings of the AAAI Fall Symposium: AI in Eldercare: New Solutions to Old Problems*, Arlington, VI, USA, 2008.

[10] V. W. Zheng, Q. Yang, W. Xiang, and D. Shen, "Transferring localization models over time," in *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*, pp. 1421–1426, Chicago, IL, USA, July 2008.

[11] G. Bingtao, Z. Yang, and L. Bin, "BioTrHMM: biomedical named-entity recognition algorithm based on transfer learning," *Computer Application Research*, vol. 36, no. 1, pp. 51–54, 2019.

[12] D. K. Kim and N. S. Kim, "Maximum a posteriori adaptation of HMM parameters based on speaker space projection," *Speech Communication*, vol. 42, no. 1, pp. 59–73, 2004.

[13] N. S. Kim, J. S. Sung, and D. H. Hong, "Factored MLLR adaptation," *IEEE Signal Processing Letters*, vol. 18, no. 2, pp. 99–102, 2011.

[14] O. Siohan, C. Chesta, and C.-H. Chin-Hui Lee, "Joint maximum a posteriori adaptation of transformation and HMM parameters," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 4, pp. 417–428, 2001.

[15] K. Ait-Mohand, T. Paquet, and N. Ragot, "Combining structure and parameter adaptation of HMMs for printed text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 9, pp. 1716–1732, 2014.

[16] X. Fei, Y. Xie, S. Tang, and J. Hu, "Identifying click-requests for the network-side through traffic behavior," *Journal of Network and Computer Applications*, vol. 173, Article ID 102872, 2021.

[17] F. Liu, D. Janssens, J. Cui, G. Wts, and M. Cools, "Characterizing activity sequences using profile hidden markov model," *Expert Systems with Applications*, vol. 42, no. 13, pp. 5705–5722, 2015.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6659118, 6 pages
https://doi.org/10.1155/2020/6659118

Research Article

Finite-Element Analysis on Compressive Performance of a Novel Glue-Laminated Cornstalk Scrimber

Wei Tian, Yongmei Qian, Zunpeng Liu, and Yiming Wang

Jilin Jianzhu University, Changchun 130118, China

Correspondence should be addressed to Yongmei Qian; qianyongmei@jlju.edu.cn

Received 26 October 2020; Revised 23 November 2020; Accepted 4 December 2020; Published 24 December 2020

Academic Editor: Gengxin Sun

Novel glue-laminated cornstalk scrimber is a new timber substitute produced by special techniques, without damaging the original fibers in cornstalks. This novel material outperforms ordinary timber in the resistance to water, damping, insect, and fire and provides a desirable green building material. However, glue-laminated cornstalk scrimber has not been widely implemented in the building industry, because the application of cornstalk products is limited to decoration panels. With the aid of the finite-element software Abaqus, this paper simulates the glue-laminated cornstalk scrimber specimens with different slenderness ratios under axial compression and analyzes the compressive performance of such specimens. The results show that the height of glue-laminated cornstalk scrimber is negatively correlated with the buckling load and nonlinear load under axial compression and positively correlated with the transverse displacement and axial displacements induced by axial compression. The research results provide a good reference for improving the design and application of glue-laminated cornstalk scrimber.

1. Introduction

Konjac is a perennial herbaceous plant in the genus of *Amorphophallus*, Araceae family. The main active component is konjac glucomannan (KGM), a water-soluble natural polymer polysaccharide [1]. The KGM is generally considered as a polysaccharide formed by linking glucose and mannose at the ratio of 1 : 1.69 or 1.4 : 1 through β-1, 4 glycosidic bonds. With a molecular weight of 200–20,000 kDa, the KGM has an acetyl group at the C6 position for every 19 sugar units on the main chain [2, 3].

In recent years, China has highlighted the importance of ecological civilization and further restricted timber consumption to protect forest resources. As a result, a diverse array of timber substitutes has emerged, such as scrimber and glulam [1, 2].

Much research has been done on timber substitutes around the world. For instance, Uzel et al. [3] compared the flexural behavior of glulam beams with that of large-volume timber beams. The University of Catania [4] described the structural behaviors of glued laminated (GL) timber and laminated veneer lumber (LVL) and introduced a new technique called post-tensioning of timber beams. Through tests and evaluation, Graz University of Technology [5] proposed a generic design approach generally applicable for all linear and planar, unidirectional, and orthogonal laminated structural timber products. The Institute for Mechanics of Materials and Structures [6] developed a stochastic engineering framework capable of accounting for how the random fluctuations of the stiffness of wooden boards work on the performance of glued laminated timber (GLT). Anshari and Guan [7] established a finite-element model to simulate the prestress performance of compressed wood (CW) glulam beams and further examined the structural properties of prestressed beams under destructive bending.

Cornstalks, the most common type of crop stalks, can be reconstituted and recombined into a novel timber substitute. This helps to protect the air environment and alleviate the shortage of forest resources. Therefore, the application of glue-laminated cornstalk scrimber not only boasts profound theoretical value but also has great economic and social significance [8–10]. Despite its excellent features (e.g., high density, high strength, fire resistance, corrosion resistance,

moisture resistance, and insect resistance), glue-laminated cornstalk scrimber has only been applied as decorative panels. Few studies have attempted to implement the novel material in the building industry [11–14].

In the light of the above, this paper carries out a finite-element simulation on the compressive performance of glue-laminated cornstalk scrimber, shedding theoretical new lights on the mechanical performance of this novel material. The research results could promote the design and application of glue-laminated cornstalk scrimber.

2. Finite-Element Analysis on Glue-Laminated Cornstalk Scrimber under Axial Compression

2.1. Constitutive Relationship. Glue-laminated cornstalk scrimber is a nonsymmetrical anisotropic material. To measure its compressive performance, the elastic moduli, Poisson's ratios, and shear moduli of three surfaces were recorded in the elastic phase: the fiber direction 1, the scrimber direction 2, and the bonding direction 3. Referring to the definitions of timber yield phase, the generalized Hill yield criterion [15–19] was introduced for the yield phase. Table 1 lists the performance parameters of glue-laminated cornstalk scrimber.

2.2. Finite-Element Model. The axial compression specimens of glue-laminated cornstalk scrimber were meshed into three-dimensional (3D) eight-node hexagonal elements C3D8R (Figure 1). The load was applied concentratedly on the top of the specimen. The load was fixed for the elastic model and gradually stepped up for the nonlinear model. At the bottom of the specimen, the support was allowed to rotate in one direction only; at the top, the support could rotate in one direction and displace in the axial direction. Through structural optimization, the finite-element model of glue-laminated cornstalk scrimber mainly consists of hexagonal elements [20, 21].

2.3. Model Size. This research only considers the overall stable bearing capacity, without tackling local stability. Therefore, the slenderness ratio of the finite-element model was changed to study its effect on the ultimate bearing capacity.

Drawing on Technical Code of Glued Laminated Timber Structures (GBT50708-2012) and Code for Design of Timber Structures (GB50005-2017), the model size was designed as follows: the sectional area was fixed at 150 mm × 150 mm, while the length was set to 1,200 mm, 1,600 mm, 2,000 mm, 2,400 mm, and 2,800 mm, respectively. The resulting specimens were coded Z-1 to Z-5 in turn (Table 2).

2.4. Results Analysis

2.4.1. Eigenvalue Buckling Analysis. The eigenvalue buckling analysis aims to calculate the ultimate load that drives an elastic member from a stable state to the critical state of instability. The analysis results are consistent with the classical Euler solution. Through Abaqus calculation, the cloud atlas of buckling equivalent stresses and buckling equivalent displacements of Z-1 to Z-5 were obtained (Figures 2 and 3).

Table 1: The performance parameters of glue-laminated cornstalk scrimber

Elastic modulus (MPa)			Shear modulus (MPa)			Poisson's ratio			Density (kg/m^3)
E_1	E_2	E_3	G_{12}	G_{13}	G_{23}	μ_{12}	μ_{13}	μ_{23}	ρ
8.817	925	548	805	584	183	0.21	0.45	0.43	1220

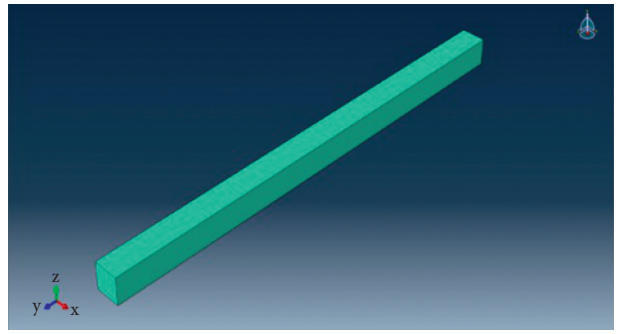


Figure 1: The finite-element model of glue-laminated cornstalk scrimber under axial compression.

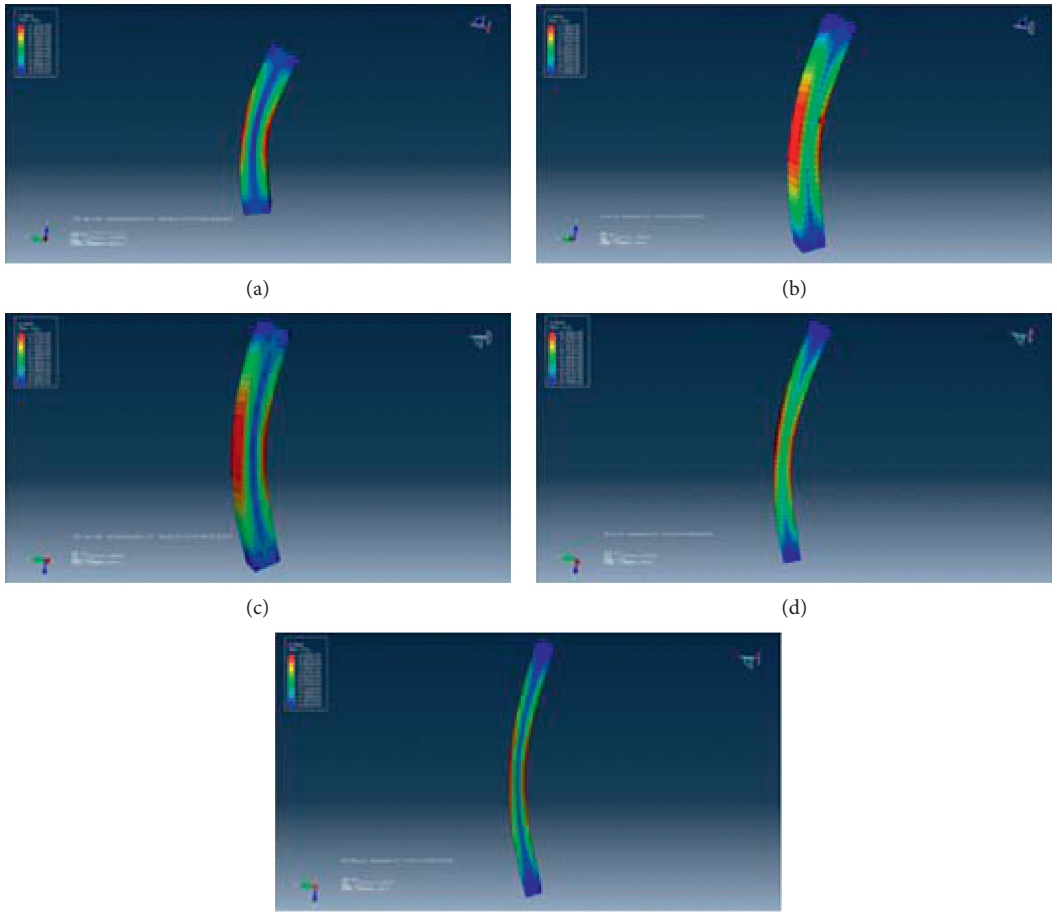
As shown in Figures 2 and 3, the buckling equivalent stress mainly concentrated in the middle of the specimens under axial compression. When the sectional area was fixed, Z-5 had the greatest deflection under ultimate load, while Z-1 had the smallest deflection. The deflection of the axial compression specimen gradually increased with the growth in specimen height. From Z-1 to Z-5, the eigenvalue buckling load was 605 kN, 545 kN, 526 kN, 502 kN, and 456 kN, respectively.

2.4.2. Nonlinear Buckling Analysis. The five axial compression specimens were subject to nonlinear analysis. In this way, the nonlinear buckling load of Z-1 to Z-5 was obtained as 564 kN, 523 kN, 504 KN, 465 kN, and 425 kN, respectively. As shown in Figure 4, the nonlinear buckling load was slightly smaller than the eigenvalue buckling load; the two loads exhibited roughly the same trend: the greater the slenderness ratio, the smaller the load and the weaker the resistance to lateral displacement.

As shown in Figures 5 and 6, for Z-1 to Z-5, the transverse displacement and axial displacements changed similarly under compression load: both of them gradually increased with the load and with the slenderness ratio. In other words, the bearing capacity decreased with the growing load and slenderness ratio. In the early phase of loading, transverse displacement and axial displacements increased almost linearly; the transverse displacement had the greater increasing rate. Later, as the load continued to grow, the traverse displacement tended to be stable, while the axial displacement still exhibited a linear growth.

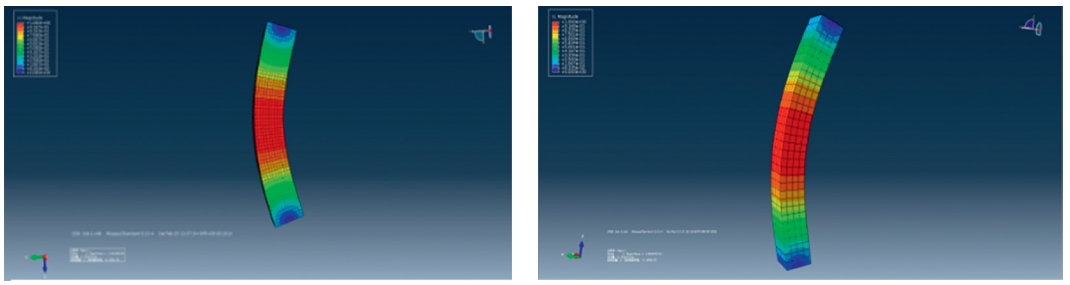
Table 2: The model size.

Specimen number	Sectional area (mm^2)	Length (mm)	Slenderness ratio
Z-1		1200	27.7
Z-2		1600	37.0
Z-3	150×150	2000	46.2
Z-4		2400	54.2
Z-5		2800	63.2



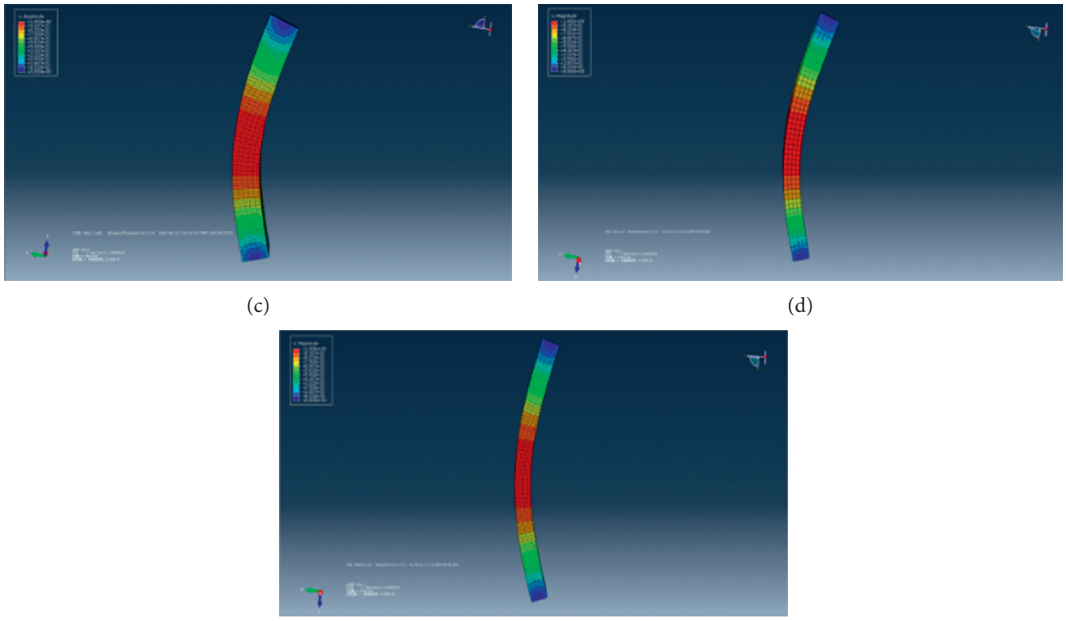
(a) (b) (c) (d) (e)

Figure 2: The cloud atlas of buckling equivalent stresses.



(a) (b)

Figure 3: Continued.



(c) (d)

(e)

Figure 3: The cloud atlas of buckling equivalent displacements.

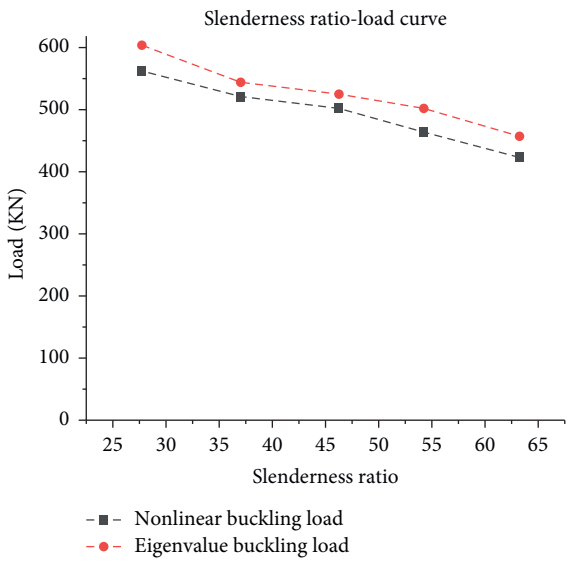


Figure 4: The slenderness ratio-load curves.

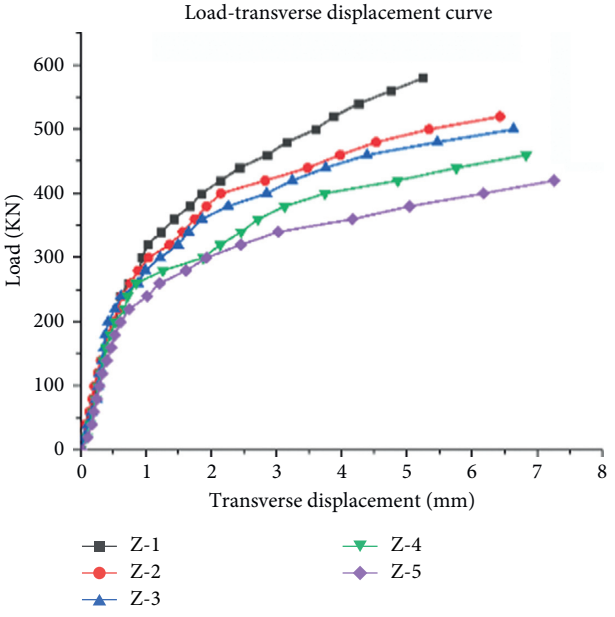


Figure 5: The load-transverse displacement curves.

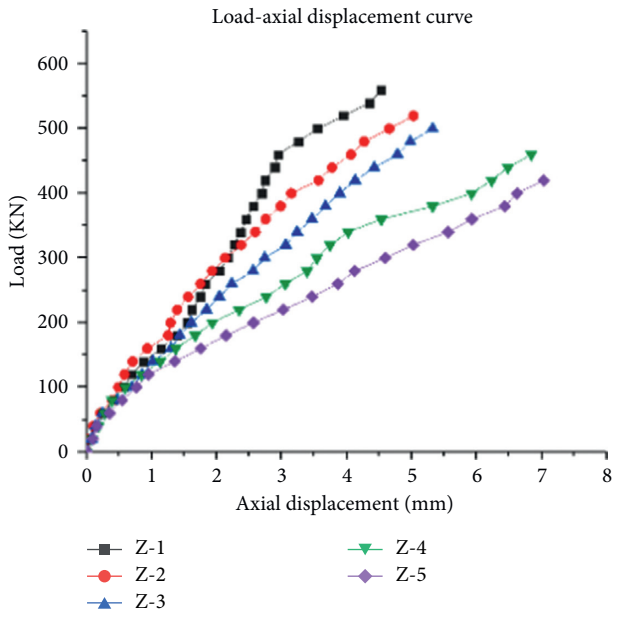


Figure 6: The load-axial displacement curves.

3. Conclusions

This paper carries out Abaqus finite-element simulation of glue-laminated cornstalk scrimber specimens with different slenderness ratios under axial compression. The generalized Hill yield criterion was introduced for elastoplastic analysis, in view of the nonlinear property of the material; through eigenvalue and nonlinear buckling analyzes, the cloud atlas of buckling equivalent stress and displacement and nonlinear load-displacement curves were obtained. The main conclusions are as follows:

(1) With the growing slenderness ratio, both eigenvalue buckling load and nonlinear buckling load gradually decreased, and the former was slightly greater than the latter. The eigenvalue buckling analysis assumes that the specimens are ideal elastic members, while the nonlinear buckling analysis considers the initial defects of the members. Therefore, the nonlinear buckling analysis better conforms to the actual situation and outputs more reliable results.

(2) According to the nonlinear buckling load-displacement curves, both transverse displacement and longitudinal displacements gradually increased with slenderness ratio. In the early phase of loading, transverse displacement and axial displacements increased almost linearly; the transverse displacement had the greater increasing rate. With further growth of the load, the traverse displacement tended to be stable, while the axial displacement continued to increase linearly at the original rate.

(3) The buckling load and nonlinear load of the axial compression specimens gradually decreased with the growth in specimen height. Moreover, the transverse displacement and axial displacements changed similarly under axial compression: both of them gradually increased with the height of specimen.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the Major Science and Technology Project Foundation of Jilin Province in China (20180201066SF).

References

[1] G. B. Hu, Z. T. Zhang, and X. B. Xiao, "China import wood situation and developing trend," *Forest Industry*, vol. 42, no. 1, pp. 5–9, 2015.

[2] X. Sun, M. He, and Z. Li, "Novel engineered wood and bamboo composites for structural applications: state-of-art of manufacturing technology and mechanical performance evaluation," *Construction and Building Materials*, vol. 249, Article ID 118751, 2020.

[3] M. Uzel, A. Togay, Ö. Anil, and C. Sogutlu, "Experimental investigation of flexural behavior of glulam beams reinforced with different bonding surface materials," *Construction and Building Materials*, vol. 158, pp. 149–163, 2018.

[4] A. D'Aveni and G. D'Agata, "Post-tensioned timber structures: new perspectives," *Construction and Building Materials*, vol. 153, pp. 216–224, 2017.

[5] R. Brandner, "Cross laminated timber (CLT) in compression perpendicular to plane: testing, properties, design and recommendations for harmonizing design provisions for structural timber products," *Engineering Structures*, vol. 171, pp. 944–960, 2018.

[6] G. Kandler, M. Lukacevic, C. Zechmeister, S. Wolff, and J. Füssl, "Stochastic engineering framework for timber structural elements and its application to glued laminated timber beams," *Construction and Building Materials*, vol. 190, pp. 573–592, 2018.

[7] B. Anshari and Z. Guan, "FE modelling of optimization on strengthening glulam timber beams by using compressed wood blocks," *Procedia Engineering*, vol. 171, pp. 857–864, 2017.

[8] D. S. Huang, A. P. Zhou, and Y. L. Bian, "Experimental and analytical study on the nonlinear bending of parallel strand bamboo beams," *Construction and Building Materials*, vol. 44, pp. 585–592, 2013.

[9] A. K. Ray, S. Mondal, S. K. Das, and P. Ramachandrarao, "Bamboo-A functionally graded composite-correlation between microstructure and mechanical strength," *Journal of Materials Science*, vol. 40, no. 19, pp. 5249–5253, 2005.

[10] M. H. Ge, *Research on Current Situation and Comprehensive Utilization of Straw Resources in Jilin Province*, Master thesis of Northeast Normal University, 2011.

[11] H. Liu, L. M. Zhang, and N. Q. Wang, "Research on current situation and development of straw panel," *China New Technologies and New Products*, vol. 2014, no. 5, p. 131, 2014.

[12] G. H. Chen, "Study on surface properties and gluing recombination technology of crop straw," Master thesis, of Central South University of Forestry and Technology, Changshu, China, 2006.
[13] C. Martins, A. M. P. G. Dias, and H. Cruz, "Bonding performance of Portuguese Maritime pine glued laminated timber," *Construction and Building Materials*, vol. 223, pp. 520–529, 2019.
[14] X. Li, Z. Cai, and J. Z. Winandy, "Effect of oxalic acid and steam pretreatment on the primary properties of UF-bonded rice straw particleboards," *Industrial Crops and Products*, vol. 33, no. 3, p. 665~669, 2012.
[15] J. A. Newlin and J. M. Gahagan, "Tests of large timber columns and presentation of the forest products laboratory column formula," United States Department of Agriculture, Washington. D.C, USA, Technical Bulletin No.167, 1930.
[16] B. Sharma and K. A. Harries, "Effect of fiber gradation on the edge bearing strength of bamboo culms," *Key Engineering Materials*, vol. 517, pp. 63–70, 2012.
[17] T. Reynolds, B. Sharma, K. Harries, and M. Ramage, "Dowelled structural connections in laminated bamboo and timber," *Composites Part B: Engineering*, vol. 90, pp. 232–240, 2016.
[18] J. J. Zahn and D. R. Rammer, "Design of glued laminated timber columns," *Journal of Structural Engineering*, vol. 121, no. 12, pp. 1789–1794, 1995.
[19] Z. Li, M. He, D. Tao, and M. Li, "Experimental buckling performance of scrimber composite columns under axial compression," *Composites Part B: Engineering*, vol. 86, pp. 203–213, 2016.
[20] P. Luna, C. Takeuchi, C. Alvarado, and I. Moreno, "Glued laminated Guadua angustifolia bamboo columns," in *I International Symposium on Genetic Resources of Bamboos and Palms and III International Symposium on Ornamental Palms*vol. 1003, pp. 125–130, 2013.
[21] H.-t. Li, J.-W. Su, Q.-S. Zhang, A. J. Deeks, and D. Hui, "Mechanical performance of laminated bamboo column under axial compression," *Composites Part B: Engineering*, vol. 79, pp. 374–382, 2015.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6625435, 11 pages
https://doi.org/10.1155/2020/6625435

Research Article

Forecast and Early Warning of Regional Bus Passenger Flow Based on Machine Learning

Wusheng Liu,[1] Qian Tan ,[2] and Wei Wu[3]

[1]*Engineering Research Center of Catastrophic Prophylaxis and Treatment of Road & Traffic Safety of Ministry of Education, Changsha University of Science & Technology, Changsha 410114, China*
[2]*School of Traffic & Transportation Engineering, Central South University, Changsha 410075, China*
[3]*School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China*

Correspondence should be addressed to Qian Tan; tq.helen@csu.edu.cn

Received 3 November 2020; Revised 25 November 2020; Accepted 4 December 2020; Published 15 December 2020

Academic Editor: Hussein Abulkasim

This paper mainly forecasts the short-term passenger flow of regional bus stations based on the integrated circuit (IC) card data of bus stations and puts forward an early warning model for regional bus passenger flow. Firstly, the bus stations were aggregated into virtual regional bus stations. Then, the short-term passenger flow of regional bus stations was predicted by the machine learning (ML) method of support vector machine (SVM). On this basis, the early warning model for regional bus passenger flow was developed through the capacity analysis of regional bus stations. The results show that the prediction accuracy of short-term passenger flow could be improved by replacing actual bus stations with virtual regional bus stations because the passenger flow of regional bus stations is more stable than that of a single bus station. The accurate prediction and early warning of regional bus passenger flow enable urban bus dispatchers to maintain effective control of urban public transport, especially during special and large-scale activities.

1. Introduction

Urban public transport is a traffic mode to alleviate traffic congestion and make efficient use of road resources. To realize intelligent dispatching of buses, it is important for decision makers to know well the change law of bus passenger flow and accurately predict the passenger flow in the short term. Burst passenger flows cause a huge amount of traffic demand in a short time, which may bring great pressure to public security. Early warning of passenger flow is required to have some preparation by the administration. However, the urban bus stops lack real-time early warning tools with good accuracy nowadays.

The short-term bus passenger flow is affected by various random, complex, and space-varying factors. In areas with large passenger flow, it is difficult to forecast the short-term change of bus passenger flow. Currently, short-term traffic is generally predicted by traditional statistical methods, novel intelligent methods, and hybrid methods.

Originating from time series analysis in the 1980s, the traditional statistical methods are built on the data collected by manual surveys. Over the years, these methods have been evolving and intellectualized in the context of traffic flow prediction. In the forecast of short-term passenger flow at a bus station, the entry and exit volumes mainly come from the card swiping records of the automatic fare collection (AFC) system. Taking such continuous data as a time series, many models have been designed for traffic flow prediction based on statistical principles, including autoregressive (AR) model, moving average (MA) model, autoregressive integrated moving average (ARIMA) model, and seasonal ARIMA (SARIMA) model [1–4]. In addition, many have introduced k-nearest neighbors (k-NN) [5], nonparametric regression [6, 7], and Kalman filter [8] to predict short-term traffic flow.

Since the dawn of the big data era, the amount of public transport data has exploded with the application of novel techniques such as the integrated circuit (IC) card. The traditional statistical methods could no longer adapt to the

complex environment and changing passenger flow at bus stations. With the rapid development of computational intelligence (AI) and data mining, data-driven intelligent methods have become popular in the prediction of short-term passenger flow [9]. The novel intelligent methods include long short-term memory (LSTM) [10–12], neural network (NN) [13–20], random forest (RF) [21, 22], support vector machine (SVM) [23, 24], fusion convolutional LSTM (FCL Net) [25], agent-based model (ABM) [26], and Bayesian network [27].

Recent years have witnessed the emergence of many hybrid forecast methods, most of which are combinations of novel intelligent methods such as NN and ML. For instance, Ke et al. [25] fused the FCL Net into a new deep learning (DL) method for the projection of short-term passenger demand. Xiao et al. [28] developed a new hybrid forecast strategy for air transport demand, which couples singular spectrum analysis (SSA), adaptive network fuzzy inference system (ANFIS), and optimized particle swarm optimization (OPSO). Sun et al. [29] combined wavelet transform (WT) with SVM into a hybrid prediction model for passenger flow; the model decomposes, predicts, and reconstructs the data on passenger flow in three stages and inherits the merits of both WT and SVM. Tan et al. [30] put forward a total traffic flow prediction method based on NN, MA, exponential smoothing (ES), and autoregressive MA (ARIMA).

In addition, Hinton and Salakhutdinov [31] applied DL to solve short-term prediction. Hu et al. [32] optimized the parameters of support vector regression (SVR) through particle swarm optimization (PSO), introduced historical momentum to reduce the impact of noise in traffic flow data, and then established a PSO-SVR model for the forecast of short-term traffic flow. Doğan [33] designed the periodic outstaring and prediction (PCP) algorithm, adopted the algorithm to improve the training set of artificial neural network (ANN), and proved that the improved ANN could predict short-term traffic flow based on selected clusters. Bagloee et al. [34] proposed a hybrid ML-based method to solve the bilevel optimization problem. Han et al. [35] derived a hybrid, optimized LSTM from Nesterov accelerated adaptive moment estimation (Nadam) and stochastic gradient descent (SGD).

To sum up, the traditional statistical methods mostly treat the current traffic state as a linear combination of the previous states and errors. On the upside, the traffic flow can be predicted simply by mathematical statistics, with relaxed data requirements. On the downside, the traditional methods fail to reflect the randomness and nonlinearity of traffic flow, consume too much manpower and financial resources in data acquisition, and have a low accuracy in the prediction of traffic flows, especially that with sudden changes. The novel intelligent methods are better than the traditional methods in data fitting and prediction accuracy, but the computing process is much more complex. The hybrid methods generally consider the features of actual traffic flow. The prediction accuracy of these methods varies with the coupled algorithms. Compared with the intelligent methods, the hybrid methods are highly complicated. The advantages, disadvantages, and applicability of common short-term prediction methods are shown in Table 1.

Despite the abundant results on short-term forecast, only a few scholars have explored the prediction or early warning of the passenger flow at urban bus stations. Gong et al. [36] proposed an ARIMA model and a Kalman filter to predict the number of passengers waiting at a bus station. Han et al. [35] created a hybrid and optimized LSTM to project the bus passenger flow. Van Oort et al. [37] converted the IC card data to the number of passengers per line, constructed an origin-destination (OD) matrix between stations, and assigned the matrix to the network to reproduce the measured passenger flow. Kumar et al. [38] developed a bus travel-time prediction method that considers both spatial and temporal variations in travel time. Wu et al. [39] built a convolutional LSTM (ConvLSTM) model with a self-attention mechanism, which accurately predicts the travel time on each segment of a trip and the waiting time at each station. Considering the small size, strong time-variation, and extraction difficulty of short-term passenger flow at bus stations, some scholars have taken account of connected and autonomous vehicles to alleviate the variability of travel time [40, 41]. Albeit these efforts, it is still difficult to make realistic forecast of the short-term passenger flow at bus stations. Wang et al. [42] designed a new framework to solve the problem of sudden passenger flow early warning. Pereira et al. [43] detected overcrowding with a threshold-based method and defined the point whose arrivals exceed the 90% percentile as overcrowding point. Bai et al. [44] monitor passenger flow to display the distribution trend of real-time passenger flow with GIS technology. To sum up, the current research mainly focuses on the early warning of traffic flow, and there is little research on the monitoring and early warning methods of passenger flow at bus stops.

We firstly analyze the features of passenger flow at bus stations and propose a novel concept called regional bus station, which is aggregated from actual bus stations in this paper. Then, the SVM was introduced to predict the short-term passenger flow at regional bus stations. The results show that the prediction accuracy of short-term passenger flow could be improved by replacing actual bus stations with virtual regional bus stations. On this basis, we designed an early warning model for regional bus passenger flow, which monitors the passenger flow in important areas during the period of special activities (e.g., large events) and takes control measures in advance to ensure the smooth progress of these activities.

The remainder of this paper is organized as follows: Section 2 puts forward the concept of regional bus stations by analyzing the IC card data of Shenzhen from November 7 to December 4, 2016, and summarizes the features of regional bus passenger flow; Section 3 selects and trains the prediction variables and employs the SVM to make short-term prediction of the passenger flow at regional bus stations; Section 4 analyzes the accuracy of prediction results; Section 5 carries out the capacity analysis and derives the early warning model for regional bus passenger flow; Section 6 puts forward the conclusions.

Table 1: Advantages, disadvantages, and applicability of common short-term prediction methods.

Prediction method	Advantages	Disadvantages	Applicable conditions
ARIMA model	The model is simple and has the ability to correct local data trend	It is difficult to fit nonlinear problems	Medium- and short-term forecast
Exponential smoothing method	Flexible and simple operation	Less variables are considered and the accuracy of smoothing number is low	Medium- and short-term forecast
Trend extrapolation	The operation is simple and the fitting effect is good	It is difficult to guarantee the accuracy due to less variables	Short-term forecast
Multivariate regressive method	Multiple factors can be considered	Large amount of calculation and high requirement for data	Medium- and short-term forecast
SVM	The model is simple and the results need not be modified	It is difficult to consider the comprehensiveness of indicators	Medium- and short-term prediction of small samples
RF	It can process high-dimensional data without feature selection	The reliability of the attribute weight on the data is not high	Medium- and short-term prediction of small samples
LSTM	Strong nonlinear fitting ability	A large amount of data is needed for network training	Nonlinear prediction

2. Bus Station Distribution and IC Card Data Processing

2.1. Bus Station Data and IC Card Data. The IC card used in this paper is Shenzhen Tong. It is a kind of stored value card for consumption by Shenzhen bus and Shenzhen Metro, which is manufactured under the supervision of Shenzhen Transport Bureau and issued by Shenzhen public transport settlement management center. The bus data only displays the valid information such as the user card number, card swiping time, and bus license plate number. It is shown in Table 2.

We extracted the GPS data and the card data of passengers and mapped the bus stops according to the time-space relationship.

(1) The GPS records of each line for 20 days are selected, and the GPS records are sorted according to the license plate number and the return time of GPS to get the track of the bus.

(2) Select the track point closest to the card swiping time as the boarding point, calculate the distance between the boarding point and each bus stop on the bus operation line, and select the nearest bus stop with the actual distance less than 100 as the actual bus stop, as shown in Figure 1.

The GPS data of 28 days from November 7 to December 4, 2016, are used in this paper, which contain about 541,115,294 pieces of time and location information of 10,314 vehicles. Therefore, we can study the characteristics of working day and non working bus passenger flow.

The passenger flows of bus stations in Shenzhen were acquired from the IC card data during the bus operation time (6:00–22:00) from November 7 to December 4, 2016. Figure 2 shows the distribution of the number of card swipes on buses in Shenzhen on a typical day.

To pinpoint the bus line number of each card swipe, the geographical positioning system (GPS) data were matched with the location data of bus stations, as shown in Figure 3. Through the matching, the bus arrival time was obtained for each bus station. Then, the boarding station of each card user was identified by acquiring the boarding time and bus line number from his/her IC card and matching the bus line number with the arrival time obtained in the previous step.

According to the time-varying distribution of bus trips in many days, it is found that the bus trips have obvious peak characteristics, and the peak hours are concentrated at 7:00 a.m. and 18:00 p.m. 7:00–9:00 and 17:00–19:00 are selected as the morning and evening peak hours of bus travel, as shown in Figure 4.

2.2. Regional Bus Station. To facilitate the passenger flow prediction of all bus stations in a region of Shenzhen (E: 113.76°–114.62°; N: 22.45°–22.87°), the road network in the region was meshed into 3,612 1 km × 1 km grids. 933 grids were found to have bus stations. In this way, the 53,914 bus stations in the region were aggregated into 993 regional bus stations (the black spots in the lower part of Figure 5).

3. Short-Term Passenger Flow Prediction at Regional Bus Stations

3.1. Principles. The short-term regional passenger flow was predicted in the following steps:

Step 1. Encode all the grids, and count the number of stations in each grid.

Step 2. Taking 1 h as the time window, count the number of passengers boarding buses at each station from 6:00 to 23:00 in the four days from December 1 to December 4, 2016.

Step 3. Allocate the data in four time windows (19:00-20:00, 18:00-19:00, 17:00-18:00, and 16:00-17:00) of the first 3 days (December 1–3, 2016) to the training set and the data in the same time windows of the last day (December 4, 2016) as the test set. Train the data by the ML-based forecast procedure.

Table 2: Data format of Shenzhen Tong (bus).

User number	Card bill number	Time	License plate number
4742502	666759180	2016-08-06T11:24:31.000Z	CB372
4742503	667174358	2016-08-06T11:39:09.000Z	CB372
4742504	667290097	2016-08-06T10:41:19.000Z	CB372

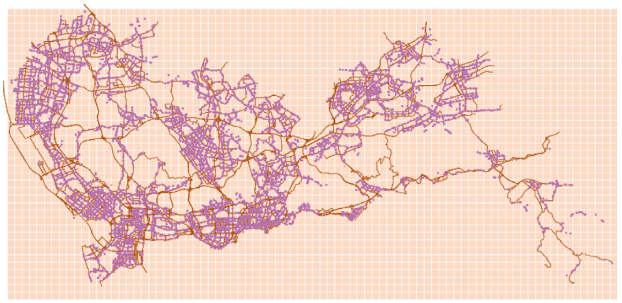


Figure 1: Location information of bus network and stations in Shenzhen.

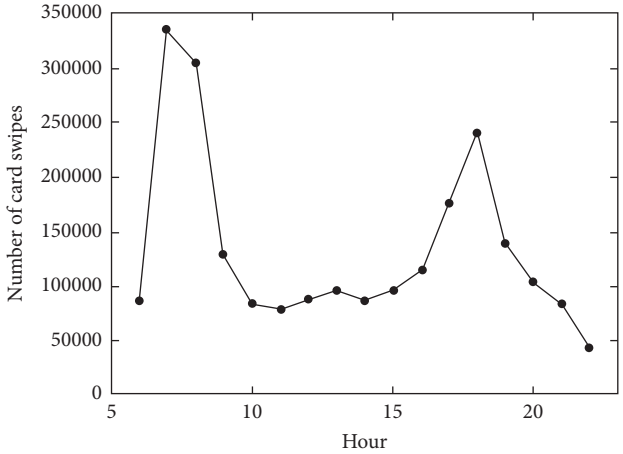


Figure 2: The distribution of the number of card swipes on buses in Shenzhen on December 4.

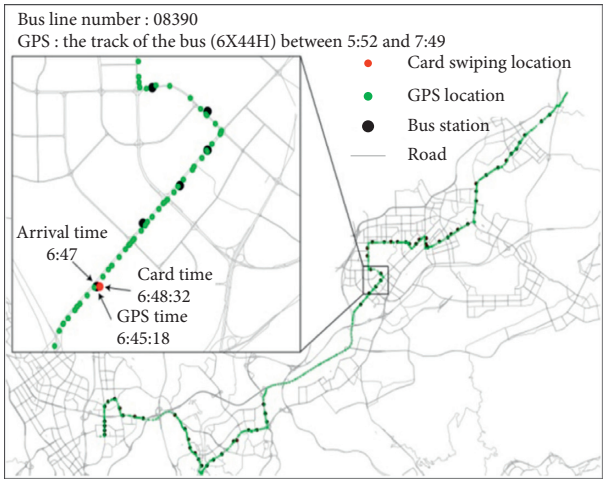


Figure 3: The acquisition of IC card information.

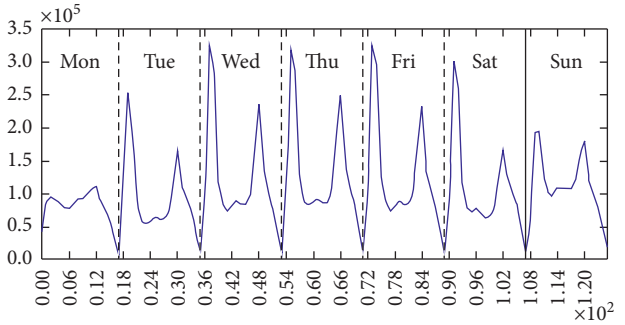


Figure 4: Time-varying distribution of bus travel.

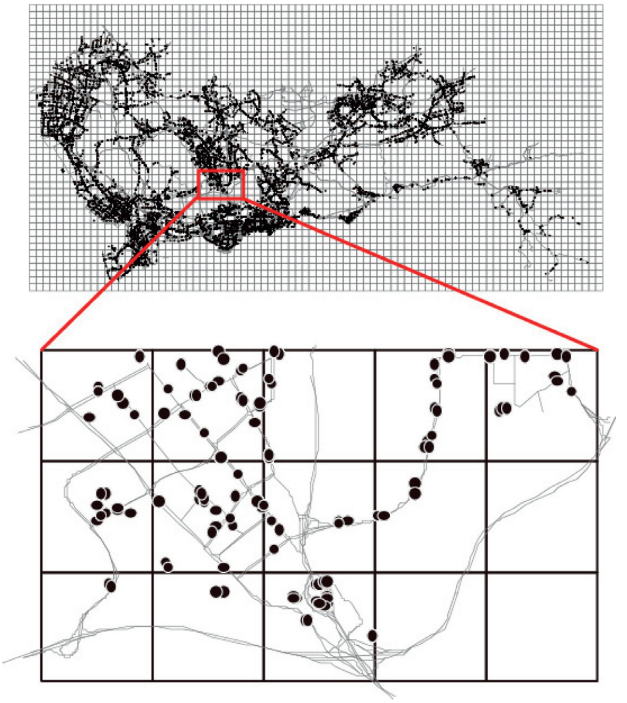


Figure 5: The distribution of bus stations in the region.

Step 4. Select the features of short-term regional bus passenger flow through ML, producing a set of valid features R.

Step 5. Select the top N features that influence the target period the most from the feature set R, and take them as the input of the regression model for SVM-based prediction.

3.2. Feature Training. Let $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}$, $x \in R^d$, $y \in R$ be the collected data, where (x_l, y_l) is

the training data of the lst day, d is feature dimension, x_l is the set of passengers boarding on the lst day under feature d, and y_l is the test data of the lst day.

Details of alternative features are shown in Table 3. The alternative feature for a target grid is the 1 km × 1 km grids around that grid, and the alternative feature set for the 993 grids in the region is whether the target grids are surrounded by 1 km × 1 km grids. If there are i grids in the set, then the number of boarding passengers should be counted in the previous j time windows of each grid in the set. Therefore, the feature dimension can be expressed as $d = i \times j$.

The number of alternative features varies from grid to grid, that is, the d value is a variable in $X^{(l\times d)} = (x_1, x_2, \ldots, x_l)^T$. Here, the value of j is set to 3. The details of alternative features are presented in Table 3.

Let T be the target time window. Then, the $N = 3$ most important features were selected through recursive feature elimination (Algorithm 1).

3.3. SVM-Based Regression Prediction. The regression prediction takes the N features that influence the target period the most as the inputs. Thus, the dimension of sample x_i is equal to N.

Let $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}, x \in R^N, y \in R$ be the training set. Then, the training samples were mapped into a higher dimension by function Φ, turning nonlinear regression inputs into high-dimensional inputs for linear regression.

The nonlinear mapping function can be expressed as $k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. Then, the SVM-based nonlinear regression can be defined as

$$\begin{aligned} \min_{\alpha^* \in R^{2l}} \quad & \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)k\left(x_i \cdot x_j\right) + \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) \\ & -y_i\sum_{i=1}^{l}(\alpha_i - \alpha_i^*), \\ \text{s.t} \quad & \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0, \alpha_i, \alpha_i^* \geq 0 \\ & i = 1, 2, \ldots, l, \alpha_i^* \leq C, \end{aligned} \tag{1}$$

where C is the penalty parameter and ε is the error term.

The optimum solution can be calculated as $\overline{\alpha}^{(*)} = (\overline{\alpha}_1, \overline{\alpha}_1^*, \ldots, \overline{\alpha}_l, \overline{\alpha}_l^*)$. The positive subsector $\overline{\alpha}_j > 0$ of $\overline{\alpha}$ or $\overline{\alpha}_j^* > 0$ of $\overline{\alpha}^*$ was chosen. Then, the $\overline{b}$ can be calculated by

$$\overline{b} = y_j - \sum_{i=1}^{l}(\overline{\alpha}_i^* - \overline{\alpha}_i)K\left(x_i, x_j\right) - \varepsilon. \tag{2}$$

Then, the decision function can be described as

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)k\left(x_i \cdot x\right) + \overline{b}. \tag{3}$$

Table 3: The details of alternative features.

Grid number	$t-3$	$t-2$	$t-1$	T
1	Pop_1_$t-3$	Pop_1_$t-3$	Pop_1_t_1	Grid
2	Pop_2_$t-3$	Pop_2_$t-3$	Pop_2_$t-1$	
...	...	...	...	
i	Pop_i_$t-3$	Pop_i_$t-3$	Pop_i_$t-1$	

For target grids, the number of passengers boarding in time window t can be forecasted based on the same N most important features on the given day. The radial basis function (RBF) was chosen as the kernel function: $k(x_i \cdot x) = \exp(-(\|x - x_i\|^2/2\sigma^2))$.

4. Prediction Results and Model Test

4.1. Prediction Results. Two time windows were randomly selected, namely, 11:00-12:00 and 19:00-20:00 on December 4, 2016, and the number of boarding passengers in the 993 grids was counted. The results predicted by our method are compared with the actual data in Figures 6 and 7.

4.2. Verification of Time Window. The target time window for prediction was set as 1 h. To verify the correctness of the time window, the real number of passengers boarding in 19:00-20:00 on December 4, 2016, was selected as the current period and compared with the real number of the previous period (18:00-19:00), the real number of the subsequent period (20:00-21:00), and the predicted value of the current period (19:00-20:00) (Figure 7).

As shown in Figures 8 and 9, the points were scattered, and the degree of linear fitting was poor, indicating that the number of passengers at each station in the grid fluctuates significantly within the time window of 1 h. Hence, it is necessary to set 1 h as the time window.

Moreover, the regression result of Figure 8 was $\beta_1 > 1$ (β_1 is the slope of the linear fitting line), that is, the result is greater than the actual boarding number in 19:00–20:00 of all grids. The regression result of Figure 9 is $\beta_1 < 1$, indicating that the actual boarding number in 19:00–20:00 of all grids is more than that in the following hour. Therefore, the data in the current time window cannot be replaced by the data of the previous or subsequent time window. In addition, in Figures 6 and 7, the slope β of linear fitting approaches to 1, indicating that it is better to use 1 hour as the time window. This further confirms the rationality of the 1 h time window.

4.3. Verification of Grid Size. We cross-grained the road network into 3,612 square grids with an area of 1 × 1 sq. km. Within the partitioned 3,612 square grids, 993 had at least one bus station inside. We cross-grained the road network with an area of 0.5 × 0.5 sq, and we found that 10% of the area had no bus stops. Moreover, the division of regional bus stations should be combined with the average length of urban road sections, and the area enclosed by each road section should be taken as the regional bus station as far as possible, and the average length of the road section in this paper is 0.78 km. In addition, we forecast the passenger flow

Input: training set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}, x \in R^d, y \in R$ and linear regression model
(1) Initialization: original feature set $S = \{1, 2, \ldots, D\}$, and feature sort set $R = \varnothing$
(2) While $S \neq \varnothing$ do
(3) Obtain the training samples of candidate feature set
(4) Obtain the weight of each feature by linear regression model, i.e., the coefficient of linear regression model $\omega_k, k = 1, 2, \ldots, |S|$
(5) Find out the features of the minimum score of sorting criteria: $p = \arg\min_k \omega_k$
(6) Update feature set $R = \{p\} \cup R$
(7) Exclude feature $S = S/p$ in S
(8) End while
Output: feature sort set R

Algorithm 1: Recursive feature elimination.

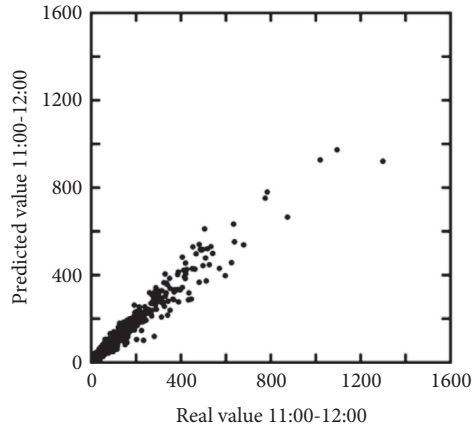


Figure 6: The predicted value vs. real value of the number of regional bus passenger flow (11:00-12:00).

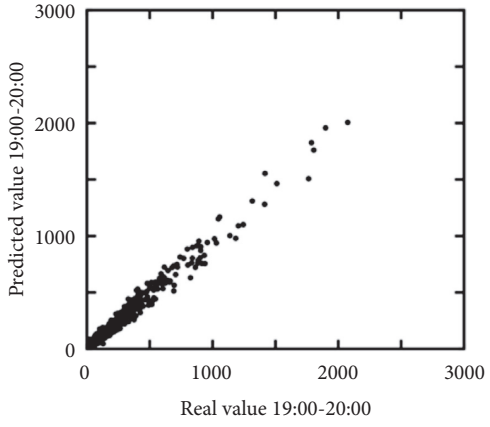


Figure 7: The predicted value vs. real value of the number of regional bus passenger flow (19:00-20:00).

of regional bus stations with different flows in different regions. We find that, with the increase of passenger flow of regional bus stations, the prediction accuracy increases, as shown in Figure 10. Therefore, it is appropriate to select 1×1 sq as the size of regional bus stations.

4.4. Model Test. It can be directly inferred from Figure 6 that the actual value and the predicted value follow a linear

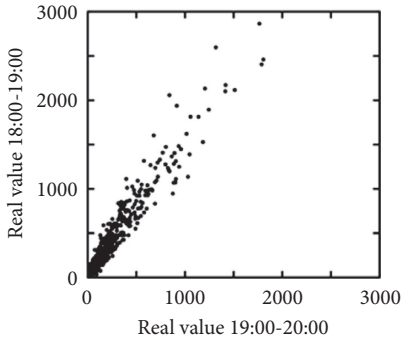


Figure 8: The real value vs. real value of the boarding number in different time windows (a).

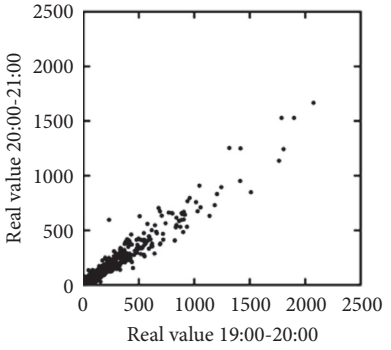


Figure 9: The real value vs. real value of the boarding number in different time windows (b).

relationship. Thus, unary linear regression was adopted for data fitting:

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad \varepsilon \sim \sigma^2, \tag{4}$$

where β_0 and β_1 both obey the normal distribution. If β_0 approaches to 0 and $\beta_1$approaches to 1, then the model has a high degree of regression. The dataset to be regressed was defined as $(x_i, y_i)(i = 1, 2, \ldots, n, n = 993)$.

Next, the linear regression problem was solved by the least squares (LS) method. The results of time window 11:00-12:00 were $\beta_0 = 4.20$ with its confidence being (2.42, 5.94) on the level of 95%, $\beta_1 = 0.90$ with its confidence being (0.885,

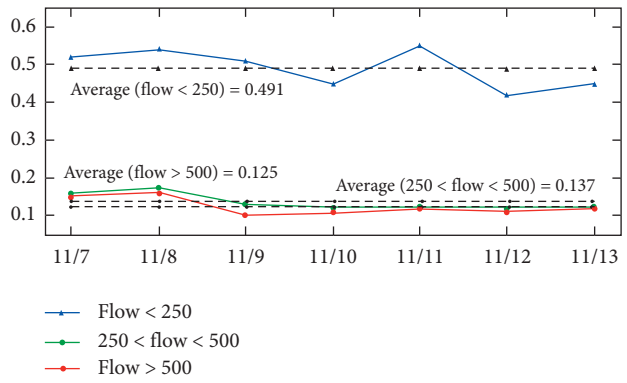


Figure 10: Prediction accuracy of station flow in different grids.

0.907) on the level of 95%, and coefficient of correlation $R^2 = 0.96$.

The predicted results were measured by three metrics: the mean absolute percentage error (MAPE), mean absolute error (MAE), variance of absolute percentage error (VAPE), and root mean square error (RMSE):

$$\begin{aligned} \text{MAPE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times 100\%, \\ \text{MAE} &= \frac{1}{n}\sum_{n=1}^{n}\left|\widehat{y}_i - y_i\right|, \\ \text{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right)^2}, \\ \text{VAPE} &= \text{Var}\left(\frac{\left|y_i - \widehat{y}_i\right|}{y_i}\right) \times 100\%. \end{aligned} \tag{5}$$

For the predicted results on the time window 11:00-12:00, the MAPE, MAE, VAPE, and RMSE were 0.18, 35.16, 0.12, and 65.42, respectively.

The results of time window 19:00-20:00 is that $\beta_0 = 2.23$ with its confidence being (−0.08, 4.55) on the level of 95%, β_1 =0.96 with its confidence being (0.95, 0.97) on the level of 95%, and coefficient of correlation $R^2 = 0.98$. The MAPE, MAE, and RMSE of these results were 0.16, 26.29, and 54.29, respectively.

The R^2 values of the two time windows demonstrate the significant correlation between variables X and Y. In both time windows, the β_0 value approached to 0, and the β_1 values approached 1. The error metrics of the results on the two time windows were both satisfactory. Therefore, the proposed ML method is favorable for predicting bus passenger flow in the short term.

4.5. Comparative Analysis of Prediction Accuracy between Single Bus Stop and Regional Bus Stop. In order to analyze the prediction accuracy of single bus stop and regional bus stop, we select the window of the world regional station, in which there are three bus stops: world window stop, Baishizhou stop, and Meilu Jinyuan stop. Bus passenger flow of November 28 is predicted, and the prediction results are shown in Figures 11 and 12.

The accuracy of regional bus station prediction results is significantly higher than that of single stops, especially for stops with relatively little passenger flow, the prediction accuracy is obviously low, and the prediction results have no practical application value.

5. Early Warning Model for Regional Bus Passenger Flow

5.1. Model Construction. The passenger capacity of a regional bus station can be calculated by

$$W = \max\left(C, P_j\right), \tag{6}$$

where C is the capacity of the regional bus station and P_j is the residual capacity of the regional bus line.

According to the *Traffic Engineering Manual*, the capacity C_i of station i can be computed by

$$C = \sum_{i=1}^{m} C_i = \sum_{i=1}^{m} \frac{3600 \times (g/c)_i \times R \times N_i}{\left[(g/c)_i \times D + t_c\right]}, \tag{7}$$

where $(g/c)_i$ is the green light ratio of the intersection in front of station i; R is the adjustment coefficient reflecting the degree of impact from bus arrival time and boarding time on station capacity (the empirical value is 0.833); D is the average boarding/alighting time (the value is generally 20–50 s); t_c is the average clearance time of bus station, i.e., the time required for the former bus to depart from and the current bus to arrive at the same position at the station (the value is generally 9–20 s); m is the total number of regional bus stations; N_i is the effective berth on the ith station.

The residual capacity P_j of regional bus line j can be calculated by

$$P_j = \sum_{j=1}^{n}\left[\frac{\left(S_j - Q_j\right) \times 60 \times H_j}{f_j}\right], \tag{8}$$

where S_j is the load factor of line j at acceptable service level (%), i.e., 80% of the rated passenger capacity; f_j is the departure interval of line j (min); H_j is the capacity of single bus online j (person) (the value is generally 60 persons); Q_j is the actual capacity of line j (%).

The early warning coefficient K can be calculated by

$$K = \frac{f(x)}{W \times \alpha}, \tag{9}$$

where $f(x)$ is the predicted short-term passenger flow of a regional bus stop and α is the proportion of IC card swiping passengers in the total number of bus passengers in the region.

5.2. Value of Early Warning Coefficient. Early warning generally refers to the prewarning of potential dangerous situation information. The early warning coefficient is used to judge what kind of situation needs to be issued and what

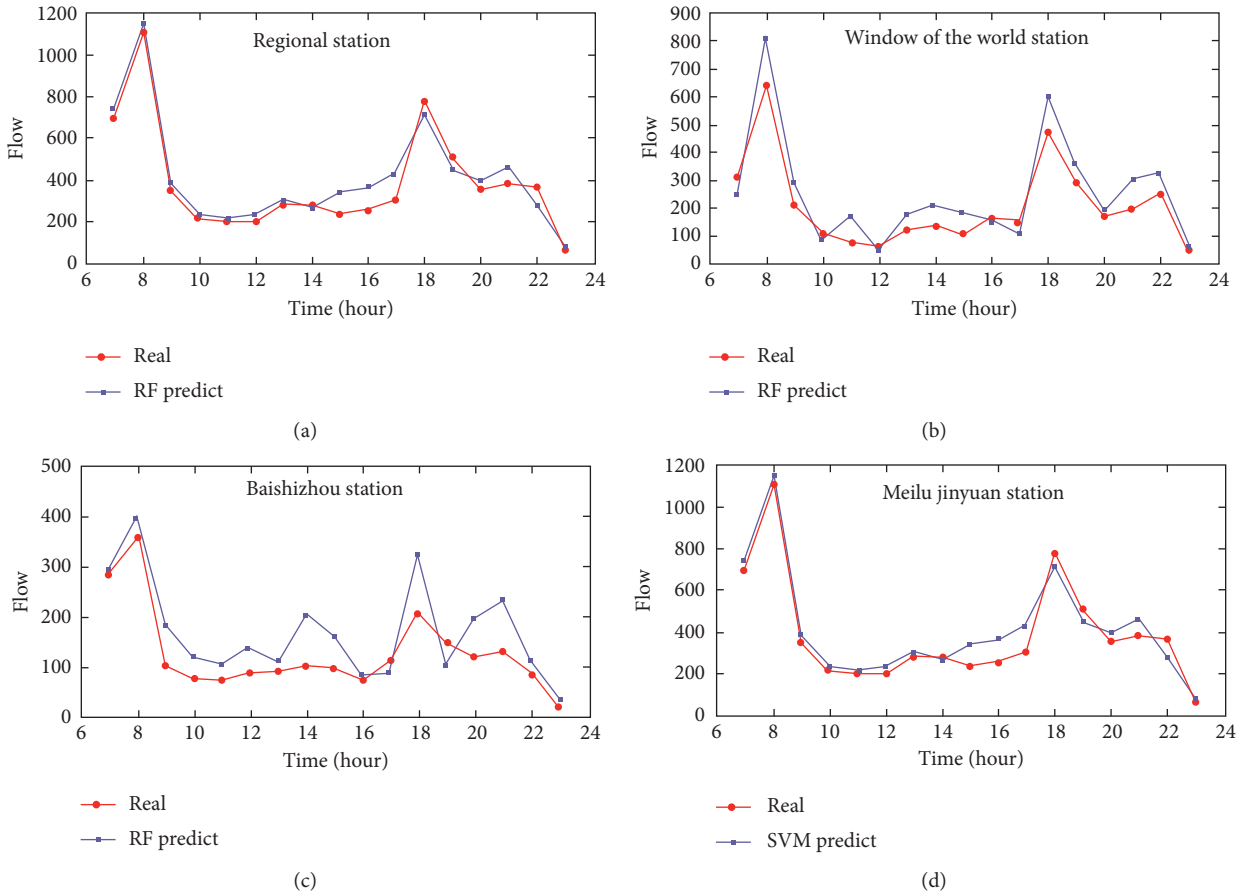


Figure 11: Prediction result between single bus stop and regional bus stop.

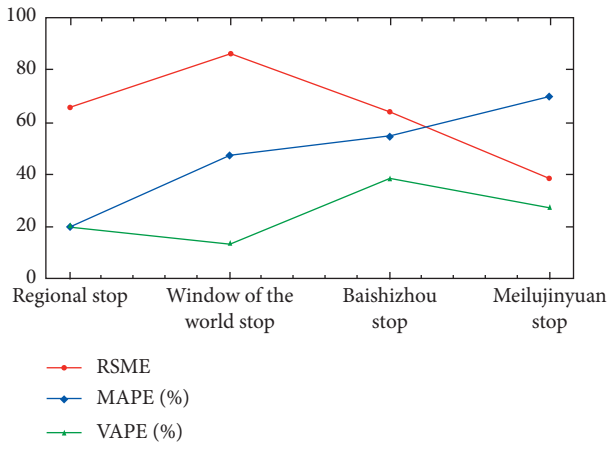


Figure 12: MAPE between single bus stop and regional bus stop.

degree of alarm is issued. In this paper, the early warning method adopts the method of combining the early warning index and the warning limit interval. If the value of the early warning index is in the corresponding warning limit interval, it corresponds to the alarm of this degree. Combined with the setting method of traffic flow warning interval [45], considering the level of public transport service, that is, considering the passenger comfort, reliability, and safety of operation service [46], the early warning interval of regional bus stops is divided into four levels. The first-level warning interval corresponds to the first level of service. The overall full-load rate of regional public transport vehicles is extremely low, and the passengers' comfort is very high, and the safety factor is very high. The second-level warning interval corresponds to the second-level service level. The overall full-load rate of regional public transport vehicles is not high, the waiting time of passengers is small, the congestion is low, and the safety factor is relatively high. The third level of warning interval corresponds to the third level of service, and the public transport vehicles are at the edge of full load. At this time, the safety factor is low, and the bus is crowded, which represents the general safety level; the fourth-level of warning interval corresponds to the fourth level of service, the safety factor is very low, the departure frequency cannot meet the needs of passengers, and the bus congestion is high, which represents the danger. The warning coefficients of the four levels are shown in Table 4.

If K falls in (0, 0.75), the region is safe; if K falls in (0.75, 0.9), the region needs to be monitored; if K surpasses 0.9, the region must be alarmed.

5.3. Case Analysis. The case analysis targets the grid at the junction of Shennan Avenue and Qiaocheng Road

Table 4: Early warning level of regional bus stops.

Early warning level	The first level	The second level	The third level	The fourth level
Warning interval	0–0.25	0.25–0.50	0.50–0.75	>0.75
Explanation	Safe	Relatively safe	Relatively danger	Danger

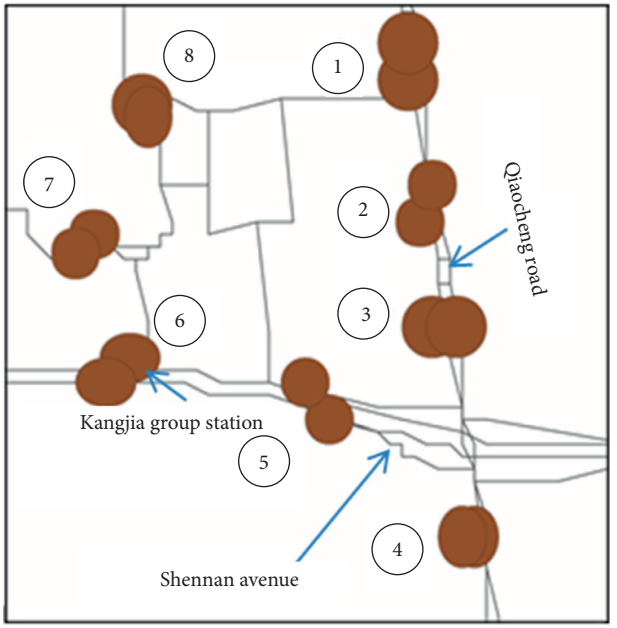


Figure 13: The target grid.

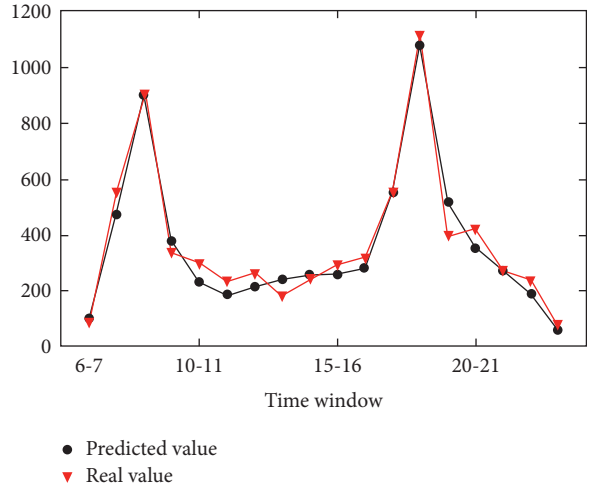


Figure 14: The predicted values vs. real values of the daily number of boarding passengers of 20 days (6:00–23:00).

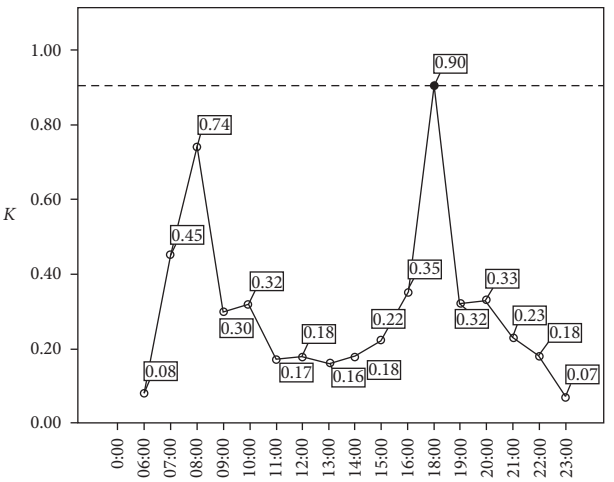


Figure 15: The early warning of passenger flow in the target grid.

(Figure 13). Shennan Avenue is one of the busiest roads in Shenzhen. Involving 16 bus stations and many bus lines, the target grid is highly representative of the public transport in Shenzhen. As shown in Figure 13, there are 35 bus lines at Kongjia Group Station on Shennan Avenue alone.

The proposed SVM-based algorithm was adopted to train and learn the historical boarding number of stations in the target grid and predict the daily number of boarding passengers of 20 days (6:00–23:00). Then, the predicted results were compared with the real values in the grid (Figure 14).

It can be seen from Figure 15 that our algorithm achieved a good prediction effect, as the predicted results followed the same trend as the real values. The MAE and MAPE were small, and the RMSE was smaller than 10%.

Through capacity analysis, the capacity of regional bus stations in 6:00–24:00 was 4,113 person times/h, while the residual capacity of bus lines was 3,840 in 6:00-7:00, 5,760 in 7:00–9:00, 4,032 in 9:00–17:00, 5,760 in 17:00–19:00, and 4,860 in 19:00–24:00. The regional bus capacity and K value in each period are presented in Figure 15.

According to the K values in Figure 15, the early warning coefficient in the target grid was 0.9 in 18:00-19:00. Hence, in this period, the bus capacity in the region is saturated and needs to be monitored. To reduce the K value, it is necessary to increase the regional bus capacity by stepping up departure frequency, improving the capacity of some stations, and expanding the effective parking spaces at stations.

6. Conclusions

This paper analyzes the IC card swiping data of all buses in Shenzhen during November 7–4, 2016 (6:00–22:00), and introduces the ML method of SVM to predict the short-term passenger flow of urban bus stations. The main conclusions are as follows:

(1) The training samples of candidate feature sets and the weight of each feature were through linear regression. The N features with the greatest impact on the target period were selected as the input of the regression model. The SVM-based regression prediction was adopted to predict the bus passenger flow

in the target time window. The model achieved high prediction accuracy when the time window is 1 h. The MAPE, MAE, and RMSE of these results were 0.16, 26.29, and 54.29, respectively.

(2) The capacity and early warning coefficient K of regional bus passenger was analyzed in detail. According to the K values, the capacity of some bus stations and bus lines in the target region should be improved to further optimize the service of public transport.

(3) The concept of regional bus stop is put forward in this paper; a suitable short-term prediction method for the passenger flow of regional bus stops is constructed, and the classification method and early warning coefficient of regional bus stop service level are developed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The project was supported by Open Fund of Engineering Research Center of Catastrophic Prophylaxis and Treatment of Road & Traffic Safety of Ministry of Education (Grant no. kfj180401, Changsha University of Science and Technology); Natural Science Foundation of Hunan Province, China (Grant no. 2019JJ40306); and National Natural Science Foundation of China (Grant no. 61773077).

References

[1] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Transportation Research Record*, vol. 722, pp. 1–9, 1979.

[2] B. M. Williams, P. K. Durvasula, and D. E. Brown, "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1644, no. 1, pp. 132–141, 1998.

[3] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1678, no. 1, pp. 179–188, 1999.

[4] B. Ghosh, B. Basu, and M. O'Mahony, "Multivariate short-term traffic flow forecasting using time-series analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 246–254, 2009.

[5] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[6] G. A. Davis and N. L. Nihan, "Using time-series designs to estimate changes in freeway level of service, despite missing data," *Transportation Research Part A: General*, vol. 18, no. 5-6, pp. 431–438, 1984.

[7] B. Yoon and H. Chang, "Potentialities of data-driven non-parametric regression in urban signalized traffic flow forecasting," *Journal of Transportation Engineering*, vol. 140, no. 7, pp. 14–27, 2014.

[8] A. Emami, M. Sarvi, and S. Asadi Bagloee, "Using Kalman filter algorithm for short-term traffic flow prediction in a connected vehicle environment," *Journal of Modern Transportation*, vol. 27, no. 3, pp. 222–232, 2019.

[9] M. Abadi, A. Agarwal, P. Barham et al., "Tensorflow: large-scale machine learning on heterogeneous distributed systems," 2016, https://arxiv.org/abs/1603.04467.

[10] N. C. Petersen, F. Rodrigues, and F. C. Pereira, "Multi-output bus travel time prediction with convolutional LSTM neural network," *Expert Systems with Applications*, vol. 120, pp. 426–435, 2019.

[11] X. Ran, Z. Shan, Y. Fang, and C. Lin, "An LSTM-based method with attention mechanism for travel time prediction," *Sensors*, vol. 19, no. 4, p. 861, 2019.

[12] Y. Hou and P. Edara, "Network scale travel time prediction using deep learning," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 45, pp. 115–123, 2018.

[13] M. Wajeed and V. Sreenivasulu, "Image based tumor cells identification using convolutional neural network and auto encoders," *Traitement du Signal*, vol. 36, no. 5, pp. 445–453, 2019.

[14] Z. L. Zhang, Y. F. Wang, and Y. Li, "Inventory control model based on multi-attribute material classification: an integrated grey-rough set and probabilistic neural network approach," *Advances in Production Engineering & Management*, vol. 14, no. 1, pp. 93–111, 2019.

[15] B. S. Kim and T. G. Kim, "Cooperation of simulation and data model for performance analysis of complex systems," *International Journal of Simulation Modelling*, vol. 18, no. 4, pp. 608–619, 2019.

[16] K. Krishna and B. Prakash, "Intrusion detection system employing multi-level feed forward neural network along with firefly optimization (FMLF2N2)," *Ingénierie des systèmes d information*, vol. 24, no. 2, pp. 139–145, 2019.

[17] Y. Li, D. Shi, and F. Bu, "Automatic recognition of rock images based on convolutional neural network and discrete cosine transform," *Traitement du Signal*, vol. 36, no. 5, pp. 463–469, 2019.

[18] Z. Zhang, Z. L. Guan, J. Zhang, and X. Xie, "A novel job-shop scheduling strategy based on particle swarm optimization and neural network," *International Journal of Simulation Modelling*, vol. 18, no. 4, pp. 699–707, 2019.

[19] K. Gorur, M. Bozkurt, M. Bascil, and F. Temurtas, "GKP signal processing using deep CNN and SVM for tongue-machine interface," *Traitement du Signal*, vol. 36, no. 4, pp. 319–329, 2019.

[20] H. M. Afify, K. K. Mohammed, and A. E. Hassanien, "Multi-images recognition of breast cancer histopathological via probabilistic neural network approach," *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 53–68, 2020.

[21] B. Yu, H. Wang, W. Shan, and B. Yao, "Prediction of bus travel time using random forests based on near neighbors," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 4, pp. 333–350, 2018.

[22] Q. Tan, X. Ling, M. Chen, H. Lu, P. Wang, and W. Liu, "Statistical analysis and prediction of regional bus passenger

flows," *International Journal of Modern Physics B*, vol. 33, no. 11, Article ID 1950094, 2019.

[23] U. Reddy, P. Dhanalakshmi, and P. Reddy, "Image segmentation technique using svm classifier for detection of medical disorders," *Ingénierie des systèmes d information*, vol. 24, no. 2, pp. 173–176, 2019.

[24] S. Deore and A. Pravin, "Histogram of oriented gradients based off-line handwritten devanagari characters recognition using SVM, K-NN and NN classifiers," *Revue d'Intelligence Artificielle*, vol. 33, no. 6, pp. 441–446, 2019.

[25] J. Ke, H. Zheng, H. Yang, and X. Chen, "Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 591–608, 2017.

[26] L.-M. Kieu, N. Malleson, and A. Heppenstall, "Dealing with uncertainty in agent-based models for short-term predictions," *Royal Society Open Science*, vol. 7, no. 1, Article ID 191074, 2020.

[27] Z. Li, M. N. Janardhanan, Q. Tang, and P. Nielsen, "Mathematical model and metaheuristics for simultaneous balancing and sequencing of a robotic mixed-model assembly line," *Engineering Optimization*, vol. 50, no. 5, pp. 877–893, 2018.

[28] Y. Xiao, J. J. Liu, Y. Hu, Y. Wang, K. K. Lai, and S. Wang, "A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting," *Journal of Air Transport Management*, vol. 39, pp. 1–11, 2014.

[29] Y. Sun, B. Leng, and W. Guan, "A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system," *Neurocomputing*, vol. 166, pp. 109–121, 2015.

[30] M. C. Tan, S. C. Wong, J. M. Xu, Z. R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 60–69, 2009.

[31] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[32] W. Hu, L. Yan, K. Liu, and H. Wang, "A short-term traffic flow forecasting method based on the hybrid PSO-SVR," *Neural Processing Letters*, vol. 43, no. 1, pp. 155–172, 2016.

[33] E. Doğan, "Short-term traffic flow prediction using artificial intelligence with periodic clustering and elected set," *Promet—Traffic & Transportation*, vol. 32, no. 1, pp. 65–78, 2020.

[34] S. A. Bagloee, M. Asadi, M. Sarvi, and M. Patriksson, "A hybrid machine-learning and optimization method to solve bi-level problems," *Expert Systems with Applications*, vol. 95, pp. 142–152, 2018.

[35] Y. Han, C. Wang, Y. Ren, S. Wang, H. Zheng, and G. Chen, "Short-term prediction of bus passenger flow based on a hybrid optimized LSTM network," *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, p. 366, 2019.

[36] M. Gong, X. Fei, Z. H. Wang, and Y. J. Qiu, "Sequential framework for short-term passenger flow prediction at bus stop," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2417, no. 1, pp. 58–66, 2014.

[37] N. Van Oort, T. Brands, and E. de Romph, "Short-term prediction of ridership on public transport with smart card data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2535, no. 1, pp. 105–111, 2015.

[38] B. A. Kumar, L. Vanajakshi, and S. C. Subramanian, "Bus travel time prediction using a time-space discretization approach," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 308–332, 2017.

[39] J. Wu, Q. Wu, J. Shen, and C. Cai, "Towards attention-based convolutional long short-term memory for travel time prediction of bus journeys," *Sensors*, vol. 20, no. 12, p. 3354, 2020.

[40] W. Wu, L. Huang, and R. Du, "Simultaneous optimization of vehicle arrival time and signal timings within a connected vehicle environment," *Sensors*, vol. 20, no. 1, p. 191, 2020.

[41] W. Wu, F. Zhang, W. Liu, and G. Lodewijks, "Modelling the traffic in a mixed network with autonomous-driving expressways and non-autonomous local streets," *Transportation Research Part E: Logistics and Transportation Review*, vol. 134, Article ID 101855, 2020.

[42] H. Wang, L. Li, P. Pan, Y. Wang, and Y. Jin, "Early warning of burst passenger flow in public transportation system," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 580–598, 2019.

[43] F. C. Pereira, F. Rodrigues, E. Polisciuc, and M. Ben-Akiva, "Why so many people? Explaining nonhabitual transport overcrowding with internet data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1370–1379, 2015.

[44] L. Bai, F. Z. Wang, and M. Zhang, "Urban rail transit network passenger flow monitoring and early warning system based on GIS," *Urban Rapid Rail Transit*, vol. 26, no. 6, pp. 56–59, 2013.

[45] L. Cheng, *Research on the Prediction and Early Warning of Short-Term Traffic Flow of Expressway*, Kunming University of Technology, Kunming, China, 2014.

[46] Transportation Research Board (TRB), *Transit Capacity and Quality of Service Manual*, 2nd edition, 2010.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 8819079, 7 pages
https://doi.org/10.1155/2020/8819079

Research Article

Application of an Estimation Method in the Lure System

Zhaoxia Huang

Department of Mathematics and Statistics, Ankang University, Ankang 725000, China

Correspondence should be addressed to Zhaoxia Huang; 281890896@qq.com

Received 21 September 2020; Revised 22 October 2020; Accepted 16 November 2020; Published 30 November 2020

Academic Editor: Hussein Abulkasim

The presence of set-valued mapping affects the stability of the output of the lure system, adding to the difficulty in observer design. To overcome the difficulty, the author mapped the system output error to the nonlinear term of the framer, creating a framer of the extended Luenberger structure, and analyzed the coordination of the error system by the monotonic system theory. On this basis, the interval observer was designed for the lure system. Then, the lure system and its observer systems were proved as asymptotically stable. Finally, it is proved that the observer system trajectory always followed the original state trajectory through the simulation under the different selections of set-valued mapping.

1. Introduction

Due to the wide range of uncertainties in the actual system, the research on uncertainty has received extensive attention. The reasons for the system uncertainty are as follows: modeling error, measurement error, environmental noise, unknown input, and control factors such as failure of the actuator or actuator, external disturbance, and parameter changes. Since the system state cannot be accurately estimated in real time, only the lower and the upper bounds about the state can be given, so the concept of the interval observer is born. The research and application of the interval observer can realize the determination of the state change interval and solve some uncertain problems based on the uncertain method. The main design methods of interval observers so far are limited, and most of them are for linear systems. Since most actual control systems are nonlinear in nature, the interval observer design theory for uncertain nonlinear systems has been developed. It is a very important research significance.

In recent years, there has been more and more research on interval observers, and different research methods have been proposed. The research objects have also expanded from linear time-invariant systems and planar systems to LPV systems, feedback linearization systems, chaotic systems, and Lipschitz nonlinear systems. This chapter mainly studies the design method of the interval observer for the lure nonlinear system. The lure nonlinear system is usually regarded as a relatively simple nonlinear system, but its application is wide and its output is uncertain, which leads to the difficulty of the observer design. This paper will output the error of the system, the nonlinear term of the frame phaser is mapped to the frame phaser of the extended Luenberger structure, and the coordination of the error system is analyzed based on the monotonic system theory. The interval observer design method of the lure system is given, and its algorithm is verified. Effectiveness.

State depicts the information of the dynamic behaviour of a system. To obtain the accurate state of the system is very difficult, due to the presence of various uncertainties. To overcome the difficulty, an effective solution is to take the input and output of the original system as the input to the new system and to use the output of the new system to estimate or reconstruct the state of the original system. The new system is known as the state observer. Many state observers were designed to eliminate the random disturbances [1–5], including several robust designs under bounded uncertainty [6–12].

With set-valued mapping, the lure system has uncertain output, which is a new type of uncertainty. In this paper, an interval observer method was proposed based on the lure system, aiming to disclose the relationship between system

structure and framer ($\underline{x}(t)/\overline{x}(t)$) configuration and find the way to worsen the state and minimize $x^+(t) - x^-(t)$. Specifically, the system output error was mapped to the nonlinear term of the framer, producing a framer of the extended Luenberger structure. Meanwhile, the coordination of the error system was analyzed by the monotonic system theory. On this basis, the interval observer design of the lure system was put forward and verified by an algorithm.

2. Problem Description

Suppose there is a system with lure type

$$\dot{x}(t) = Ax(t) + Br(t) - G\omega(t), \tag{1}$$

$$\omega(t) \in v(Hx(t)), \tag{2}$$

$$y(t) = Cx(t), \tag{3}$$

where, $x(t) \in R^n$ is the state, $y(t) \in R^r$ is the output, and $r(t) \in R^m$ is the input of the system; A, B, and C are real number matrices with proper dimensions; $G \in R^{n\times q}$ is the output matrix with proper dimensions; $v(\cdot)$ is the set-valued mapping; $H \in R^{q\times n}$ is the input matrix; and $\omega(t) \in R^q$ is the output of the set-valued mapping. Here, we assume B and G are column full rank, and C is row full rank. The set-valued mapping is defined below with an example.

As the fundamental concept of the differential inclusion theory, set-valued mapping was further described as follows.

Set-valued mapping exists in many actual systems. A typical example is the friction in mechanical systems and the diode in electronic systems. Set-valued mapping is needed to describe these basic elements in a specific manner. Hence, set-valued mapping and the ensuing differential inclusion are indispensable for the creation of an accurate model for dynamic systems.

In 1902, Stribeck discovered the variation curve of the friction coefficient when he explored dry friction (i.e., the friction without lubricant). This curve is now known as the Stribeck curve Figure 1. With the change in speed, the variation in friction coefficient goes through four phases, namely, static friction, Stribeck friction, Coulomb friction, and viscous friction. In the phase of static friction, there is no relative motion between the friction pairs; in the phase of Stribeck friction, relative motion occurs between the friction pairs, and the friction coefficient gradually declines. The four-phase process exhibits an obvious nonlinearity. Set-valued mapping appears in the phase of static friction. Despite the absence of relative motion, the friction force still exists in the same magnitude as the external force but points to the opposite direction.

Glocker described the friction force λ with the following formula of set-valued mapping:

$$\lambda \in F_f(\dot{q}) = -\mu\lambda_N \operatorname{sgn}(\dot{q}) + F_S(\dot{q}), \tag{4}$$

where $F_S(\dot{q})$ is a function of the speed; q is the displacement; μ is the friction coefficient; and λ_N is the positive pressure; we view $\operatorname{sgn}(\dot{q})$ as a set-valued mapping:

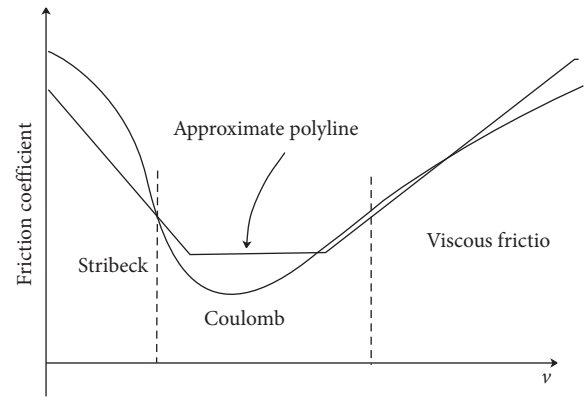


Figure 1: Stribeck curve.

$$\operatorname{sgn}(\dot{q}) = \begin{cases} 1, & \dot{q} > 0 \\ [-1, 1], & \dot{q} = 0 \\ -1, & \dot{q} < 0. \end{cases} \tag{5}$$

Only a controlled mechanical system with dry friction can be described as follows:

$$M\ddot{q} + D\dot{q} + Kq = Su + T\lambda, \tag{6}$$

where K, D, and M, are the stiffness matrix, damping matrix, and mass matrix of the mechanical system; q is the generalized speed containing both translation and rotation; S is an input matrix; and we assume u as the control variable; T is the friction matrix; λ is the vector composed of friction forces. The ith component of λ stands for the friction at the ith contact point, which can be expressed as the set-valued mapping below:

$$\lambda_i \in -\mu_i\lambda_{Ni}\operatorname{sgn}\left(T_i^T\dot{q}\right) + F_{Si}\left(T_i^T\dot{q}\right), \tag{7}$$

where T_i is the ith column of T and $T_i^T\dot{q}$ is the relative sliding of the ith contact point. The $M\ddot{q} + D\dot{q} + Kq = Su + T\lambda$ can be expressed as the following state space:

$$\begin{aligned} \dot{x} &= Ax + Gw + Bu, \\ z &= Hx, \\ y &= Cx, \\ w &\in -\varphi(z), \end{aligned} \tag{8}$$

where $x = (q^T\dot{q}^T)^T$ is the state; w and z are the output and input of $\varphi(\cdot)$; and y is the measurable output. It can be seen that set-valued mapping is another form of uncertainty. This paper focuses on the design of interval observers for the systems with such uncertainty.

The differences between differential inclusion and traditional differential equations are described below.

Let f: $[0, T] \times R^n \longrightarrow R^n$ be a known single-valued mapping ($T \in R(>0)$), $t \in [0, \mathrm{T}]$ be time, $x(t)$ be an unknown single-valued mapping of $[0, T] \longrightarrow R^n$, and $\dot{x}(t)$ be a derivative of $x(t)$ relative to t.

If F: $[0, T] \times R^n \longrightarrow R^n$ is a set-valued mapping, $x(t)$ is an unknown function of $[0, T] \longrightarrow R^n$. The solving of $x(t)$ from $\dot{x}(t) \in F(t, x(t))$ is a differential inclusion problem. If

F is only $R^n \longrightarrow R^n$, without explicitly containing t, then the differential inclusion becomes $\dot{x}(t) \in F(x(t))$.

The above formula is known as the time-invariant differential inclusion, while the previous formula is known as the time-varying differential inclusion.

If there is a Luenberger observer satisfying (1), then H must be row full rank. In the lure system, we assume the set-valued mapping $v(\cdot)$ as monotonic. In other words, $\omega_1 - \omega_2, Hx_1 - Hx_2 \geq 0$.

The author further investigated the Cauchy problem of (1) and (2) under the initial condition of $x(0) = x_0$ with $\omega \in v(\cdot)$. For the design of differential inclusion observer, the adaptive problem can be solved by two methods. First, it is assumed that, for any $\omega(t) \in v(Hx(t))$, the equation $\dot{x}(t) = Ax(t) + Br(t) - G\omega(t)$, $x(0) = x_0$, or that the set-valued mapping $v(\cdot)$ is a closed convex function with Lipshitz continuity. Second, we assume $v(\cdot)$ as maximal monotonic, such that the solution not only exists for $\dot{x}(t) = Ax(t) + Br(t) - G\omega(t)$, $x(0) = x_0$, but also is unique except for a zero measure set.

The lure type differential inclusion system has a closed-loop structure. The nonlinear part is concentrated on the feedback channel. It can be viewed as the product of nonlinear feedback of a linear system, provided that the nonlinear feedback satisfies the Popov integral inequality, i.e., $\int_{t_0}^{t} v^T(\tau) y(\tau) d_\tau \geq -r_0^2$ with v being the output of the nonlinear part and y being the output of the linear part.

3. Nature of Set-Valued Mapping

In the first two selections, the single-valued mapping is either continuous or measurable. In the approximate selection, a function sequence with good properties is selected such that it approximates $clF(x)$.

Lemma 1. *If set-valued mapping $v(\cdot): R^m \longrightarrow R^m$ is maximal monotonic, then $R^n \longrightarrow R^n$'s mapping $x \mapsto H^T v(Hx + h)$ is monotonic, $x \in R^n$, $h \in R^m$, $H \in R^{m \times n}$, and H is row full rank.*

Proof. Let $y_i \in x \mapsto H^T v(Hx + h)$ $i = 1, 2$. Then, there exists $\omega_i \in v(Hx_i + h)$ such that $y_i = H^T \omega_i$. Since H is row full rank, we only have to consider ω_i after determining y_i. Considering that

$$\begin{aligned} \langle y_1 - y_2, x_1 - x_2 \rangle &= \langle H^T\omega_1 - H^T\omega_2, x_1 - x_2 \rangle = \langle \omega_1 - \omega_2, Hx_1 - Hx_2 \rangle \\ \langle \omega_1 - \omega_2, (Hx_1 + h) - (Hx_2 + h) \rangle &= \langle \omega_1 - \omega_2, Hx_1 - Hx_2 \rangle \geq 0, \end{aligned} \tag{9}$$

we have $y_1 - y_2, x_1 - x_2 \geq 0$; that is, $x \mapsto H^T v(Hx + h)$ is monotonic.

The next step is to prove the maximal monotonicity of $x \mapsto H^T v(Hx + h)$. $v(Hx + h)$ is maximal monotonic. If $v(Hx)$ is monotonic, we only need to prove that $x \mapsto H^T v(Hx)$ is maximal monotonic.

Let $y_1 \in H^T v(Hx_1)$. If $y_2 \in lmH^T$, there exists $x_2 \in Rn$ such that $y_1 - y_2, x_1 - x_2 \geq 0$. Thus $y_2 \in lmH^T$, and there exists $\omega_2 \in Rm$ such that $y_2 = H^T\omega_2$. Since $y_1 \in H^T v(Hx_1)$, there exists $\omega_1 \in v(Hx_1)$ such that $y_1 = H^T\omega_1$. Hence, $y_1 - y_2, x_1 - x_2 \geq 0$ is $\omega_1 - \omega_2, Hx_1 - Hx_2 \geq 0$. Since $v(Hx)$ is maximal monotonic, we have $\omega_2 \in v(Hx_2)$; that is, $y_2 \in H^T v(Hx_2)$.

Lemma 2. *The linear and nonlinear parts of the loop in the lure type differential inclusion system are separable. If the nonlinear part satisfies the Popov integral inequality, then the necessary and sufficient condition for the closed-loop system to be ultrastable is that the linear part is positive and real; that is, the transfer function matrix is positive and real.*

Proof. Adequacy. Let $V(x) = x^T P x$ be a positive definite function. We have

$$\begin{aligned} \dot{V}(x) &= (Ax - Bv)^T Px + x^T P(Ax - Bv) \\ &= x^T\left(A^T P + PA\right)x - v^T B^T Px - x^T PBv \\ &= -x^T L^T Lx - 2v^T\left(C - K^T L^T\right)x \\ &= -x^T L^T Lx + 2v^T K^T L^T - v^T K^T Kv - 2v^T Cx \\ &\quad + v^T K^T Kv \\ &= -\|Lx - Kv\|^2 - v^T Cx - x^T C^T v + v^T Dv + v^T D^T v \\ &= -\|Lx - Kv\|^2 - 2v^T y \leq -2v^T y. \end{aligned} \tag{10}$$

Through the integration on both sides, we have

$$\int_0^t \dot{V} d_t = x^T(t) P x(t) - x^T(0) P x(0) \leq -2\int_0^t v^T(\tau) y(\tau) d_t \leq 2r_0^2. \tag{11}$$

According to the inequality above, we have

$$\lambda_m\|x(t)\|^2 \le \lambda_M\|x(0)\|^2 + 2r_0^2 \le K\left(\|x(0)\|^2 + r_0^2\right) \le K\left(\left\|x(0) + r_0\right\|\right)^2, \tag{12}$$

where $K = \max(\lambda_M, 2)$. Thus, we have $x(t, x_0, t_0) \le K(x_0 + r_0)$, indicating that the system is ultrastable.

Necessity. If there exists $\sigma_0 + i\omega_0$, in which $\sigma_0 > 0$, such that $G(\sigma_0 + i\omega_0) + G^*(\sigma_0 + i\omega_0)$ is not positive semidefinite. In other words, there exists $u_0 \in Cm$, where C is a complex number field, $u_0 = 1$, and $u_0^*[G(\sigma_0 + i\omega_0) + G^*(\sigma_0 + i\omega_0)]u_0 < 0$. Under the input $u(t) = \mathrm{Re}u_0 e^{(\sigma_0 + i\omega_0)t}$, since G(s) means the system is asymptotically stable and $f(t) = \sum_k \mathrm{Re}\, sF(p_k)\, e^{p_k t}$ $y(t) = \mathrm{Re}G(\sigma_0 + i\omega_0)u_0 e^{(\sigma_0 + i\omega_0)t}$

Considering $u^T(t)y(t) = \mathrm{Re}u_0^T e^{(\sigma_0 + i\omega_0)t}\mathrm{Re}\, G(\sigma_0 + i\omega_0)$ $u_0 e^{(\sigma_0 + i\omega_0)t}$, it can be concluded that $\int_0^t u^T(\tau)y(\tau)d_\tau \longrightarrow \infty$ through in-depth discussion. Thus, the Popov integral inequality $\int_{t_0}^t v^T(\tau)y(\tau)d_\tau \ge -r_0^2$ is valid.

4. Design Method for Lure System State Observer

For the lure type differential inclusion system (1), the Luenberger interval observers can be designed as follows:

$$\dot{x}^+ = Ax^+ + Br(t) + L\left(y - Cx^+\right) - G\omega^+(t), \tag{13}$$

$$\dot{x}^- = Ax^- + Br(t) + L\left(y - Cx^-\right) - G\omega^-(t), \tag{14}$$

where $\omega(t) \in v(Hx(t))$, with x being the state and ω being the output of set-valued mapping; $L \in R^{n\times r}$ is the observation gain which ensures that $A - LC$ is a Hurwitz matrix. According to the linear system, when (A, C) can be observed, there $A - LC$ is a Hurwitz matrix.

Subtracting $\dot{x}(t) = Ax(t) + Br(t) - G\omega(t)$ from $\dot{x}^+ = Ax^+ + Br(t) + L(y - Cx^+) - G\omega^+(t)$, we have

$$\begin{aligned} \dot{x}^+ - \dot{x} &= (A - LC)x^+ + G\omega(t) - G\omega^+(t), \\ \omega(t) &\in v(Hx(t)), \\ \omega^+(t) &\in v\left(Hx^+(t)\right), \\ y(t) &= Cx(t). \end{aligned} \tag{15}$$

Let $e^+ = x^+ - x$ be the observer error. Then, the above formula can be rewritten as follows:

$$\begin{aligned} \dot{e}^+ &= (A - LC)x^+ + G\left(\omega(t) - \omega^+(t)\right), \\ \omega(t) &\in v(Hx(t)), \\ \omega^+(t) &\in v\left(Hx^+(t)\right). \end{aligned} \tag{16}$$

Similarly, subtracting $\dot{x}^- = Ax^- + Br(t) + L(y - Cx^-) - G\omega^-(t)$ from $\dot{x}(t) = Ax(t) + Br(t) - G\omega(t)$, we have

$$\begin{aligned} \dot{x} - \dot{x}^- &= (A - LC)x^- + G\omega^-(t) - G\omega(t), \\ \omega(t) &\in v(Hx(t)), \\ \omega^-(t) &\in v(Hx^-(t)), \\ y(t) &= Cx(t). \end{aligned} \tag{17}$$

Let $e^- = x - x^-$ be the observer error. Then, the above formula can be rewritten as follows:

$$\begin{aligned} \dot{e}^- &= (A - LC)x^- + G(\omega^-(t) - \omega(t)), \\ \omega(t) &\in v(Hx(t)), \\ \omega^-(t) &\in v(Hx^-(t)). \end{aligned} \tag{18}$$

To find the L that makes the error system stable is our aim.

Lemma 3. *For differential inclusion systems (1)–(3), (13), and (14) are the Luenberger interval observers of the system (1):*

(i) There exists an L such that $(H, A - LC, G)$ is observable, strictly positive and real and controllable

(ii) $v(\cdot)$ is monotonic

(iii) The differential inclusions (16) and (18) have solutions

Proof. Let $z = Hx$, $\hat{z}^+ = H\hat{x}^+$, and $\hat{z}^- = H\hat{x}^-$ be the auxiliary outputs of the system (1), interval observer (13), and interval observer (14), respectively. Since $v(\cdot)$ is monotonic, we have $\omega - \omega^+, z - z^+ \ge 0$ and $\omega^- - \omega, z^- - z \ge 0$. Let $\mu_1 = \omega - \omega^+$, $\mu_2 = \omega^- - \omega$, $\zeta_1 = H(x^+ - x)$, and $\zeta_2 = H(x - x^-)$. Then, (16) and (18) can be expressed as follows:

$$\begin{aligned} \dot{e}^+ &= (A - LC)x^+ + G\mu_1, \\ \zeta_1 &= He^+, \\ \mu_1 &\in v(Hx) - v\left(Hx^+\right), \\ \dot{e}^- &= (A - LC)x^- + G\mu_2, \\ \zeta_2 &= He^-, \\ \mu_2 &\in v(Hx^-) - v(Hx). \end{aligned} \tag{19}$$

Let

$$v\left(e^+, x\right) = v\left(Hx^+\right) - v(Hx) = v(Hx) - v\left(H\left(x - e^+\right)\right). \tag{20}$$

Obviously, $0 \in v(0, x)$ is valid for all x. Let μ_1 and μ_2 be the inputs of the linear parts of systems (16) and (18), respectively, and ξ_1 and ξ_2 be the outputs of the linear parts of systems (16) and (18), respectively. Then, $\omega - \omega^+, z - z^+ \ge 0$, and $\omega^- - \omega, z^- - z \ge 0$ imply that $\mu_1, \zeta_1 \ge 0$ and $\mu_2, \zeta_2 \ge 0$. According to Lemma 2, $(H, A - LC, G)$ is real and strictly positive. Meanwhile, $e^- \longrightarrow 0$ and $e^+ \longrightarrow 0$, indicating that all dynamic systems are asymptotically stable and positive. Thus, we have $x^- \le x \le x^+$. This proves that (13) and (14) are interval observers of the system (1).

Lemma 4. *Let P be a given $m \times n$ real symmetric matrix, $U \in R^{m\times n}$, and $V \in R^{m\times l}$ be given column full rank matrices, and $U_\perp$ $V_\perp$ be the orthogonal complement matrix of U and V, respectively. Then, the necessary and sufficient conditions for the existence of a Q, such that*

$$P + UQV^T + VQ^TU^T < 0, \tag{21}$$

are

$$\begin{aligned}(V_\perp)^T PV_\perp < 0,\\ (U_\perp)^T PU_\perp < 0.\end{aligned} \tag{22}$$

The significance of Lemma 3 lies in V and U that appear in the same term in $P + UQV^T + VQ^TU^T < 0$, when matrix Q is unknown. By contrast, $U_\perp$ and $V_\perp$ appear in two independent inequalities, when matrix Q is known. This makes the problem easier to solve. The proof of the lemma and the general formula of Q satisfying $P + UQV^T + VQ^TU^T < 0$ are available in [13–17].

U is a reversible matrix because it is column full rank. In this case, $U_\perp$ is an element zero, and $(U_\perp)^T PU_\perp < 0$ should be removed from the lemma.

If U is a reversible matrix, the unique Q can be determined by $K = UQ$. In this case, if K is set to kV, then

$$P + UQV^T + VQ^TU^T = P + 2kVV^T. \tag{23}$$

According to the Finsker theorem, the necessary and sufficient condition for the existence of an k such that $P + 2kVV^T < 0$ is $V_\perp^T PV_\perp < 0$ [18].

Thus, we have the following theorem:

Theorem 1. *If G is column full rank, the necessary and sufficient conditions for the existence of an L such that $(H, A - LC, G)$ is strictly positive and real and are as follows:*

(1) $HG > 0$, i.e., H is row full rank

(2) $M_\perp^T C_\perp (NA + A^T N) C_\perp^T M_\perp < 0$ and $(C_\perp A^T G_\perp) C_\perp^T C_\perp (NA + A^T N) C_\perp^T C_\perp A^T G_\perp < 0$, where $M = C_\perp G_\perp$ and $N = H^T (HG)^{-1} H$

Conversely, if the lemma conditions are valid, we can obtain a P such that $G^T P = H$. According to condition (13) of the lemma, we have $C_\perp (PA + A^T P) C_\perp^T < 0$. Then, it can be seen from Lemma 3 that $(H, A - LC, G)$ is strictly positive and real.

Note that the unknown term in the inequality of condition (13) is the matrix X, and the inequality is linear relative to X. Thus, the inequality tool of the linear matrix can be adopted for the solution.

For practice, a simpler method is mentioned to solve the inequality $C_\perp (PA + A^T P) C_\perp^T < 0$ independently, which is irrelevant to $G^T P = H$.

Further results can be derived from Lemma 3 through the analysis of

$$\begin{aligned}C_\perp \Big[&\left(H^T (HG)^{-1} H + G_\perp X G_\perp^T\right) A \\ &+ A^T \left(H^T (HG)^{-1} H + G_\perp X G_\perp^T\right) \Big] C_\perp^T < 0.\end{aligned} \tag{24}$$

For simplicity, let $M = C_\perp G_\perp$ and $N = H^T (HG)^{-1} H$. Then, we have $M \in R^{(n-r)\times(n-q)}$, $N \in R^{n\times n}$, and N is a semipositive definite. Then, the above inequality can be rewritten as follows:

$$C_\perp \left(NA + A^T N\right) C_\perp^T + MXG_\perp^T AC_\perp^T + C_\perp A^T G_\perp XM^T < 0. \tag{25}$$

According to Lemma 3, the necessary and sufficient conditions for $C_\perp (NA + A^T N) C_\perp^T + MXG_\perp^T AC_\perp^T C_\perp^T + C_\perp A^T G_\perp XM^T < 0$ to be valid are

$$\begin{aligned}M_\perp^T C_\perp \left(NA + A^T N\right) C_\perp^T M_\perp < 0,\\ \left(C_\perp A^T G_\perp\right) C_\perp^T C_\perp \left(NA + A^T N\right) C_\perp^T C_\perp A^T G_\perp < 0.\end{aligned} \tag{26}$$

5. Simulation

The following lure system is established for simulation:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-10 & -3 & -1\\ 6 & -5 & 4\\ 1 & 0 & -9\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} - \begin{bmatrix}1\\ 2\\ 2\end{bmatrix}\omega + \begin{bmatrix}2\\ 0\\ 1\end{bmatrix} r, \tag{27}$$

where $y = x_1$, $v(\cdot)$ can be expressed as follows:

$$v\left(x_1 + 3x_2 + 2x_3\right) = \begin{cases} x_1 + 3x_2 + 2x_3 + 3\operatorname{sgn}\left(x_1 + 3x_2 + 2x_3\right), & x_1 + 3x_2 + 2x_3 \neq 0,\\ [-3, 3], & x_1 + 3x_2 + 2x_3 = 0,\end{cases} \tag{28}$$

where

$$\begin{aligned}A &= \begin{bmatrix}-10 & -3 & -1\\ 6 & -5 & 4\\ 1 & 0 & -9\end{bmatrix},\\ B &= \begin{bmatrix}2\\ 0\\ 1\end{bmatrix},\\ C &= \begin{bmatrix}1 & 0 & 0\end{bmatrix},\\ G &= \begin{bmatrix}1\\ 2\\ 2\end{bmatrix}.\end{aligned} \tag{29}$$

Let $L = \begin{bmatrix}-6 & 2 & 1\end{bmatrix}^T$. We have $A - LC = \begin{bmatrix}-4 & -3 & -1\\ 4 & -5 & 4\\ 0 & 0 & -9\end{bmatrix}$.

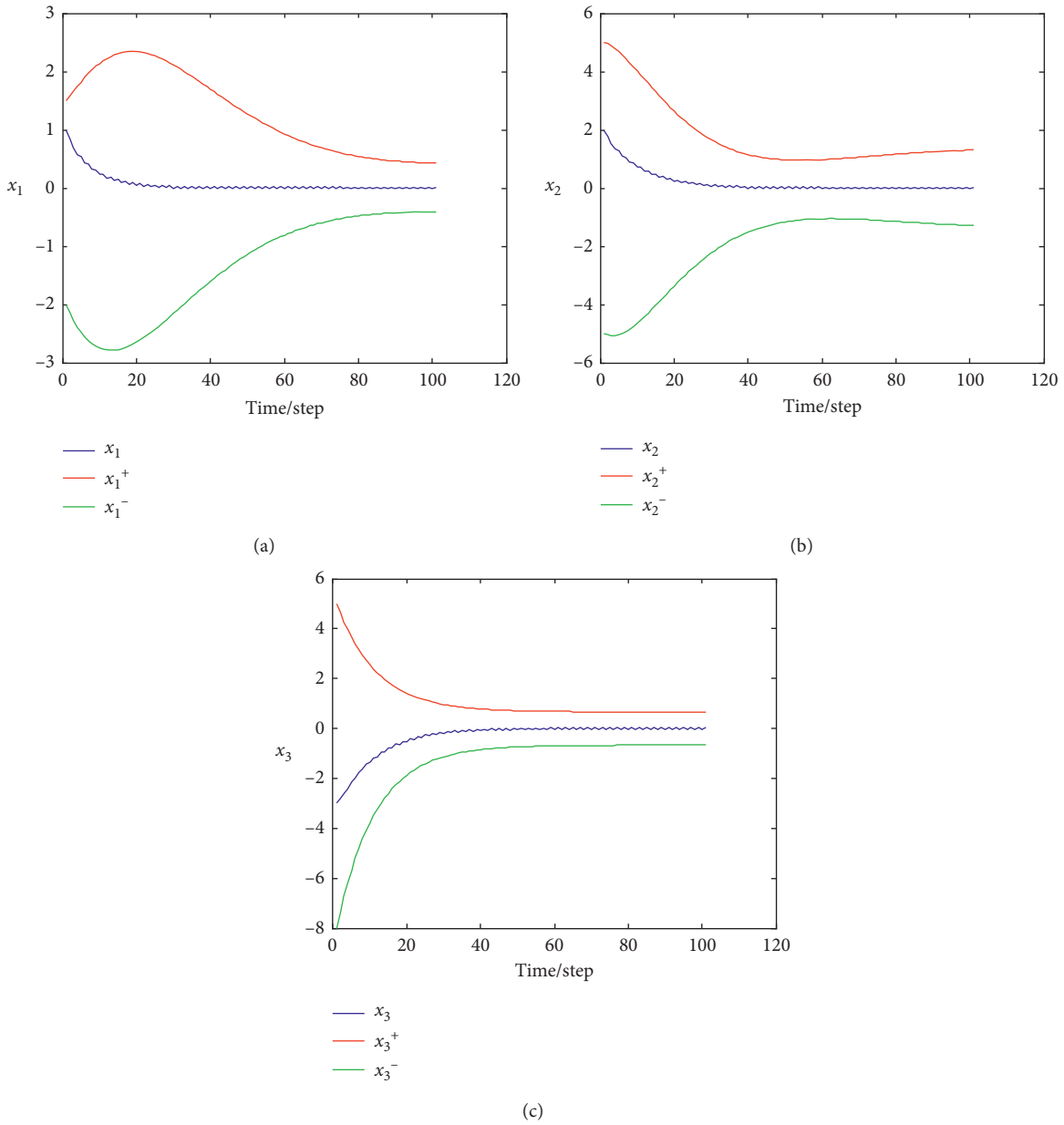


Figure 2: State trajectories of the lure system and the interval observers.

It can be calculated that $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}$ and $H = [\,1 \;\; 3 \;\; 2\,]$.

Figure 2 shows the estimated interval state of the system. It can be seen that $x_i^-(t) \le x_i(t) \le x_i^+(t)$, $i = 1, 2, 3$ at any time. This means the proposed interval observers (13) and (14) can estimate the interval state of the system well.

As shown in Figure 2, the lure system and its observer systems were asymptotically stable under different selections of set-valued mapping, and the observer system trajectory always followed the original state trajectory.

6. Conclusions

This paper explores the interval observer design in the lure system. Interval state estimation is a hot and frontier content in control theory research, and it has a wide range of applications in many engineering fields. At present, the international research on interval observers has just started, and the research objects are mainly focused on linear systems, with limited research methods and conclusions. Therefore, to carry out research on interval observers and control of uncertain systems not only has important

scientific significance and theoretical innovation but also has a certain engineering application value.

The presence of set-valued mapping affects the stability of the output of the lure system, adding to the difficulty in observer design. To overcome the difficulty, the author mapped the system output error to the nonlinear term of the framer, creating a framer of the extended Luenberger structure, and analyzed the coordination of the error system by the monotonic system theory. On this basis, a design method was proposed for the interval observer of the lure system. Then, the lure system and its observer systems were proved as asymptotically stable. Finally, it is proved that the observer system trajectory always followed the original state trajectory through the simulation under the different selections of set-valued mapping.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We are grateful to Prof Qian Fucai who made analysis for the paper. We also greatly appreciate Prof. Qian's useful discussion. This work was supported by the Foundation for Higher level Talents in Ankang University (2019AYQDZR01).

References

[1] T. Koide and T. Kodama, "Generalization of uncertainty relation for quantum and stochastic systems," *Physics Letters A*, vol. 382, no. 22, pp. 1472–1480, 2018.

[2] A. Gupta, C. Chakraborty, and B. Gupta, "Monitoring of epileptical patients using cloud-enabled health-IoT system," *Traitement du Signal*, vol. 36, no. 5, pp. 425–431, 2019.

[3] G. Jin, K. Bai, Y. Zhang, and H. He, "A smart water metering system based on image recognition and Narrowband Internet of Things," *Revue d'Intelligence Artificielle*, vol. 33, no. 4, pp. 293–298, 2019.

[4] B. Loutfi, Z. Samir, D. Ali, and G. M. Zinelaabidine, "Real time implementation of type-2 fuzzy backstepping sliding mode controller for twin rotor MIMO system (TRMS)," *Traitement du Signal*, vol. 36, no. 1, pp. 1–11, 2019.

[5] Y. Li, "Design and implementation of intelligent travel recommendation system based on internet of things," *Ingénierie des systèmes d'information*, vol. 23, no. 5, pp. 159–173, 2018.

[6] X. Li, C. Lin, and X. Xu, "A target tracking model for enterprise production monitoring system based on spatial information and appearance model," *Traitement du Signal*, vol. 36, no. 4, pp. 369–375, 2019.

[7] D. Efimov, L. Fridman, T. Raïssi, A. Zolghadri, and R. Seydou, "Interval estimation for LPV systems applying high order sliding mode techniques," *Automatica*, vol. 48, no. 9, pp. 2365–2371, 2012.

[8] Z. Zhang and S. Xu, "Observer design for uncertain nonlinear systems with unmodeled dynamics," *Automatica*, vol. 51, pp. 80–84, 2015.

[9] F. Zhu, "State estimation and unknown input reconstruction via both reduced-order and high-order sliding mode observers," *Journal of Process Control*, vol. 22, no. 1, pp. 296–302, 2012.

[10] J. Bodapati, V. Sajja, N. B. Mundukur, and N. Veeranjaneyulu, "Robust cluster-then-label (RCTL) approach for heart disease prediction," *Ingénierie des systèmes d information*, vol. 24, no. 3, pp. 255–260, 2019.

[11] I. R. Petersen, "Robust unobservability for uncertain linear systems with structured uncertainty," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1461–1469, 2007.

[12] K. Kumar and R. K. Mishra, "A robust mRMR based pedestrian detection approach using shape descriptor," *Traitement du Signal*, vol. 36, no. 1, pp. 79–85, 2019.

[13] P. Gahinet and P. Apkarian, "A linear matrix inequality approach toH∞ control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 4, pp. 421–448, 1994.

[14] W. P. M. H. Heemels, J. Daafouz, and G. Millerioux, "Observer-based control of discrete-time LPV systems with uncertain parameters $ $," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2130–2135, 2010.

[15] D. Ichalal and S. Mammar, "On unknown input observers for LPV systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 9, pp. 5870–5880, 2015.

[16] D. Zhao, Q. Zhang, X. Li, and L. Kong, "Event-triggered $H\infty$ control for continuous-time nonlinear system," *Advances in Neural Networks - ISNN 2015*, vol. 9377, pp. 62–70, 2015.

[17] S. Yasini, M. B. Naghibi Sitani, and A. Kirampor, "Reinforcement learning and neural networks for multi-agent nonzero-sum games of nonlinear constrained-input systems," *International Journal of Machine Learning and Cybernetics*, vol. 7, no. 6, pp. 967–980, 2016.

[18] Y. Wang, V. Gupta, and P. J. Antsaklis, "On passivity of a class of discrete-time switched nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 692–702, 2014.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 7210547, 7 pages
https://doi.org/10.1155/2020/7210547

Research Article

Feature Analysis on Mixed Traffic Flow of Manually Driven and Autonomous Vehicles Based on Cellular Automata

Xinghua Hu,[1] **Mengyu Huang**,[1] **and Jianpu Guo**[2]

[1]*School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China*
[2]*Chongqing Productivity Promotion Center, Chongqing 401147, China*

Correspondence should be addressed to Xinghua Hu; xhhoo@cqjtu.edu.cn

Received 17 August 2020; Revised 25 September 2020; Accepted 23 October 2020; Published 6 November 2020

Academic Editor: Hussein Abulkasim

This paper attempts to disclose the features of the mixed traffic flow of manually driven vehicles (MVs) and autonomous vehicles (AVs). Considering dynamic headway, the mixed traffic flow was modelled based on the improved single-land cellular automata (CA) traffic flow model (DHD) proposed by Zhang Ningxi. The established CA model was adopted to obtain the maximum flow of the mixed traffic flow and was analyzed under different proportions of AVs. On this basis, the features of the mixed traffic flow were summarized. The main results are as follows: the proportion of AVs has a significant impact on the mixed traffic flow; when the proportion reached 0.6, the flow of the whole lane was twice that of the MV traffic flow. At a low density, the AV proportion has an obvious influence on mixed traffic flow. At a high density, the mixed traffic flow changed very little, as the AV proportion increased from 0 to 5. The reason is that the flow of the whole lane is constrained by the fact that MVs cannot move faster. However, when the AV proportion reached 0.8, the flow of the whole lane became three times that at the proportion of 0.6. At the speed of 126 km/h, the flow rate was 2.5 times the speed limit of 54 km/h. The findings lay a theoretical basis for the modelling of multilane mixed traffic flow.

1. Introduction

In traffic engineering, the traffic flow model is an important research tool to facilitate the understanding of complex traffic phenomena. For more than half a century, many have theorized the traffic flow of manually driven vehicles (MVs), creating numerous traffic flow models and traffic flow prediction models [1–5]. Among them, the traffic flow model based on cellular automata (CA) stands out for its abilities to reproduce various complex phenomena in the traffic flow system and reflect the features of traffic flow.

In recent years, artificial intelligence (AI) and machine learning (ML) have made rapid progress [6, 7]. Thanks to the constant updates of algorithms and sensing techniques, autonomous vehicles (AVs) are poised to make up an important part of road traffic. In addition to their excellence in driving, AVs may bring a huge impact on the safety and efficiency of the operation of the traffic system [8–10]. However, the proportion of AVs will be limited early after their deployment. MVs and AVs are expected to travel together on roads for a long period of time.

The mixture of MVs and AVs will complicate the features of the traffic flow, which originally only involves MVs. The Institute of Electrical and Electronic Engineers (IEEE) predicted that the AV proportion will reach 75% by 2040. With the advancement of autonomous driving technology, the growth in the number of AVs will gradually pick up speed, due to their ultrashort reaction delay, low requirement for spacing or headway, and broad spectrum of speed. The mixed traffic flow of MVs and AVs will exert an enormous influence on the traditional road traffic system.

So far, only a few scholars have explored the mixed traffic flow of MVs and AVs. In 2003, Bose and Ioannou [11] analyzed the flow densities of MVs and AVs and discussed the shock waves in the mixed traffic flow, using the car following model and the semiautomatic vehicle model. Davis [12] applied the car following model to simulate the traffic phenomena of the mixed traffic flow in the confluence

area of the ramp, evaluated the local stability of traffic flow in the presence of AVs, and suggested AVs be provided with interconnected information, such as to debottleneck the traffic congestion on the reduced road. Kim and Liu [13] presented the concept of cooperative autonomous driving, which provides the drivers with the traffic situation ahead, enabling them to make better decisions that favor the efficiency and safety of traffic flow.

Through modelling, Levin and Boyles [10] found that AVs improve the traffic pattern in links and intersections, and the degree of improvement depends on the proportion of such vehicles. However, Levin's model only takes account of the driver's reaction time, without considering the interaction between MVs and AVs. Sharma et al. [14] and van Lint et al. [15] simulated the driving behavior of drivers in the mixed traffic flow model and identified the features of mixed traffic flow but failed to measure the restriction effect of MVs on AVs. Ngoduy [16] noticed the significant improvement to the capacity of the traffic system and travel time, when the AV proportion in the mixed traffic flow reached 30%.

In China, Jiang and Wu [17] were the first to propose a single-land CA model involving AVs. However, Jiang's model retains the random deceleration step reflecting the psychological changes of the driver, causing fluctuations to the speed of AVs. Inspired by Gipps' safety distance rule, Qiu et al. [18] put forward a CA model for AVs based on safety distance. Through simulation, Qin et al. [19] learned that the traditional traffic flow became more stable with the growing proportion of AVs. Liu et al. [20] set up a vehicle following model and proved that connected vehicles improve the stability of traffic flow. Many other simulation experiments [21–25] agree that introducing AVs to the road greatly improves the traffic flow and free flow speed, and the degree of improvement increases with the AV proportion.

Our research is based on the improved single-land CA traffic flow model (DHD) proposed by Zhang et al. [26]. The DHD assumes that daredevil and skillful drivers will resort to dynamic headway to adjust speed and follow the front car. The DHD model can simulate free flow, synchronous flow, and wide-range motion congestion in traffic flow, indicating that synchronous flow will also occur where there is no traffic bottleneck. The AVs, fully controlled by a computer, are similar to vehicles driven by the daredevil drivers in the DHD. In this paper, the mixed traffic flow of MVs and AVs is modelled based on the CA, in the light of the dynamic headway. After that, the maximum flow of the mixed traffic flow was analyzed under different proportions of AVs. In this way, the authors summarized the features of the mixed traffic flow.

2. Modelling

2.1. Motion Rules of MVs. CA traffic flow model is very popular because of its simplicity and ease of computer operation. One of the most important CA traffic flow models is the NaSch model proposed by Nagel and Schrekenberg [1] in 1992. Being a random single-lane CA traffic flow model, the NaSch model regards the lane as a pattern of one-dimensional (ID) discrete lattice points with a length of L and assumes that N vehicles are randomly distributed on the lane. At each moment, a lattice point is either empty or occupied by a vehicle. The state of each vehicle is characterized by its speed V_n $(n = 1, 2, \ldots, N)$, which falls within the interval $\{0, 1, 2, \ldots, V_{\max}\}$, where $V_{\max}$ is the maximum allowable speed. The position of the n-th vehicle is represented as X_n. The distance between the n-th vehicle and the $n+1$ previous vehicles is represented as $d_n = X_{n+1} - X_n - 1$. From t to $t+1$, the position and state of all vehicles are updated in parallel at each time step by the following rules:

Step 1. Acceleration: $V_n(t+1) = \min(V_n(t) + 1, V_{\max})$

Step 2. Deterministic deceleration: $V_n(t+1) = \min(V_n(t), d_n(t))$

Step 3. Probabilistic random deceleration: $V_n(t+1) = \max(V_n(t) - 1, 0)$

Step 4. Position updates: $X_n(t+1) = X_n(t) + V_n(t+1)$

Albeit its simplicity, the NaSch model can depict actual traffic phenomena like spontaneous congestion and stop-and-go wave in crowded traffic. In this paper, the motion rules of MVs are the same as those of ordinary drivers in the DHD; that is, the traffic flow evolves in accordance with the NaSch model.

2.2. Motion Rules of AVs. Considering the motions of the front car, the DHD focuses on the change law of the synchronization between the front and rear cars. Unlike the NaSch model, the DHD no longer treats the headway at the current moment as that at the next moment. Let $V_{n-1}(t+1)$ be the expected speed of the front vehicle. Then, Step 2 of the NaSch model can be modified as

$$V_n(t+1) = \min\left(V_n(t), d_n(t) + V_{n-1}(t+1)\right). \quad (1)$$

In this way, the headway increases from $d_n(t)$ to $d_n(t) + V_{n-1}(t+1)$, setting a rule for high-speed following. Obviously, AVs can follow the front car at a high speed just like daredevil drivers, owing to the numerous sensors onboard and advanced techniques of information processing. Therefore, the DHD was selected as the basis for the motion rules of AVs.

Moreover, Step 3 (probabilistic random deceleration) was deleted, for AVs are completely controlled by computers, which make no uncertain behavior as traditional drivers do [27]. Considering dynamic headway, the motion rules of AVs can be summed up as follows:

Step 1. Acceleration: $V_n(t+1) = \min(V_n(t) + 1, V_{\max})$

Step 2. Deterministic deceleration: $V_n(t+1) = \min(V_n(t), d_n(t) + V_{n-1}(t+1))$

Step 3. Position updates: $X_n(t+1) = X_n(t) + V_n(t+1)$

The AV proportion in the traffic flow was defined as ρ. If $\rho = 1$, the traffic flow is purely AVs that evolve according to the above three steps; if $\rho = 0$, the traffic flow is purely MVs that evolve according to the rules of the NaSch model; if $0 < \rho < 1$, the traffic flow is a mix of MVs and AVs, and the flow change varies with the ρ value.

3. Numerical Simulation

The mixed traffic flow of MVs and AVs on a 5000 m-long single-land road was simulated on MATLAB. The entire road was simulated under periodic boundary condition, and the vehicle density was strictly controlled: once a vehicle leaves the road, it will reenter the road from the starting point. The road was divided into 1000 cells, each of which is 5 m long. The cell length is approximately the average length of a vehicle. Each cell could be empty or occupied by a vehicle.

The vehicle speed was divided into 6 discrete levels: 0, 1, 2, 3, 4, and 5 cell(s) per time step. Level 0 means that the vehicle is at a standstill; level 5 means the vehicle travels across 5 cells in a time step, corresponding to the speed of 90 km/h. For MVs, because human drivers have some random behaviors during driving, they will cause braking measures, and the probability P of random deceleration was set to 0.1.

The traffic flow was initialized randomly for the simulation. N vehicles with different initial speeds and positions were generated on the road. That is, initial position, speed, and vehicle type are completely random. The total density of the road can be calculated as $\rho = N/L$.

3.1. Simulation of MV Traffic Flow. The flow-density diagram of MV traffic flow was plotted. Based on the speed limit $V_{\max} = 5$, the total number of MVs in the lane was determined, and then the density of the lane was obtained. The entire simulation lasted 20,000 time steps. The first 10,000 time steps were removed to eliminate the transient influence. Then, the average speeds of all MVs in each of the next 10,000 time steps were recorded and multiplied with the current density to produce 10,000 flow values. The above operations were repeated at different initial densities to create more points on the flow-density diagram. The flow-density diagram of MV traffic flow is presented in Figure 1.

As shown in Figure 1, the flow of MV traffic flow peaked at the density of 0.16, which corresponds to an actual flow of 1440 veh/h, an optimal density of 32 veh/km, and a critical speed of 45 km/h. The results shown in Figure 1 are consistent with actual traffic flow.

3.2. Simulation of Mixed Traffic Flow. Next, the flow density of the mixed traffic flow of MVs and AVs was simulated at different proportions of AVs. The simulation results are displayed in Figures 2–5.

As shown in Figures 2–5, the maximum possible flow increased with the AV proportion, indicating that introducing AVs can boost the flow of the whole lane. In Figure 2, the flow-density relationship is linear as density changed from 0 to 0.2 (i.e., actual density changed from 0 to 40 veh/km). Hence, the vehicles in the area of density 0–0.2 belong to the free flow state, and 40 veh/h is the critical density under the AV proportion of 0.2.

With the growing AV proportion, the critical density continued to increase. Once the density surpassed the critical density, the flow of the whole lane started to decline gently,

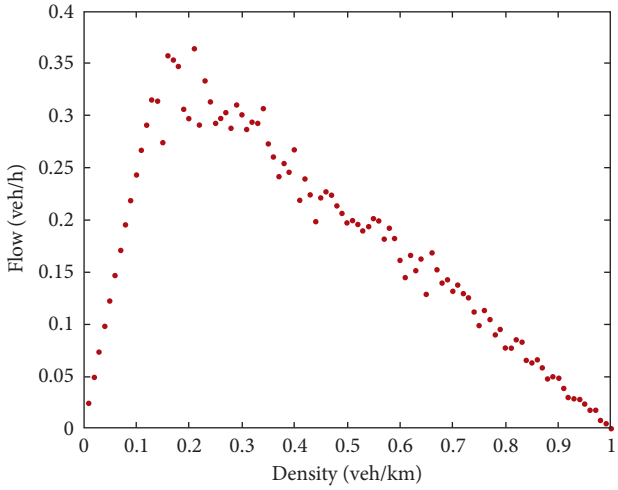


Figure 1: Flow-density diagram of MV traffic flow.

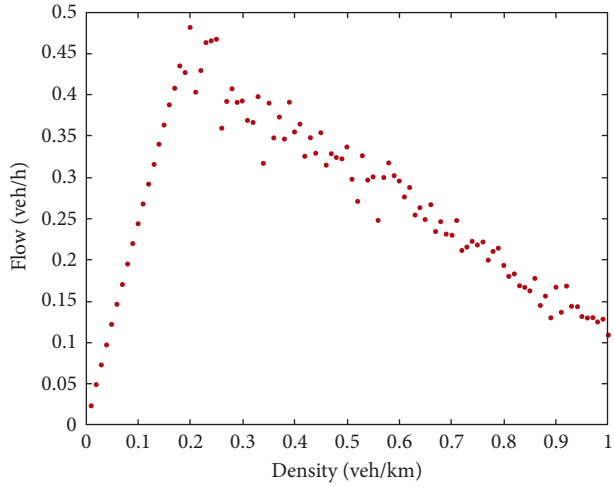


Figure 2: Flow-density diagram of mixed traffic flow at an AV proportion of 0.2.

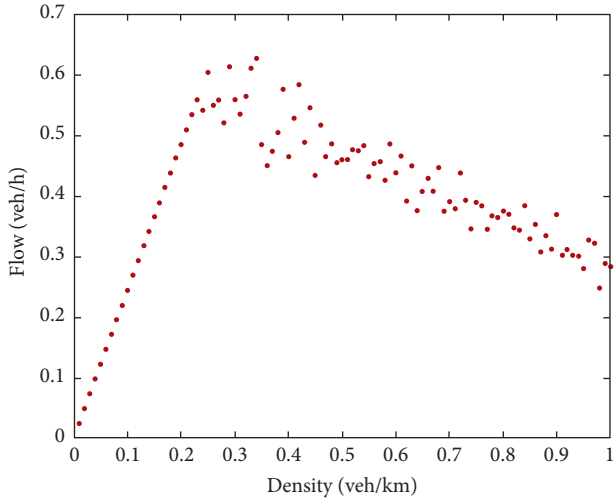


Figure 3: Flow-density diagram of mixed traffic flow at an AV proportion of 0.4.

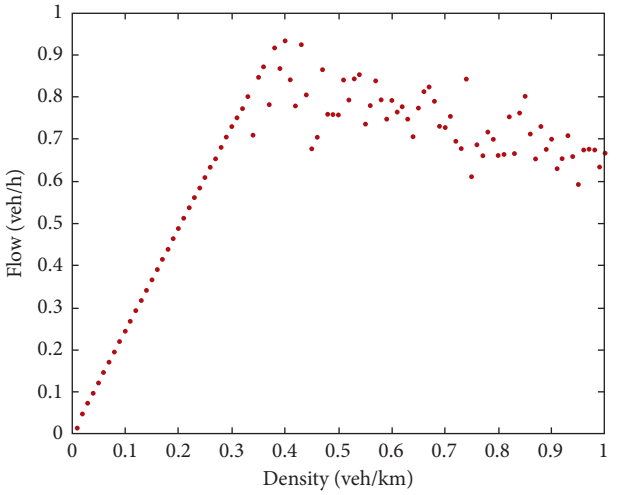


Figure 4: Flow-density diagram of mixed traffic flow at an AV proportion of 0.6.

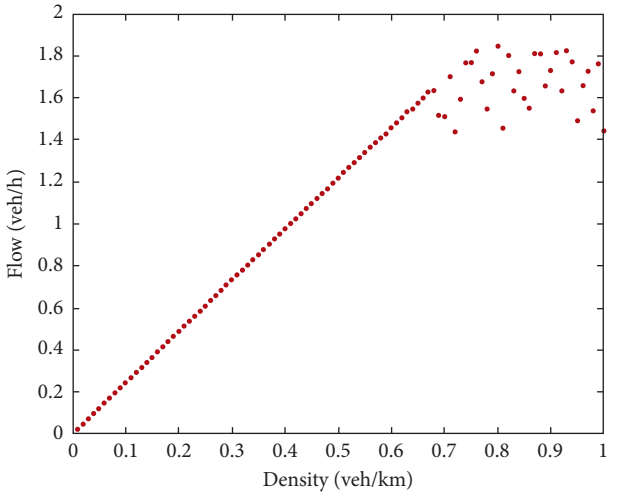


Figure 5: Flow-density diagram of mixed traffic flow at an AV proportion of 0.8.

and the points on the flow-density diagram became scattered. In addition, when the penetration reaches 0.6, the flow of the whole lane was twice that of the MV traffic flow.

3.3. Influence of AV Proportion on Mixed Traffic Flow. Figures 6–10 present the variation of the mixed traffic flow with AV proportions, as the density increased from 0.1 to 0.9 with a step length of 0.2.

As shown in Figures 6–10, the mixed traffic flow was linearly correlated with AV proportion at the density of 0.3; the mixed traffic flow peaked at 2700 veh/h at the AV proportion of 0.6; after AV proportion surpassed 0.6, the mixed traffic flow no longer changed, that is, not correlated with AV proportion. In addition, at the density = 0.9, the mixed traffic flow increased by 2 times, as the AV proportion grew from 0.6 to 0.8.

In our CA model, the AVs receive the expected speed of the front car and gain a wider driving space following the

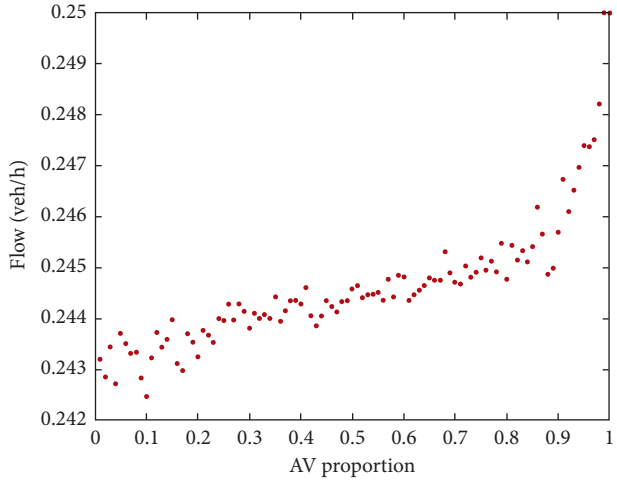


Figure 6: Flow-AV proportion diagram at density = 0.1.

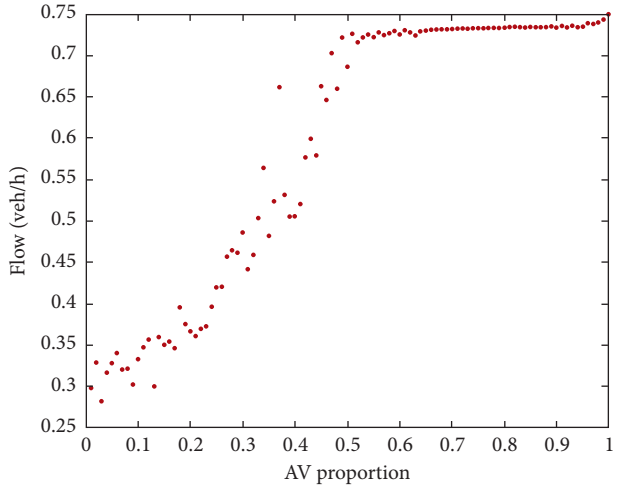


Figure 7: Flow-AV proportion diagram at density = 0.3.

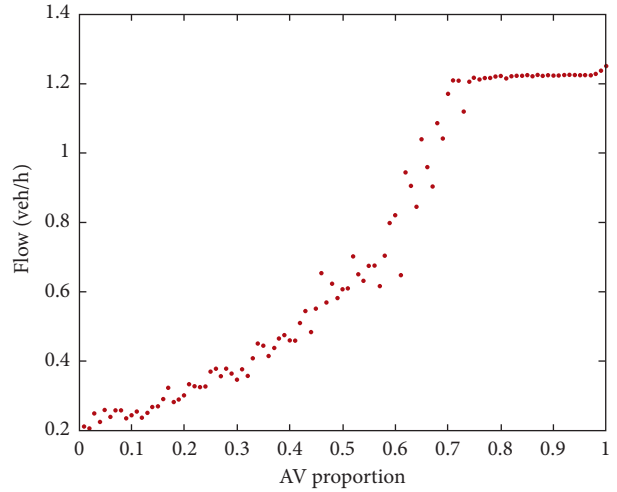


Figure 8: Flow-AV proportion diagram at density = 0.5.

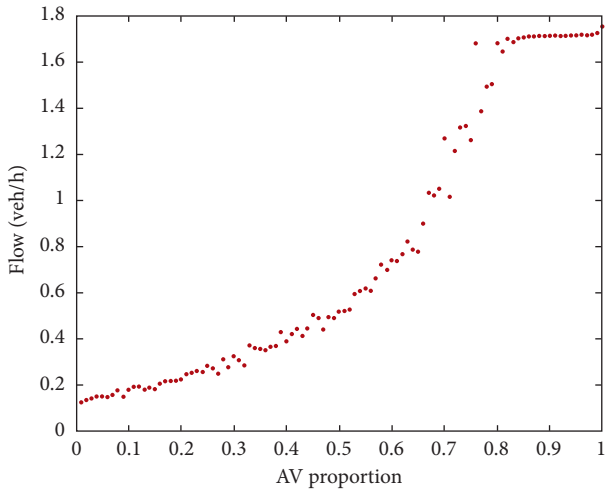


Figure 9: Flow-AV proportion diagram at density = 0.7.

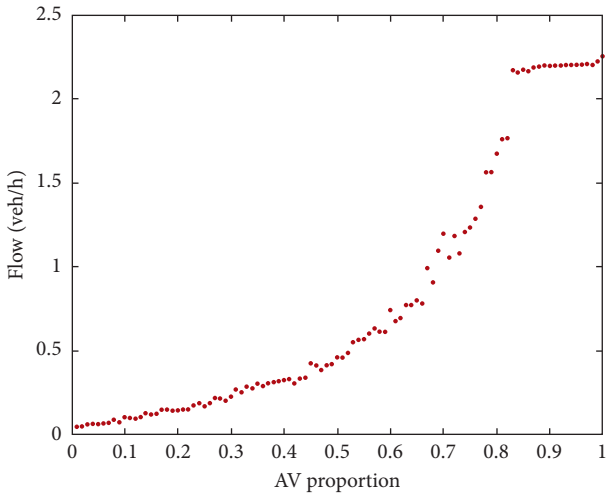


Figure 10: Flow-AV proportion diagram at density = 0.9.

strategy of dynamic headway. Therefore, the AVs can maintain the state of high-speed following. If the front car is an MV, the human driver will experience probabilistic random deceleration. This experience, coupled with the personality of the driver, makes it impossible for the front car to maintain high-speed following. Then, the MV indirectly limits the speed of the AVs behind. The limiting effect is obvious in dense mixed traffic flows with a low AV proportion. As shown in Figures 6–10, when it was smaller than 0.5, the AV proportion has little impact on mixed traffic flow, for the high-speed following of AVs was limited by the huge number of MVs.

If the front car and the rear car are both AVs, the two vehicles will drive on the road like carriages of a train, with a very small space between them, forming a self-organizing fleet in the mixed traffic flow. With the growth in the AV proportion, more and more AVs join the self-organizing fleet, which moves like the uniform flow in fluid mechanics. This explains why the critical density increased with the AV proportion.

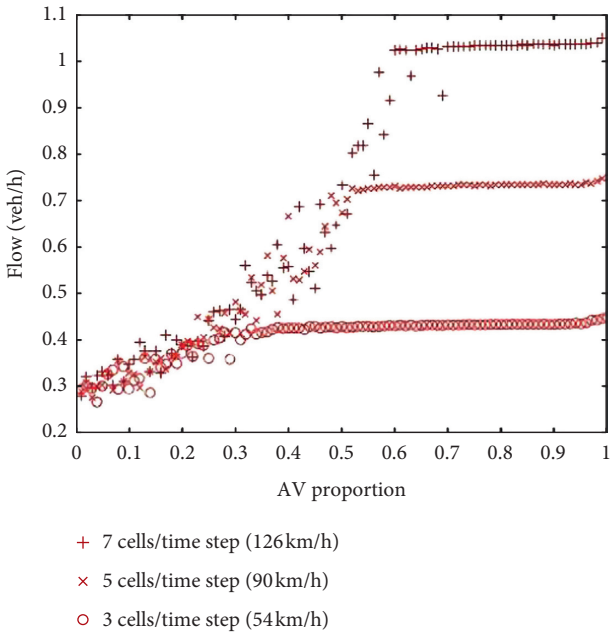


Figure 11: Flow-AV proportion relationship under the three speed limits at density = 0.3.

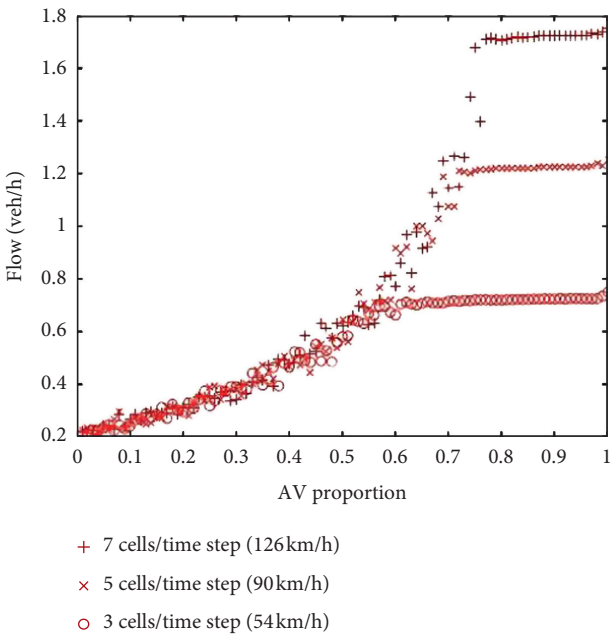


Figure 12: Flow-AV proportion relationship under the three speed limits at density = 0.5.

Focusing on the AV proportions of 0–0.2, the flow of the whole lane was the largest at density $\rho = 0.3$ and relatively small at $\rho = 0.1$ or 0.5. The results are consistent with the three-phase traffic flow theory, indicating that density = 0.5 is closer to the critical density. When the AV proportion fell in 0.6–0.8, the peak flow of the whole lane increased with the density.

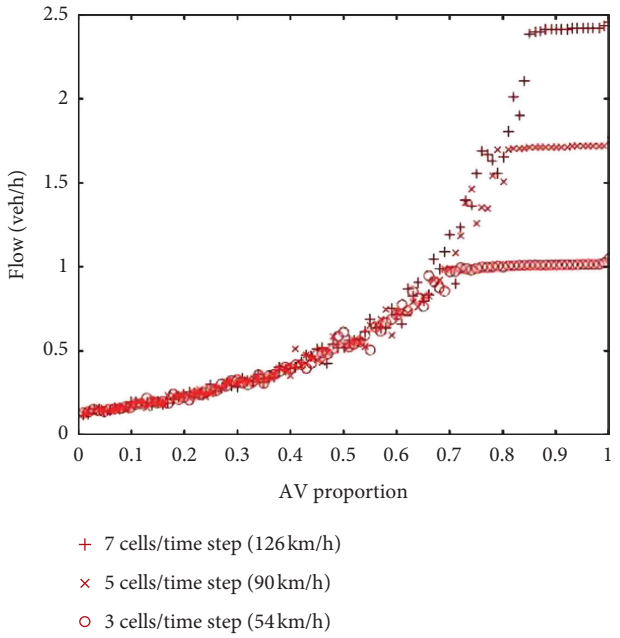


Figure 13: Flow-AV proportion relationship under the three speed limits at density = 0.7.

3.4. Influence of Speed Limits on Mixed Traffic Flow. To characterize the mixed traffic flow, different speed limits were introduced to the simulation: 3 cells/time step (54 km/h), 5 cells/time step (90 km/h), and 7 cells/time step (126 km/h). The flow-AV proportion relationships under the three speed limits at different densities are displayed in Figures 11–13, respectively.

As shown in Figures 11–13, regardless of the density, the mixed traffic flow increased with the speed limit. The peak flow under the speed limit of 7 cells/time step was 2.5 times that under 3 cells/time step. When the AV proportion was low, whichever the speed limit, most points in the diagram overlapped each other. This is because MVs are uncertain in driving, failing to maintain high-speed following. Their uncertainties affect the speed of AVs, dragging down the speed of the mixed traffic flow. However, when the AV proportion was high, the speed limit had a serious impact on the simulation results. In this case, the promoting effect of AVs cannot be ignored, while the limiting effect of MVs dwindles.

4. Conclusions

Based on the DHD, this paper designs an improved CA model, in which MVs update their positions by the rules of the NaSch model, while AVs update their positions by the dynamic headway rules. The designed model was applied to MATLAB simulation. The results show that the AV proportion has a significant impact on the mixed traffic flow; when the proportion reached 0.6, the flow of the whole lane was twice that of the MV traffic flow. At a low density, the AV proportion has an obvious influence on mixed traffic flow. At a high density, the mixed traffic flow changed very little, as the AV proportion increased from 0 to 5. However, when the AV proportion reached 0.8, the flow of the whole lane became three times that at the proportion of 0.6. In addition, speed limit also has a major impact on the mixed traffic flow. Regardless of the density, the peak flow under the speed limit of 7 cells/time step (126 km/h) was 2.5 times that under 3 cells/time step (54 km/h).

The proposed model is a single-lane model, in which the motion rules of MVs and AVs are both grounded on the CA. Our model can be extended into a multilane traffic flow model. The application prospect of our model is very broad.

Data Availability

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

References

[1] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," *Journal de Physique I*, vol. 2, no. 12, pp. 2221–2229, 1992.

[2] R. Barlovic, L. Santen, A. Schadschneider, and M. Schreckenberg, "Metastable states in cellular automata for traffic flow," *The European Physical Journal B*, vol. 5, no. 3, pp. 793–800, 1998.

[3] V. Kanagaraj and M. Treiber, "Self-driven particle model for mixed traffic and other disordered flows," *Physica A: Statistical Mechanics and its Applications*, vol. 509, pp. 1–11, 2018.

[4] N. Davoodi, A. R. Soheili, and S. M. Hashemi, "A macromodel for traffic flow with consideration of driver's reaction time and distance," *Nonlinear Dynamics*, vol. 83, no. 3, pp. 1621–1628, 2016.

[5] A. I. Delis, I. K. Nikolos, and M. Papageorgiou, "A macroscopic multi-lane traffic flow model for ACC/CACC traffic dynamics," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 20, pp. 178–192, 2018.

[6] W. Meng, C. Mao, J. Zhang, J. Wen, and D. Wu, "A fast recognition algorithm of online social network images based on deep learning," *Traitement du Signal*, vol. 36, no. 6, pp. 575–580, 2019.

[7] U. J. Reddy, B. R. V. R. Reddy, and B. E. Reddy, "Recognition of lung cancer using machine learning mechanisms with fuzzy neural networks," *Traitement du Signal*, vol. 36, no. 1, pp. 87–91, 2019.

[8] A. Talebpour and H. S. Mahmassani, "Influence of connected and autonomous vehicles on traffic flow stability and throughput," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 143–163, 2016.

[9] Y.-M. Yuan, R. Jiang, M.-B. Hu, Q.-S. Wu, and R. Wang, "Traffic flow characteristics in a mixed traffic system consisting of ACC vehicles and manual vehicles: a hybrid modelling approach," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 12, pp. 2483–2491, 2009.

[10] M. W. Levin and S. D. Boyles, "A multiclass cell transmission model for shared human and autonomous vehicle roads," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 103–116, 2016.

[11] A. Bose and P. Ioannou, "Mixed manual/semi-automated traffic: a macroscopic analysis," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 6, pp. 439–462, 2003.

[12] L. C. Davis, "Improving traffic flow at a 2-to-1 lane reduction with wirelessly connected, adaptive cruise control vehicles," *Physica A: Statistical Mechanics and its Applications*, vol. 451, pp. 320–332, 2016.

[13] S.-W. Kim and W. Liu, "Cooperative autonomous driving: a mirror neuron inspired intention awareness and cooperative perception approach," *IEEE Intelligent Transportation Systems Magazine*, vol. 8, no. 3, pp. 23–32, 2016.

[14] A. Sharma, Y. Ali, M. Saifuzzaman, Z. Zheng, and M. M. Haque, "Human factors in modelling mixed traffic of traditional, connected, and automated vehicles," in *Proceedings of the International Conference on Applied Human Factors and Ergonomics*, pp. 262–273, Los Angeles, CA, USA, July 2017.

[15] H. van Lint, W. Schakel, G. Tamminga, P. Knoppers, and A. Verbraeck, "Getting the human factor into traffic flow models: new open-source design to simulate next generation of traffic operations," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2561, no. 1, pp. 25–33, 2016.

[16] D. Ngoduy, "Application of gas-kinetic theory to modelling mixed traffic of manual and ACC vehicles," *Transportmetrica*, vol. 8, no. 1, pp. 43–60, 2012.

[17] R. Jiang and Q.-S. Wu, "The adaptive cruise control vehicles in the cellular automata model," *Physics Letters A*, vol. 359, no. 2, pp. 99–102, 2006.

[18] X. P. Qiu, L. N. Ma, X. X. Zhou, and D. Yang, "Research on manual-automatic vehicle mixed traffic flow based on safety distance," *Journal of Transportation Systems Engineering and Information*, vol. 16, no. 4, pp. 101–108, 2016.

[19] Y. Y. Qin, H. Wang, W. Wang, and Q. Wan, "Stability analysis and fundamental diagram of heterogeneous traffic flow mixed with cooperative adaptive cruise control vehicles," *Acta Physica Sinica*, vol. 66, no. 9, Article ID 094502, 2017.

[20] L. Liu, C. Li, Y. Li, S. Peeta, and L. Lin, "Car-following behavior of connected vehicles in a mixed traffic flow: modeling and stability analysis," in *Proceedings of the 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 1085–1088, Tianjin, China, July 2018.

[21] R. Mohajerpoor and M. Ramezani, "Mixed flow of autonomous and human-driven vehicles: analytical headway modeling and optimal lane management," *Transportation Research Part C: Emerging Technologies*, vol. 109, pp. 194–210, 2019.

[22] A. Ghiasi, O. Hussain, Z. Qian, and X. Li, "A mixed traffic capacity analysis and lane management model for connected automated vehicles: a Markov chain method," *Transportation Research Part B: Methodological*, vol. 106, pp. 266–292, 2017.

[23] Y. Liu, J. Guo, J. Taplin, and Y. Wang, "Characteristic analysis of mixed traffic flow of regular and autonomous vehicles using cellular automata," *Journal of Advanced Transportation*, vol. 2017, Article ID 8142074, 2017.

[24] T. Wietholt and J. Harding, "Influence of dynamic traffic control systems and autonomous driving on motorway traffic flow," *Transportation Research Procedia*, vol. 15, pp. 176–186, 2016.

[25] S. Gong and L. Du, "Cooperative platoon control for a mixed traffic flow including human drive vehicles and connected and autonomous vehicles," *Transportation Research Part B: Methodological*, vol. 116, pp. 25–61, 2018.

[26] N. X. Zhang, H. B. Zhu, H. Lin, and M. Y. Huang, "One-dimensional cellular automaton traffic flow model considering dynamic vehicle spacing," *Acta Physica Sinica*, vol. 64, no. 2, pp. 303–309, 2015.

[27] X. L. Ma, Y. H. Hu, and Q. Y. Yang, "Modeling and simulation of mixed traffic flow in two-lane general vehicle-automatic truck," *Transportation Systems Engineering and Information*, vol. 18, no. 6, pp. 72–80, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9653058, 1 page
https://doi.org/10.1155/2020/9653058

Corrigendum

Corrigendum to "Mathematical Analysis of Heat Transfer in Peristaltic Transport through a Rough Nonuniform Inclined Channel"

R. Shukla,[1] **A. Medhavi**,[1] **S. S. Bhatt**,[2] **and R. Kumar**[3]

[1]*Department of Mechanical Engineering, Kamla Nehru Institute of Technology, Sultanpur, India*
[2]*Independent Researcher, India*
[3]*B. N. College of Engineering and Technology, Lucknow, India*

Correspondence should be addressed to S. S. Bhatt; shankarbhatt56@gmail.com

Received 10 September 2020; Accepted 10 September 2020; Published 30 September 2020

In the article titled "Mathematical Analysis of Heat Transfer in Peristaltic Transport through a Rough Nonuniform Inclined Channel" [1], an incorrect affiliation was given for S. S. Bhatt. The corrected author list is shown above.

References

[1] R. Shukla, A. Medhavi, S. S. Bhatt, and R. Kumar, "Mathematical analysis of heat transfer in peristaltic transport through a rough nonuniform inclined channel," *Mathematical Problems in Engineering*, vol. 2020, Article ID 4715082, 16 pages, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 8305475, 10 pages
https://doi.org/10.1155/2020/8305475

Research Article

Destination Estimation for Bus Passengers Based on Data Fusion

Wusheng Liu,[1] Qian Tan ,[2] and Lisheng Liu[3]

[1]*Engineering Research Center of Catastrophic Prophylaxis and Treatment of Road & Traffic Safety of Ministry of Education, Changsha University of Science & Technology, Changsha 410114, China*
[2]*School of Traffic & Transportation Engineering, Central South University, Changsha 410075, China*
[3]*School of Urban Construction, Wuchang Institute of Technology, Wuhan 430223, China*

Correspondence should be addressed to Qian Tan; tq.helen@csu.edu.cn

Received 6 July 2020; Revised 21 August 2020; Accepted 6 September 2020; Published 16 September 2020

Academic Editor: Hussein Abulkasim

The planning and operation of urban buses depend heavily on the time-varying origin-destination (OD) matrix for bus passengers. In most cities, however, only boarding information is recorded, while the alighting information is not available. This paper proposes a novel method to predict the destination of a single bus passenger based on bus smartcard data, metro smartcard data, and global positioning system (GPS) bus data. First, the attractiveness of each bus stop in a bus line was evaluated, considering the attractiveness of nearby metro stations. Then, the exploration and preferential return (EPR) model was employed to estimate the probability of a bus stop to be the alighting stop, i.e., the destination, of a passenger. The estimation result was obtained through a simulation based on the Monte Carlo (MC) algorithm. The effectiveness of our method was proved through a case study on the bus network in Shenzhen, China.

1. Introduction

The origin-destination (OD) estimation of bus passengers is essential to the network planning and operation of buses. Traditionally, the ODs of bus passengers were estimated by questionnaire surveys, which are small in sample size and low in precision. In recent years, the OD estimation of bus passengers has progressed rapidly, owing to the proliferation and use of geoinformation systems and smartcard techniques. GPS and bus IC card data are widely used. A bus IC card provides a large amount of data for the travel characteristic analysis of bus passengers, which has large amount of data, low cost, and high accuracy. However, it can only provide the relevant information of passengers' boarding stops. In order to obtain the passenger travel OD, it is necessary to calculate the alighting stops.

Many scholars have investigated the OD estimation problem. The early studies mainly derived the OD matrix of public transit passengers from the passenger flow at each station. For instance, Ben-Akiva et al. [1, 2] proposed an OD derivation method based on the survey data on ODs. After analyzing the passenger flow at each bus stop, Navick et al. [3, 4] estimated the travel patterns of bus passengers and constructed the OD matrix of these passengers. Tsygalnitsky [5] predicted the destinations of public transit passengers under the assumption that all the passengers boarding the same type of public transit vehicle at the same station have the same probability to alight at another station.

Moreover, Li and Cassidy [6] designed an algorithm that does not need a seed matrix to estimate the ODs of public transit passengers. Based on boarding and alighting counts at each stop along the route, the designed algorithm deduces an OD matrix for the entire trip and forecasts the probabilities for passengers to board and alight at each stop along the route. These probabilities tend to remain fixed throughout the trip. Compared with Tsygalnitsky's prediction technique, Li and Cassidy's algorithm is highly suitable for general use. Using boarding and alighting counts, Li [7] developed an efficient statistical inference method for a closed-form OD matrix of a travel route: the Markov chain model was adopted to capture the correlations between matrix elements and reduce the number of unknown parameters; then, the unknown parameters of the Markov chain model were inferred by Bayesian analysis.

In practice, the OD estimation is often realized through iteration proportional fitting, in which the OD matrix is adjusted continuously based on survey data, and the passenger flow of each bus stop is obtained by adding up the row and column vectors of the OD matrix [8, 9]. This approach is efficient and easy to implement but costly in terms of labor and money. In addition, a very few cities have adopted an automatic passenger counter (APC) system to collect the information of bus passenger flows in an automatic manner.

Recently, new methods have emerged to estimate the destinations of bus passengers based on global positioning system (GPS) data and smartcard data. Using the APC data, Barry et al. [10] estimated the destinations of bus passengers in New York City based on the trip chain theory. Zhao [11] derived the destinations of passengers that transfer between metro lines or between metro and bus. Seaborn et al. [12] explored the time span for transfer between public transit modes and relied on the time span to identify multimode transfers. Hofmann and O'Mahony [13] determined the transfer stations according to the time difference between two card swipes.

In addition, based on GPS data, Giannotti et al. [14] researched the detailed trajectory of vehicles and the frequent mode of urban residents' travel, the prediction method of traffic intensive areas, and the description method of traffic congestion; Jiang et al. [15] found that human travel distance obeyed power-law distribution using GPS data. Based on the travel survey data, Garske et al. [16] studied the residents' travel in two cities with different economic development levels in China. Kölbl and Helbing [17] found that under different travel modes, people's travel distance follows a general distribution law. In [18], the daily travel logs of 230 volunteers in Frauenfeld, Switzerland, were analyzed. The author found that the travel distance of the group obeyed a power-law distribution with exponential truncation, which was very close to the empirical research results based on mobile phone data, and most individual travel distance does not conform to power-law distribution. Cui [19] deduced the boarding stations of passengers from the data collected by the automatic fare collection system and the automatic vehicle location system [20]. Farzin [21] analyzed the boarding stations of passengers in São Paulo, Brazil, referring to information of integrated circuit (IC) bus cards, GPS data, and bus stop data. Xu et al. [22] estimated the destinations of public transit passengers, in the light of travel distance distribution and bus stop features. Xu et al. [23] clustered smart card data before estimating the alighting stops of passengers.

The current research focuses on two aspects: Firstly, the main method is to analyze the distribution characteristics of passenger travel distance to estimate alighting stop of a single bus route. Secondly, OD was obtained with using the difference of swipe card time in passenger travel to analyze the transfer between different modes of transportation. In terms of individual travel characteristics, it is found that the individual average travel distance is significantly different, the individual visit frequency to the location also follows the power-law distribution, and the individuals with different average travel distance have high similarity in spatial motion location distribution [24].

Therefore, it is reasonable to use the distribution characteristics of travel distance to study the single bus stop without transfer with subway, but this study uses the mobile phone signaling data; the traffic mode includes walking and nonmotor vehicle. The travel distance is a continuous variable, which is from the origination to the destination, not the next stop. When passengers travel by bus, the travel distance is a discrete variable, and the next station may not be the destination, so the transfer must be considered. In addition, according to the habits of human travel activities, travel behavior tends to choose to return to places that have been visited more times in history, such as home and office [25]. For the passengers with a large number of historical travel records, it is more accurate to analyze the historical travel law to predict the next stop.

In general, the destinations of bus passengers on a single bus line are estimated from the trip chain of boarding stations and the historical smartcard data, or from the land use attributes of all bus stops. In this paper, the historical travel features of a single passenger and the attractiveness of the nearby metro stations of a bus stop are analyzed in detail; the alighting stop, i.e., the destination, of that passenger was estimated by using the exploration and preferential return (EPR) model [26] and the Monte Carlo (MC) algorithm; and we also try to estimate the destinations of bus passengers in the future under connected and autonomous vehicle environment [27–31].

2. Data Sources

The research data can be divided into five datasets: bus line data, bus smartcard data, bus GPS data, metro smartcard data, and road network data.

2.1. Bus Line Data. Bus line data contain the information of each bus stop on a bus line, including but not limited to coordinates, name, and bus line number. In total, the data were collected from 1,516 bus lines in Shenzhen.

2.2. Bus Smartcard Data. The bus smartcard data refer to the transaction information captured by the smart card fare collector, once a passenger swipes his/her smartcard upon boarding a bus. The data cover passenger identity (ID), boarding time, bus ID, and bus line number. Each bus smartcard carries a unique passenger ID that can be identified easily. In this research, the bus smartcard data were provided by ShenZhenTong, the largest public transit service provider in Shenzhen, and collected in the 21 days from October 11th to 31st, 2014. These data are for the three weeks after the National Day holiday. The three weeks include 18 working days and 4 nonworking days. The average number of card swipes per week is the largest in the whole year in Figure 1, which is representative.

2.3. Bus GPS Data. Each bus has a GPS tracker that records the bus position in real time. The bus ID and bus line

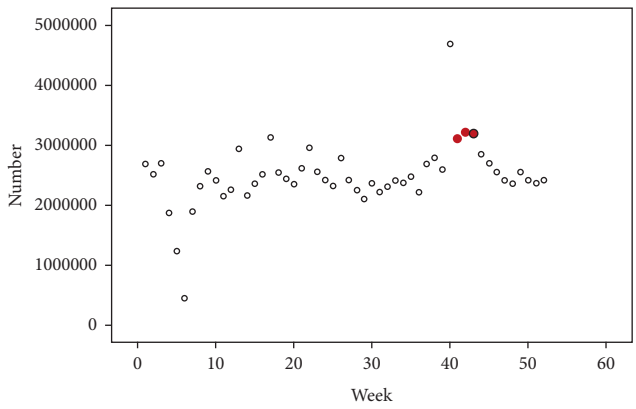


Figure 1: The average number of card swipes per week.

number recorded by the GPS tracker are unique, allowing us to match bus GPS data with bus smartcard number.

2.4. Metro Smartcard Data. Metro smartcard data include the name of boarding and alighting metro stations, passenger ID, and boarding and alighting time. In this research, the metro smartcard data were also provided by ShenZhenTong and collected in the same period.

2.5. Road Network Data. The road network data of Shenzhen (N: 22°46′–22°83′; E: 113°78′–114°59′) was simplified into 21,115 sections.

3. Methodology

Based on the above data, this paper aims to estimate the destination for every single passenger boarding a bus line. The estimation process can be broken down into three tasks. The first task is to determine the arrival time and location of the bus at each stop of the bus line by matching bus line data with bus GPS data. The second task is to identify the boarding station of the passenger by matching GPS data with bus smartcard data. The third and final task is to estimate the destination of the passenger by using the EPR model, with metro smartcard data as the basis for exploration.

3.1. Data Matching. The bus smartcard only records the boarding time and vehicle of the passenger, but not the boarding location (station). To determine the boarding station, it is necessary to match bus smartcard data with bus GPS data, i.e., complete the first two tasks mentioned at the beginning of this section:

Task 1: matching bus line data with bus GPS data

In general, a bus line is composed of an upline and a downline; that is, a bus moves in one of the two opposite directions at a time. Considering the inevitable errors in bus GPS data and the close proximity between stops with similar names, it is impossible to match bus line data with bus GPS data based on the distance indicated by position feedbacks from the GPS trackers on the bus and the GPS device at each stop. Instead, the moving direction of the bus should be identified in the light of its trajectory from the GPS, and then the locations be matched based on the direction. The matching process is illustrated in Figure 2, where L_u and L_d are the upline and downline, respectively.

Let S and S' be the departure stations of L_u and L_d, respectively. Ten consecutive tracking points were chosen from the GPS data of a bus M1 that operates along line L. Then, the distance from each tracking point to S or S' was calculated. If the distance of these points increases from station S (S'), then the bus is on the up (down) line, that is, $L_{M1} = L_u$ (L_d). Next, the bus GPS data were matched with stops along the right direction. The location matching was deemed successful, when the GPS location fell within 100 m of the boarding stop S_0. The matching time was taken as the arrival time at the stop. In this way, the arrival time of bus M_1 at all stops along line L can be obtained as T_i, $i = 1,2,3, \ldots, n$.

Task 2: matching bus smartcard data with bus GPS data

Suppose a passenger P_1 boards M_1 at station S_i and swipes his/her smartcard at time T_p. Then, T_p was taken as the boarding time of passenger P_1. Comparing T_p with T_i, the boarding stop S_0 (as shown in Figure 3, i.e., origin, of passenger P_1 can be identified as the station S_i with the lowest time difference from T_p. Any of the stops following the origin could be the destination of the passenger.

3.2. Destination Estimation of a Single Bus Passenger. The destination of a bus passenger depends on various factors, namely, travel distance and attractiveness of each stop. The latter, referring to the possibility that a passenger alights at the stop, encompasses the attractiveness of the stop itself and extra attractiveness. Here, extra attractiveness is measured by the attractiveness of metro stations near the stop because buses often serve as the feeders for metro in the large and dense transit network in Shenzhen. The two mechanisms of the EPR model and the MC algorithm are introduced to estimate the destination of a single bus passenger.

3.2.1. Explore Mechanism. The explore mechanism applies to the scenario where there is no card swiping record at the previous boarding stops of the passenger. For such a passenger, the possibility that he/she alights at a stop is the probability P_{ij} for the stop to be the destination. According to the definition and composition of stop attractiveness, the probability P_{ij} covers two items, namely, the probability F_{ij} arising from the attractiveness of the stop itself and the probability D_{ij} stemming for the attractiveness of nearby metro stations.

The travel distance can be approximated by the number of stops passed by the bus. Without considering probability D_{ij}, the number of stops passed by a bus follows the Poisson distribution. In this case, the probability F_{ij} can be expressed as

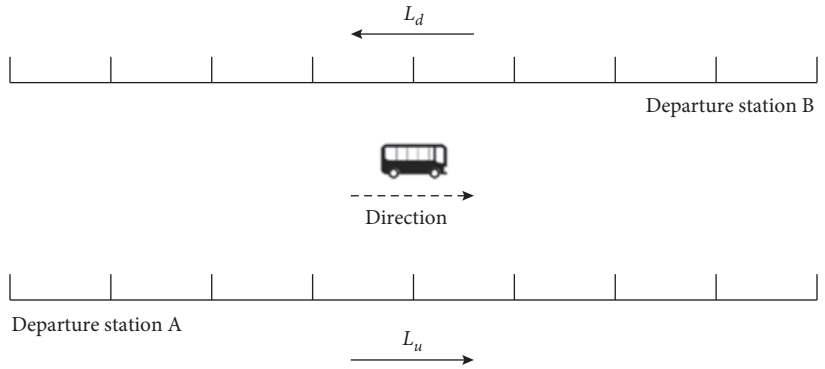


Figure 2: The matching between bus line data and bus GPS data.

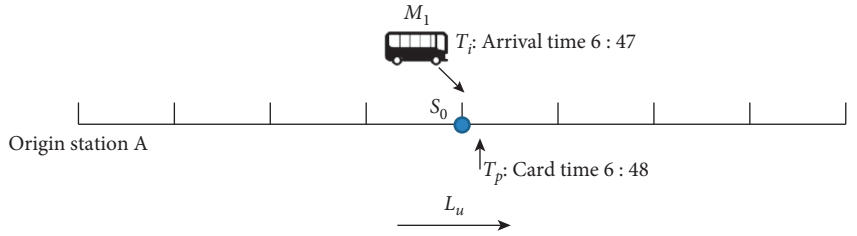


Figure 3: The matching between bus smartcard data and bus GPS data.

$$F_{ij} = \begin{cases} \dfrac{e^{-\lambda}\lambda^{(j-i)}}{(j-i)!}, & i < j, \\ 0, & i \geq j, \end{cases} \quad i, j = 1, 2, \ldots \tag{1}$$

where i and j are the serial numbers of the boarding and alighting stops, respectively (the serial numbers were assigned from the departure station of the bus line); n is the number of stops on the bus line; and λ is the mean number of stops in a bus line (λ was set to 10 for Shenzhen [32]). If the number of remaining stops after the boarding stop is fewer than λ, then $\lambda = n - i$.

The Poisson distribution can be normalized as

$$F_{ij}^{*} = \frac{\left(e^{-\lambda}\lambda^{(j-i)}\right)/(j-i)!}{\sum_{i=i+1}^{m}\left(e^{-\lambda}\lambda^{(j-i)}\right)/(j-i)!}. \tag{2}$$

However, another determinant factor of bus passengers' destination is the nature of land use. If there are shopping malls and entertainment sites nearby, the attraction is greater; especially in the stops near the transportation hub, the number of people boarding and alighting the bus is the largest. Due to the round-trip characteristics of residents' public transportation, the volume of generation and attraction of stops are basically in a balanced state; that is, the more the people get on the bus, the more the people get off the bus. The attraction intensity of each stop is calculated by counting the total number of passengers at each stop from the judgment of the previous stop.

The metro records near the boarding stop were statistically analyzed. The boarding stop was considered near a metro station, if their distance is smaller than 1,000 m. This distance may lead to a time difference in the metro records of the transfer passenger. Here, the time difference is set to 30 min; that is, the metro records generated 30 min after the passenger swiped his/her smartcard on the target bus are counted. The records at the nearby metro stations were used to measure the attractiveness of the bus stop. If there is no metro station nearby, the attractiveness of the bus stop was set to zero.

Then, the probability D_{ij} of the bus stop j can be computed as

$$D_{ij} = \frac{d_j^{*}}{\sum_{j=i+1}^{m} d_j^{*}}, \tag{3}$$

where d_j^{*} is the record at the nearby metro stations generated 30 min after the passenger swiped his/her smartcard on the target bus.

According to the literature [33], based on more than 230000 pieces of data analysis, research on the transfer time between conventional bus and subway, and analyzing the transfer data with the interval of 5 minutes, it can be found that most of the transfer behavior takes less than 20 minutes and only 2–4% of the total transfer amount exceeds 20 minutes (Figure 4). In order to ensure the integrity of transfer sample identification, 30 is selected as the transfer time threshold.

According to the agglomeration effect of public transit stations, the attractiveness of a metro station decreases with

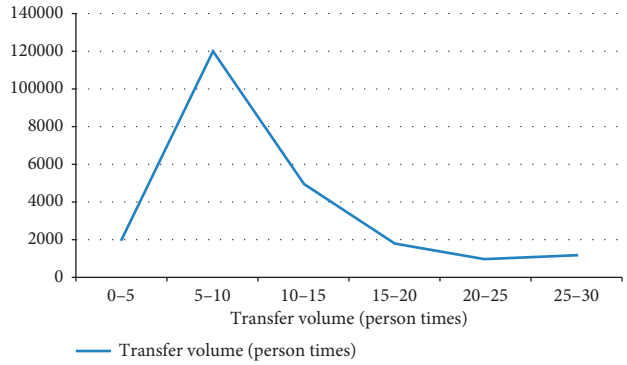


Figure 4: Transfer time distribution.

its distance to the bus stop. Hence, the D_{ij} value obeys the exponential distribution:

$$E_j = E_0 e^{-s}, \tag{4}$$

where E_j is the intensity of the agglomeration effect; $E_0 E_j = 1 E_0$ is the peak agglomeration effect; and s is the distance from a metro station to the bus stop.

Then, the d_j^* value can be obtained by

$$d_j^* = E_j \times d_j. \tag{5}$$

On this basis, the probability P_{ij} a stop to be the destination for a single bus passenger can be calculated by

$$P_{ij} = \frac{F_{ij}^* + \alpha D_{ij}}{\sum \alpha D_{ij} + \sum F_{ij}^*}, \tag{6}$$

where α is the coefficient of the attractiveness of nearby metro stations $(0 < \alpha \leq 1)$, i.e., the $0 < \alpha \leq 1$ weight of D_{ij}, and the weight of F_{ij}^* is 1. The value of α is positively correlated with the proportion of passengers taking metro instead of bus.

3.2.2. Preferential Return Mechanism. Based on the research results of the literature [25], it is found that in terms of residents' travel destination, when the historical data of residents' travel increases with increase of time, the number of residents visiting new places follows $s(t) = t^{-\mu}$, μ~0.6; the frequency of residents' visit to the place (k) is subject to $f_k = k^{-\xi}$, ξ~1.2; and the authors pointed out that the accuracy of the CTRW model is not good. The main idea of the exploratory regression model is that individuals return to the previously visited places with the probability $(\rho s^{-\gamma})$ of exploring the location, and the probability $(1^{-}\rho s^{-\gamma})$ of visiting a place is directly proportional to the probability of individuals being found in the location, as shown in Figure 5.

Next, the preferential return mechanism was employed to predict the destination under the scenario that there is no card swiping record on the bus line boarded by the passenger. In general, passengers prefer to alight at frequently visited places, such as home and workplace. Thus, the basic idea of the preferential return mechanism is that passengers tend to alight at stops with more historical card swiping records. In other words, the probability for a stop to be the destination is proportional to the historical record count of the passenger at that stop.

To eliminate the interference of stops with similar names, the smartcard records within 100 m were counted as the records of one stop, where 100 m is the return range. The stop with many historical records has a high probability of being returned, which is directly proportional to the number of historical records.

Based on the historical records of a passenger, the probability f_i of a stop to be the destination can be described as

$$f_i = \frac{m_i}{\sum_{i=1}^{n} m_i}, \tag{7}$$

where i is the serial number of stops following the boarding stop; m_i is the number of historical records of stop i; and n is the number of stops with a probability of being returned.

3.2.3. MC Algorithm. Finally, the destination of a single bus passenger was predicted by using the MC algorithm. Based on probability theory, the MC algorithm relies on a random probability model to approximate the probability through simulation and statistical testing on random variables. As shown in Figure 6, the MC algorithm is implemented in the following steps [20]:

Step 1: construct and describe the probability process as formulae (2) and (6)

Step 2: determine the sample size and samples from the probability distribution, and set the number of simulations to 1,000 for each passenger

Step 3: confirm the estimation, i.e., the alighting stop

The estimated destination is the stop with the largest number of occurrences in the 1,000 simulations. If a passenger boards at stop S_0 of bus line L, then the stops after S_0 are numbered as $S_1, S_2, \ldots, S_n$ in turn. The number of occurrences of each stop in the 1,000 simulations is denoted as $x_1, x_2, \ldots, x_n$, respectively, and the estimated destination as S_i, where $x_i = \max(x_1, x_2, \ldots, x_n)$.

4. Case Study

4.1. Estimation of the Boarding Station. The proposed method was applied to predict the destination of every single bus passenger in the 1,516 bus lines across Shenzhen. The smartcard bus data, bus GPS data, and smartcard metro data were collected by ShenZhenTong in 21 days from October 11th to 31st, 2014, including 14,109 trip records for 23 passengers, as well as the trajectories and stop locations of the 1,516 bus lines.

During the 21 days from October 11 to 31, 2014, the total number of card swipes was 268623 (excluding subway passengers). There were 113625 passengers. The method of random sampling was used to collect the sample, and the sample size is calculated as shown in Table 1.

Sample size n is shown below:

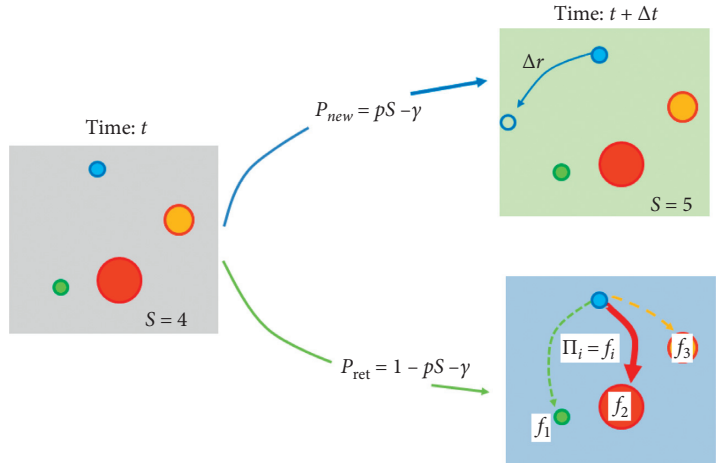


Figure 5: Preferential return mechanism [25].

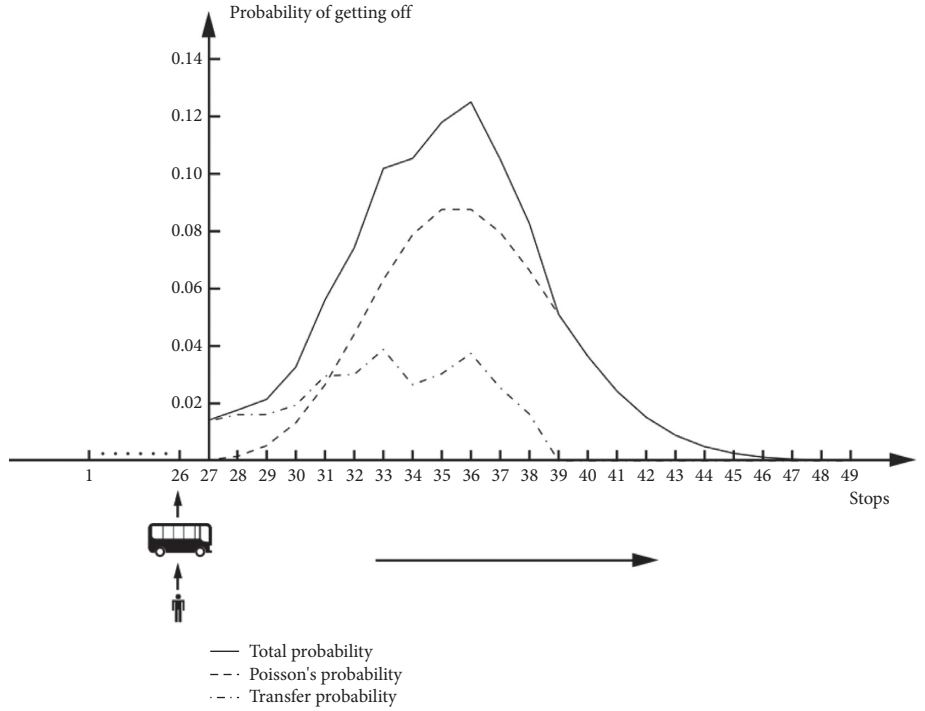


Figure 6: The MC-based estimation of destination.

Table 1: Sample-related parameters.

Sample characteristics	Standard deviation of total trips σ	Expected error E	Confidence interval (%)
Average number of trips per person	31.5	5.0	95

$$n = \frac{\left(z_{(\alpha/2)}\right)^2 \sigma^2}{E^2} = \frac{1.96^2 \times 31.5^2}{5^2} = 152. \quad (8)$$

As mentioned before, the road network data of Shenzhen (N: 22°46′–22°83′; E: 113°78′–114°59′) was simplified into 21,115 sections.

For simplicity, the downtown of Shenzhen was divided into 74 × 162 zones, each of which is 555 m by latitude and 515 m by longitude. Then, the smartcard records, GPS bus data, and stop locations were fused to obtain spatial distribution of the boarding stops for bus passengers in Shenzhen during the five days from October 27th to 31st,

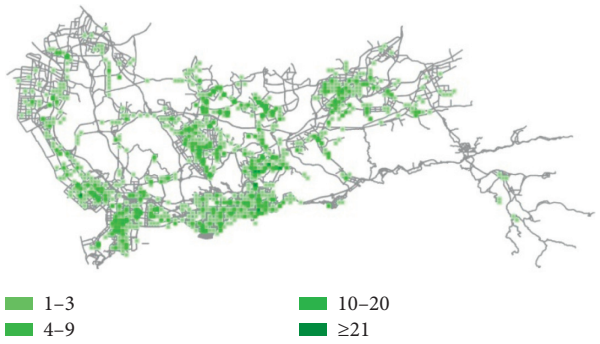


Figure 7: Spatial distribution of boarding stops.

2014. As shown in Figure 7, most passengers board buses at the center of the city; that is, the bus stops at the central area of Shenzhen have relatively high attractiveness.

To differentiate the estimated results of exploration mechanism from those of the preferential return mechanism, the boarding stops of the 293 bus passengers with and without historical records are displayed in Figures 8 and 9, respectively.

4.2. Destination Estimation

4.2.1. Determination of α Value. The value of α, that is, the weight of D_{ij} or the coefficient of the attractiveness of nearby metro stations, was set to 0.5 and 0.7 during the destination estimation. The results show that the α value has a limit effect on the estimation. Since the ratio of metro trips to bus trips in Shenzhen is 3 : 7, the value of α was set to 0.7.

4.2.2. Destination Estimation. Our method was adopted to estimate the alighting stop, i.e., destination, of every single bus passenger from October 27th to 31st, 2015. The estimated destinations are displayed as the heat map in Figure 10. It can be seen that the destinations concentrated in the central area of the city, revealing a correlation between boarding and alighting locations.

Next, the exploration mechanism and the preferential return mechanism were separately adopted to estimate the destinations of each of the 293 passengers without historical records on the bus line, with $\alpha = 0.7$. The estimated results of the two mechanisms are presented in Figures 11 and 12, respectively.

4.3. OD Distributions in Different Periods. Based on the above estimation, the distribution of origins (boarding stops) and destinations (alighting stops) was illustrated for different periods of the day (Figures 13–18).

As shown in Figures 13–18, the origin distribution in the morning peak is consistent with the destination distribution in the evening peak, and both origins and destinations are clustered in residential areas. It is learned that the bus trips of passengers in the morning and evening peaks leave from and return to their homes, respectively.

Figure 8: Spatial distribution of boarding stops for bus passengers with historical records.

Figure 9: Spatial distribution of boarding stops for bus passengers without historical records.

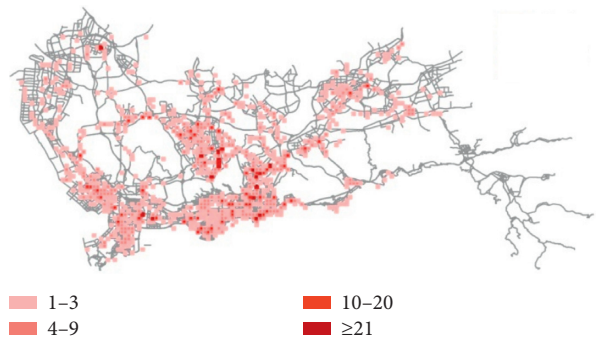


Figure 10: Heat map of destinations.

Figure 11: Spatial distribution of destinations with exploration mechanism.

Figure 12: Spatial distribution of destinations with preferential return mechanism.

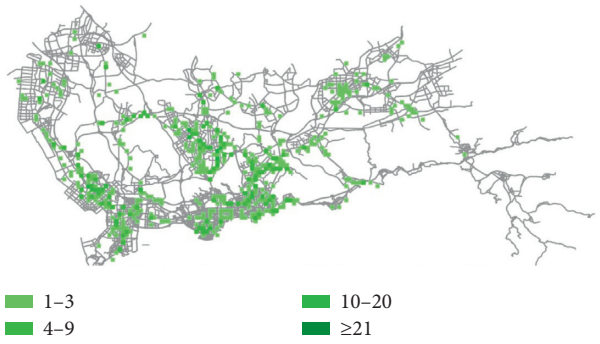


Figure 13: Origin distribution in the morning peak.

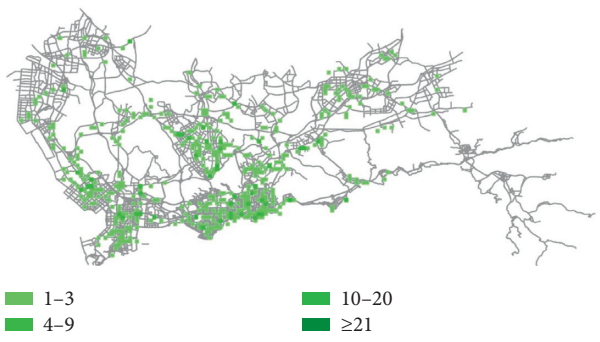


Figure 14: Origin distribution in the evening peak.

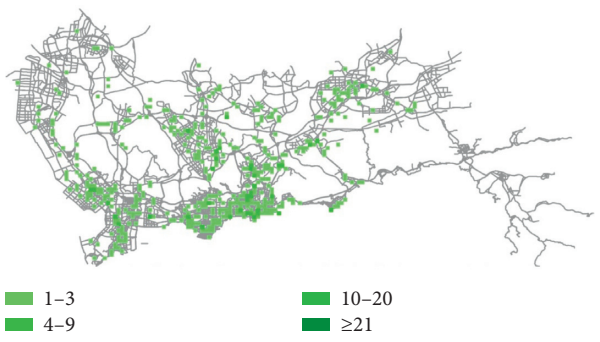


Figure 15: Origin distribution in off-peak hours.

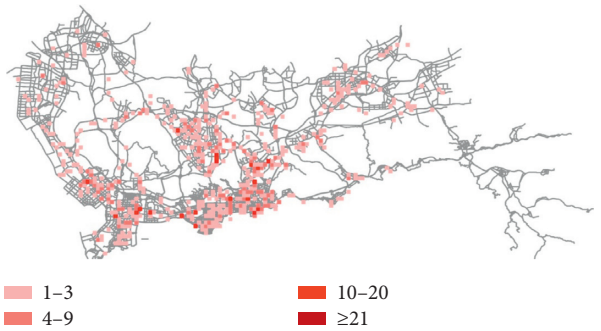


Figure 16: Destination distribution in the morning peak.

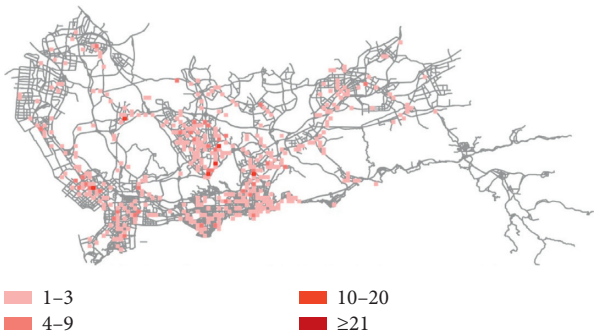


Figure 17: Destination distribution in the evening peak.

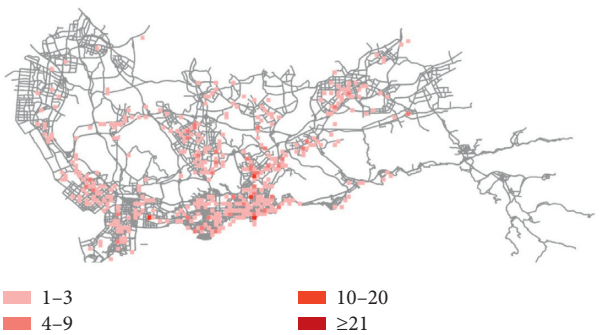


Figure 18: Destination distribution in off-peak hours.

Besides, the destinations in the morning peak mostly fell in commercial and school areas, indicating that most passengers go to work or school in the morning. The origins in the evening peak was slightly scattered, yet mainly from commercial and school areas, which are the main destinations for business and school travels.

Finally, the origins and destinations were relatively decentralized in the off-peak hours, suggesting that a certain portion of the travels are nondomestic.

Through our data analysis, 3,023 destinations (82.8%) were derived by the preferential return mechanism from the data of 3,651 passengers, while 62 (17.2%) were derived by the exploration mechanism. It can be speculated that urban residents tend to return to places they have visited before.

With the growing number of trips, the residents are more likely to prefer the historical locations over new locations. This agrees with the rule of preferential return. Our estimation shows that the travels of bus passengers concentrated in the morning and evening peaks: going downtown in the morning peak and returning home in the evening peak. This is clearly in line with the situation of urban commuters.

5. Conclusions

This paper fully integrates bus line data, bus smartcard data, bus GPS data, and metro smartcard data, with road network data, and introduces the EPR model and MC algorithm to estimate the alighting stops of bus passengers in Shenzhen. Firstly, the boarding location and time of every single passenger were estimated based on the integrated data. Then, the alighting station of the passenger was predicted under the exploration mechanism and the preferential return mechanism, which is based on the features of travel activities. During the prediction, the metro smartcard data were innovatively employed to evaluate the extra attractiveness of each bus stop. Considering the features of historical trips and bus-metro transfer, the proposed method was found to effectively solve the destination estimation problem through a case study. The future research will further explore the multimode traffic and OD estimation of multitransfer trips; based on the results of this study, we can obtain the origination and the destination of passengers and estimate the travel OD combined with multimode traffic and multiple transfer, so as to obtain the relevant characteristics of passengers' travel, including travel time, travel OD, travel distance, and station passenger flow; the travel distance and station passenger flow can provide more accurate data support for urban bus dispatching and schedule more reasonable departure time and interval of peak and flat peak vehicles. In addition, the passenger flow and travel OD can provide the basis for public transportation network adjustment.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank Shenzhen Urban Transport Planning Center for providing part of our research data and also to other scholars and research teams for their scientific contributions. This project was supported by Open Fund of Engineering Research Center of Catastrophic Prophylaxis and Treatment of Road & Traffic Safety of Ministry of Education (Grant no. kfj180401, Changsha University of Science & Technology), Natural Science Foundation of Hunan Province, China (Grant no. 2019JJ40306), and the National Natural Science Foundation of China (Grant no. 61773077).

References

[1] M. E. Ben-Akiva, "Methods to combine different data sources and estimate origin-destination matrices," in *Proceedings of the 10th international symposium on transportation and traffic theory*, Elsevier, Cambridge, UK, pp. 459–481, July 1987.

[2] M. Ben-Akiva, P. Macke, and P. Hsu, "Alternative methods to estimate route-level trip tables and expand on-board surveys," *Transportation Research Record*, vol. 1037, p. 1, 1985.

[3] D. S. Navick and P. G. Furth, "Distance based model for estimation a bus route OD matrix," *Transportation Research Record*, vol. 1433, p. 16, 1994.

[4] H. L. Dou, H. D. Liu, and X. G. Yang, "OD matrix estimation method of public transportation flow based on passenger boarding and alighting," *Computer and Communications*, vol. 25, no. 135, p. 79, 2007.

[5] S. Tsyglnitsky, "Simplified methods for transportation planning," Massachusetts Institute of technology, Cambridge, UK, 1977.

[6] Y. Li and M. J. Cassidy, "A generalized and efficient algorithm for estimating transit route ODs from passenger counts," *Transportation Research Part B: Methodological*, vol. 41, no. 1, pp. 114–125, 2007.

[7] B. Li, "Markov models for Bayesian analysis about transit route origin-destination matrices," *Transportation Research Part B: Methodological*, vol. 43, no. 3, pp. 301–310, 2009.

[8] M. Bacharach, "Biproportional matrices and input-output change," Cambridge University Press, London, UK, 1970.

[9] B. Beirami and M. Mokhtarzade, "Spatial-spectral random patches network for classification of hyperspectral images," *Traitement du Signal*, vol. 36, no. 5, p. 399, 2019.

[10] J. J. Barry, R. Newhouser, A. Rahbee, and S. Sayeda, "Origin and destination estimation in New York City with automated fare system data," *Transportation Research Record*, vol. 1817, no. 1, pp. 183–187, 2002.

[11] J. Zhao, "The planning and analysis implications of automated data collection systems: rail transit OD matrix inference and path choice modeling examples," Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2004.

[12] C. Seaborn, J. Attanucci, and N. H. Wilson, "Using smart card fare payment data to analyze multi-modal public transport journeys in London," 2009.

[13] M. Hofmann and M. O'Mahony, "Transfer journey identification and analyses from electronic fare collection data," in *Proceedings of the 2005 IEEE Intelligent Transportation Systems*, pp. 34–39, Vienna, Austria, September 2005.

[14] F. Giannotti, M. Nanni, D. Pedreschi et al., "Unveiling the complexity of human mobility by querying and mining massive trajectory data," *The VLDB Journal*, vol. 20, no. 5, p. 695, 2011.

[15] B. Jiang, J. Yin, and S. Zhao, "Charactering the human mobility pattern in a large street network," *Physical Review E*, vol. 80, Article ID 021136, 2009.

[16] T. Garske, H. Yu, and Z. Peng, "Travel patterns in China," *PLoS One*, vol. 6, Article ID e16364, 2011.

[17] R. Kölbl and D. Helbing, "Energy laws in human travel behavior," *New Journal of Physics*, vol. 5, p. 1, 2003.

[18] X. Yan, X. Han, and B. Wang, "Diversity of individual mobility patterns and emergence of aggregated scaling laws," *Scientific Reports*, vol. 3, Article ID 2678, 2013.

[19] A. Cui, "Bus passenger origin-destination matrix estimation using automated data collection systems," Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2006.

[20] X. Song, S. Gao, and C. Chen, "A novel vehicle feature extraction algorithm based on wavelet moment," *Traitement du Signal*, vol. 35, no. 3-4, p. 223, 2018.

[21] J. M. Farzin, "Constructing an automated bus origin-destination matrix using farecard and global positioning system data in são Paulo, Brazil," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2072, no. 1, pp. 30–37, 2008.

[22] W. Y. Xu, C. Y. Deng, and B. Y. Liu, "Approach on public traffic passenger flow statistics based on IC data," *China Journal of Highway and Transport*, vol. 26, no. 5, pp. 158–163, 2013.

[23] J. M. Xu, W. H. Xiong, and F. You, "The method of signals-line public transportation OD base on GPS and intelligent card," *Control & Automation*, vol. 24, no. 8, pp. 221–223, 2008.

[24] D. Brockmann, l. Hufnagel, and T. Geisel, "The scaling laws of human travel," *Nature*, vol. 439, no. 7075, p. 462, 2006.

[25] P. Wang, Z. R. Huang, and G. Hang, "Transportation engineering in the big data era," *Journal of University of Electronic Science and Technology of China*, vol. 42, p. 806, 2013.

[26] C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modelling the scaling properties of human mobility," *Nature Physics*, vol. 6, no. 10, pp. 818–823, 2010.

[27] W. Wu, L. Huang, and R. Du, "Simultaneous optimization of vehicle arrival time and signal timings within a connected vehicle environment," *Sensors*, vol. 20, no. 1, p. 191, 2020.

[28] W. Wu, F. Zhang, W. Liu, and G. Lodewijks, "Modelling the traffic in a mixed network with autonomous-driving expressways and non-autonomous local streets," *Transportation Research Part E: Logistics and Transportation Review*, vol. 134, Article ID 101855, 2020.

[29] S. Veeramalla and V. Talari, "Estimation of neural sources from EEG measurements using sequential Monte Carlo method," *Ingénierie des Systèmes d Information*, vol. 24, no. 4, p. 411, 2019.

[30] S. Ourrad, Y. Houmadi, A. Ziadi, S. Mamoune, and A. Lousdad, "Probabilistic analysis for estimating the hydrogen desorption time from steel wire rods using Monte Carlo simulation," *Annales de Chimie-Science des Matériaux*, vol. 43, no. 4, p. 241, 2019.

[31] H. M. Reda and B. Abdelylah, "Numerical investigation and solar flux distribution analysis of parabolic trough solar collector by adding secondary reflector," *Instrumentation Mesure Métrologie*, vol. 18, no. 3, p. 275, 2019.

[32] Shenzhen Traffic and Transport Committee Integrated traffic and transportation system planning of Shenzhen (2013-2030). 2013.

[33] N. N. Guan and B. Wang, "Characteristics of transfer passenger flow between metro and conventional bus based on IC data in Chengdu city," *China Metros*, vol. 2018148 pages, 2018.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 3070539, 11 pages
https://doi.org/10.1155/2020/3070539

Research Article

Application of the Artificial Fish School Algorithm and Particle Filter Algorithm in the Industrial Process Control Particle Filtering Algorithm for Industrial Process Control

Zhaoxia Huang

Department of Mathematics and Statistics, Ankang University, Ankang 725000, China

Correspondence should be addressed to Zhaoxia Huang; 2007990011@aku.edu.cn

Received 6 July 2020; Revised 27 July 2020; Accepted 4 August 2020; Published 15 September 2020

Guest Editor: Hussein Abulkasim

The Industrial Internet of Things (IIoT) is of strategic importance in the new era of industrial big data, creating a brand-new industrial ecosystem. Considering the unknown parameters in the IIoT-based industrial process control systems, this paper combines the artificial fish swarm algorithm (AFSA) and the particle filtering (PF) algorithm into the AFSA-PF algorithm based on the self-organizing state space (SOSS) model. The AFSA-PF algorithm not only can estimates the system state but also can make the sampling distribution of the unknown parameter to move the true parameter distribution. Ultimately, the true values of the unknown parameters are identified. In this way, the system model can gradually approximate the actual IIoT-based industrial process control system.

1. Introduction

The Internet of Things (IoT) has entered various areas of our lives, such as smart light bulbs and shared bikes to smart locks [1–4]. The IoT connects up a dazzlingly array of things, making our lives more and more convenient. When it comes to industry, the IoT also has lots to offer: industrial control, automation, metering, condition monitoring, and asset management, to name a few.

Coupled with emerging information technologies (ITs) such as cloud computing and big data, the IoT evolves into the industrial IoT (IIoT). The IIoT is of strategic importance in the new era of industrial big data, creating a brand-new industrial ecosystem. Under this ecosystem, intelligent automated machines, advanced predictive analysis, and human-machine collaboration are deeply integrated to enhance productivity, efficiency, and reliability. The IIoT makes it possible to connect embedded modules and products into largescale systems, laying the basis for intelligent industrial monitoring.

Traditionally, industrial processes are monitored by online systems with alarming function. There are two kinds of industrial process monitoring systems: the manufacturing execution system (MES) and process control system. The process control system, which is more popular than the MES, monitors the states of industrial processes with configuration software, focusing on the online detection of machine failures and the change trend of key parameters.

The current process control system provides many modeling methods, namely, principal component analysis (PCA), kernel-PCA (KPCA), and support vector machine (SVM). During operation, the process control system collects the operational data continuously and selects the most suitable modeling method to predict faults online and judge whether it is necessary to issue alarms. However, the prediction accuracy of the current process control system is yet to be improved.

The IIoT-based control of industrial processes has better accuracy than the current process control system. But, the fault prediction is still hindered by the heavy presence of unknown parameters in the control system, which adds to the difficulty in mechanical maintenance. To solve the problem, this paper combines the artificial fish swarm algorithm (AFSA) and the particle filtering (PF) algorithm

into the AFSA-PF algorithm based on the self-organizing state space (SOSS) model and verifies the effectiveness of the algorithm through simulation. The artificial fish school algorithm is a random search optimization algorithm that simulates the ecological behavior of fish schools in the natural environment. It includes four adaptive behaviors of fish schools: foraging behavior, grouping behavior, rear-end behavior, and random behavior. It has been used in many engineering fields and achieved good results.

With the dawn of the Fourth Industrial Revolution (Industry 4.0), the IIoT has emerged as a communication network that allows devices to be accessed easily without sacrificing security and reliability. In general, the IIoT consists of four layers, including a sensing layer, site management layer, network layer, and application layer (the copyright of Figure 1 does not belong to the author).

The sensing layer collects information from various devices, e.g., radio-frequency identification (RFID) readers, cameras, and sensors. In this layer, intelligent wireless systems with sensors can automatically perceive and exchange information between devices and control them in a remote manner.

The site management layer provides an external interface for the industrial data and manages the data locally. This layer is equivalent to a local dispatching and management centre.

The network layer mainly transmits messages and processes information. This layer connects everything together, allowing them to share information with each other.

The application layer adopts suitable technologies to optimize the production processes or industrial applications.

The IIoT is generally deployed in four phases: First, smart sensors collect the real-time industrial data anytime, anywhere. Second, the communication network uploads the collected data in real time. Third, the uploaded data are modelled and analysed based on big data, cloud computing, and data mining. Fourth, the industrial production is upgraded through information management and platform integration, making production more efficient, resource utilization more complete, and production cost lower.

The traditional monitoring systems of industrial processes mainly support online monitoring and early warning. The IIoT-based intelligent monitoring system emphasizes on the diagnosis and prediction of abnormalities through modeling the current states of industrial processes. To issue real-time alarms, the IIoT-based system pushes the alarm information to the relevant management and operation terminals via the mobile network. The distributed message queue middleware is adopted to deliver the alarm messages reliably in real time. All alarm messages are sent from the system to the middleware, according to the type and level of the information.

2. Materials and Methods

The materials and methods section should contain sufficient detail so that all procedures can be repeated. It may be divided into headed subsections if several methods are described.

Based on the Monte Carlo (MC) method and recursive Bayesian estimation, the PF provides a filtering method for nonlinear, non-Gaussian systems [5, 6]. During the PF, the MC method is employed for integral operations in the Bayesian estimation according to the law of large numbers.

The basic idea of the PF algorithm is as follows: First, a set of random samples called particles are generated based on the empirical distribution of system state vectors. Then, the weight and position of the particles are updated continuously according to the measured data. Finally, the initial empirical distribution is corrected as per the updated information.

In essence, the PF algorithm approximates the probability distribution of correlations by discrete random measures, which are composed of particles and their weights, and updates the discrete random degree recursively. If the sample size is large, the MC method will be called to approach the true posteriori probability density function of the state variable.

At present, the PF algorithm has been widely applied in positioning and tracking [7–13], including but not limited to vehicle positioning, aircraft positioning through map matching or terrain aided navigation, integrated navigation, and angle measurement. Some scholars have applied the PF algorithm to single-target air-to-sea passive tracking and multitarget tracking. Other application fields of the algorithm are target recognition, system identification, parameter estimation, automatic control, and fault detection. The wide application is attributable to its excellence in recursive Bayesian estimation of nonlinear, non-Gaussian systems.

For an industrial process monitoring system, the discrete equation of state can be expressed as

$$X_k = \varphi_{k,k-1} X_{k-1} + U_k + W_k, \tag{1}$$

where, $X_k = [x_{e(k)}, v_{e(k)}, a_{e(k)}, x_{n(k)}, v_{n(k)}, a_{n(k)}]^T$ is the state variable of the system; W_k is the state noise vector; U_k is the mean system manoeuvre at time k; and $\phi_{k,k-1}$ is the system state transition matrix.

The discrete observation equation of the system can be defined as

$$\begin{aligned} Y_k &= h(X_k) + V_k, \\ Y_k &= \left[e_{\mathrm{GPS}(k)}, n_{\mathrm{GPS}(k)}, \omega_k, s_k\right]^T, \\ h(X_k) &= \begin{bmatrix} x_{e(k)} \\ x_{n(k)} \\ \dfrac{v_{n(k)} a_{e(k)} - v_{e(k)} a_{n(k)}}{v_{e(k)}^2 + v_{n(k)}^2} \\ \xi T \sqrt{v_{e(k)}^2 + v_{n(k)}^2} \end{bmatrix}, \\ V_k &= \begin{bmatrix} v_{1(k)} \\ v_{2(k)} \\ \varepsilon_{\omega(k)} \\ \varepsilon_{s(k)} \end{bmatrix}, \end{aligned} \tag{2}$$

where Y_k is the observed vector; v_1 and v_2 are the observed noises; ε_ω is the drift of the gyro; ε_s is the observed error; ξ is the scaling factor; and ξ is an uncertain system parameter.

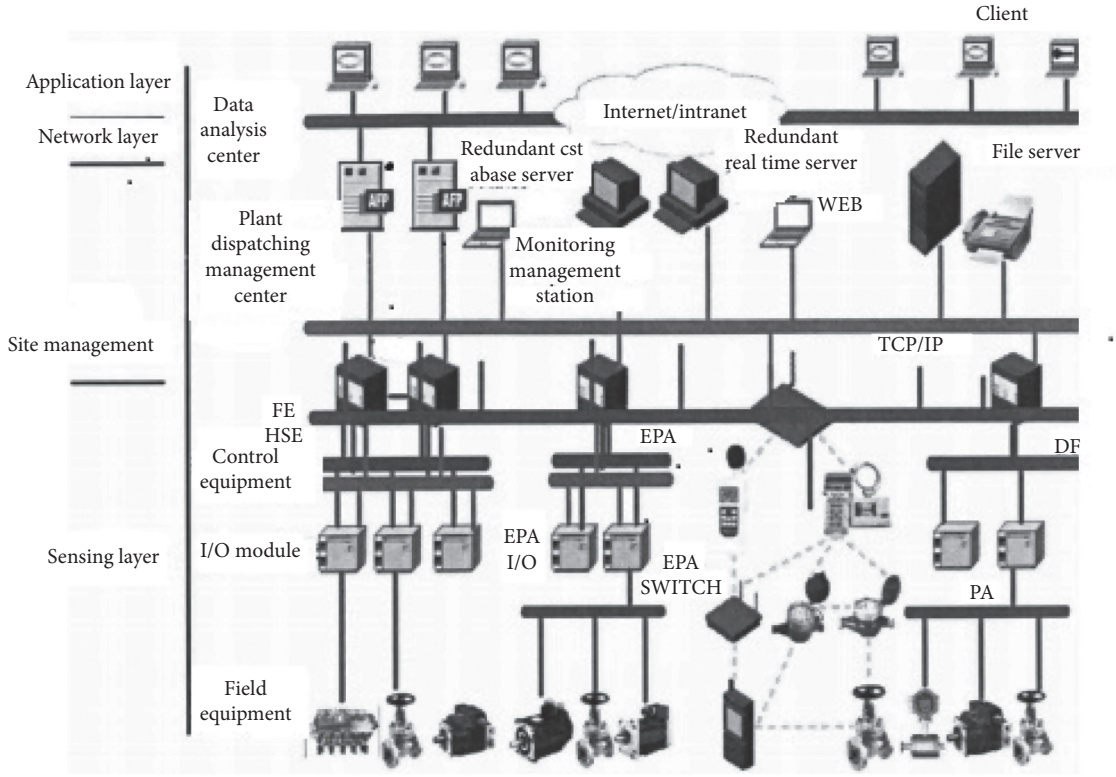


Figure 1: The structure of the IIoT.

The general nonlinear equation can be described as

$$x_{k+1} = f(x_k, w_k), \tag{3}$$

$$z_k = h(x_k, v_k), \tag{4}$$

where $x_k \in R^n$ is the system state vector at time $k \in T$; f: $R^n \times R^n \longrightarrow R^n$ is the state evolution map; $\{w_k\}_{k\in T}$ is the process noise sequence with independent distribution; h: $R^n \times R^m \longrightarrow R^m$ is the measure mapping; and $\{v_k\}_{k\in T}$ is the independent identically distributed sequence of measured noises. The initial state $x_0 \in R^n$ of the system is a random vector.

The goal of process monitoring is to estimate the state based on the measured data $z_{1:k} = \{z_1, \ldots, z_k\}$. Suppose $p(x_0)$ is the priori probability distribution of the state of the given system. Then, the posteriori probability distribution $p(x_{0:k} \mid z_{1:k})$ of the target state x_k at time k can be described as $\{x_{0:k}^i, \omega_k^i\}_{i=1}^N$. This is a set of particles corresponding to the weight $\{\omega_k^i, i = 0, \ldots, N\}$ of $\{x_{0:k}^i, i = 0, \ldots, N\}$, where $x_{0:k} = \{x_j, j = 0, \ldots, k\}$ is a set of states from time 0 to time k.

Normalizing the weight to $\sum_i \omega_k^i = 1$, the posteriori probability distribution of the target state at time k can be discretely weighted to

$$p(x_{0:k} \mid z_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i \delta(x_{0:k} - x_{0:k}^i). \tag{5}$$

The weights are selected by importance sampling. If the set $\{x_{0:k}^i\}_{i=1}^N$ can be obtained by the importance density function $q(x_{0:k} \mid z_{1:k})$, then the weight can be described as

$$\omega_k^i \propto \frac{p(x_{0:k}^i \mid z_{1:k})}{q(x_{0:k}^i \mid z_{1:k})}. \tag{6}$$

If the importance density can be resolved as

$$q(x_{0:k} \mid z_{1:k}) = q(x_k \mid x_{0:k-1}, z_{1:k})\, q(x_{0:k-1} \mid z_{1:k-1}). \tag{7}$$

Then, a new set $\{x_{0:k}^i\}_{i=1}^N$ can be obtained by the particle $\{x_k^i\}_{i=1}^N$ obtained from $q(x_k \mid x_{0:k-1}, z_{1:k})$, and the particle set $\{x_{0:k-1}^i\}_{i=1}^N$ can be obtained from $q(x_{0:k-1} \mid z_{1:k-1})$.

The posteriori probability density function can be expressed as

$$\begin{aligned} p(x_{0:k} \mid z_{1:k}) &= \frac{p(z_k \mid x_{0:k}, z_{1:k-1})\, p(x_{0:k} \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})}, \\ &= \frac{p(z_k \mid x_{0:k}, z_{1:k-1})\, p(x_k \mid x_{0:k-1}, z_{1:k-1})}{p(z_k \mid z_{1:k-1})} \\ &\quad \cdot p(x_{0:k-1} \mid z_{1:k-1}), \\ &= \frac{p(z_k \mid x_k)\, p(x_k \mid x_{k-1})}{p(z_k \mid z_{1:k-1})}\, p(x_{0:k-1} \mid z_{1:k-1}) \\ &\propto p(z_k \mid x_k)\, p(x_k \mid x_{k-1})\, p(x_{0:k-1} \mid z_{1:k-1}). \end{aligned} \tag{8}$$

Substituting (5) and (6) into (7), the importance weight update formula can be obtained as

$$\omega_k^i = \frac{p\left(z_k \mid x_k^i\right) p\left(x_k^i \mid x_{k-1}^i\right) p\left(x_{0:k-1}^i \mid z_{1:k-1}\right)}{q\left(x_k^i \mid x_{0:k-1}^i, z_{1:k}\right) q\left(x_{0:k-1}^i \mid z_{1:k-1}\right)} = \omega_{k-1}^i \frac{p\left(z_k \mid x_k^i\right) p\left(x_k^i \mid x_{k-1}^i\right)}{q\left(x_k^i \mid x_{0:k-1}^i, z_{1:k}\right)}. \tag{9}$$

If $q(x_k \mid x_{0:k-1}, z_{1:k}) = q(x_k \mid x_{k-1}, z_k)$, then the importance density function depends only on x_{k-1} and z_k. In this case, only $\{x_k^i\}_{i=1}^N$ needs to be stored in the calculation, regardless of the set $\{x_{0:k-1}^i\}_{i=1}^N$ and the past measurement $z_{1:k-1}$. The modified weight can be expressed as

$$\omega_k^i \propto \omega_{k-1}^i \frac{p\left(z_k \mid x_k^i\right) p\left(x_k^i \mid x_{k-1}^i\right)}{q\left(x_k^i \mid x_{k-1}^i, z_k\right)}. \tag{10}$$

The standard PF algorithm chooses the easiest achievable priori probability density as the importance density function:

$$q\left(x_k^i \mid x_{k-1}^i, z_k\right) = p\left(x_k^i \mid x_{k-1}^i\right). \tag{11}$$

Replacing (11) with (10), the importance weight can be reduced to

$$\omega_k^i \propto \omega_{k-1}^i p\left(z_k \mid x_k^i\right). \tag{12}$$

Then, the weight ω_k^i can be normalized:

$$\omega_k^i = \frac{\omega_k^i}{\sum_{i=1}^N \omega_k^i}. \tag{13}$$

The posteriori probability density $p(x_k \mid z_{1:k})$ can be expressed as

$$p\left(x_k \mid z_{1:k}\right) \approx \sum_{i=1}^N \omega_k^i \delta\left(x_k - x_k^i\right). \tag{14}$$

It can be seen that, when $N \longrightarrow \infty$, the abovementioned formula can approximate the real posteriori probability $p(x_k \mid z_{1:k})$ by the law of large numbers.

The pseudocode of the standard PF algorithm is given below (Algorithm 1).

The PF algorithm was improved based on the SOSS model to control the industrial processes based on the IIoT. To begin with, the vector of the unknown parameters in the IIoT-based process control system is denoted as θ and extended to the state vector, such that $Z_k = [X_k, \theta_k]^T$. Let Y_k be a stochastic noise. If the stochastic step-size model $\theta_k = \theta_{k-1} + \gamma_k$ is applied, then the state space of the process control system can be modelled as

$$Z_k = \begin{bmatrix} \varphi_{k,k-1} & 0 \\ 0 & 1 \end{bmatrix} Z_{k-1} + \begin{bmatrix} U \\ 0 \end{bmatrix} + \begin{bmatrix} W_k \\ \gamma_k \end{bmatrix}, \quad Y_k = H(Z_k) + V_k. \tag{15}$$

We suppose the initial set θ_0^i, $\{i = 1, \ldots, N\}$ of the parameter vector θ is sampled from the uniform distribution of (3):

$$\theta_{j,0}^i - U\left(P_j - r_j, P_j + r_j\right), \tag{16}$$

where $\theta_{j,0}^i$ is the j − th component of θ_0^i; $U(P_j - r_j, P_j + r_j)$ is the uniform distribution from $P_j - r_j$ to $P_j + r_j$; P_j is the centre of $(P_j - r_j, P_j + r_j)$; and r_j is a real number chosen by the user according to the range of parameter changes.

Given the priori condition probability $\rho(X_0)$ of the system state, $\{Z_k^i, w_k^i\}_{i=1}^N = (\{X_k^i, \theta_k^i\}_{i=1}^N, \{w_k^i\}_{i=1}^N)$ can be used to describe the posteriori probability distribution $\rho(Z_k \mid Y_k)$ of the target state Z_k at time k, and $\{Z_k^i, i = 1, \ldots, N\}$ can be used to represent the particle set that corresponds to the weight $\{w_k^i, i = 1, \ldots, N\}$.

Normalizing the weight to $\sum_{i=1}^N w_k^i = 1$, the pseudocode of the SOSS-based PF algorithm can be programmed as follows (Algorithm 2).

In normal cases, the SOSS-based PF algorithm can effectively filter unknown parameters of the IIoT-based process control system. However, the performance of the algorithm depends heavily on the sampling quality, which is a subproblem of initial sampling. Unfortunately, the initial distribution of unknown parameters is not easy to determine. If the initial samples only have a few values in the concentration zone of posteriori distribution, it would be difficult for the SOSS-based PF algorithm to filter the unknown parameters well.

To solve the problem, the AFSA [14–20], a stochastic search optimization algorithm, was introduced to improve the SOSS-based PF algorithm. Mimicking the behaviors of fish swarms in nature, the AFSA includes four adaptive behaviors: foraging, clustering, chasing, and random walk. This algorithm has been successfully applied in many engineering fields. The AFSA-based optimization of the SOSS-based PF algorithm is implemented as follows.

First, the search space Ω of unknown parameters was divided into several equal-size subspaces $\Omega^1, \Omega^2, \ldots, \Omega^M$, each of which is defined as an artificial fish (AF). The AFs simulate the random search process of the fish in nature. As shown in Figure 2, the relationship between each subspace and the original search space can be represented as $\Omega^1 \bigcup \cdots \bigcup \Omega^M \subseteq \Omega$. Then, the optimal distribution space of unknown parameters was obtained.

Through the abovementioned process, the SOSS-based AFSA-PF algorithm was obtained. This global search algorithm can effectively estimate system states and derive a realistic distribution of unknown parameters. The pseudocode of the SOSS-based AFSA-PF algorithm is given below (Algorithm 3).

$[\{x_k^i, \omega_k^i\}_{i=1}^N] = SIS[\{x_{k-1}^i, \omega_{k-1}^i\}_{i=1}^N, z_k]$
{
A particle $\{X_0^i\}_{i=1}^N$ is produced by a priori probability $p(x_0)$, and the total weight of all particles is $1/N$, making $k = 1$.
For $k = 1, \ldots, T$
For $i = 1, \ldots, N$
Prediction: Use equation of State (1) to predict the unknown state x_k^i.
Calculate particle weights:
$\omega_k^i = \omega_{k-1}^i p(z_k|x_k) = \omega_{k-1}^i \rho_{ek}(z_k - h(x_k^i))$, $i = 1, \ldots, N$
End For
Normalized weight,
For $i = 1, \ldots, N$
$\omega_k^i = \omega_k^i / \sum_{i=1}^N \omega_k^i$.
End For
Then, the minimum mean square estimate of the state vector at time k is obtained as $\hat{x}_k \approx \sum_{i=1}^N \omega_k^i x_k^i$. The priori condition probability density of system states is chosen as the important density function ρ_{ek}.
Resampling: A new set of particles $\{x_k^i\}_{i=1}^N$ is obtained.
End For
}

Algorithm 1: The standard PF algorithm.

$[P] = \mathrm{SOSS}(\{X_0^i\}_{i=1}^N, \{\theta_0^i\}_{i=1}^N, \Omega)$
{
A particle set $\theta_0^i \sim U(P - r, P + r)$, $(i = 1, \ldots, N)$, is initialized, and a particle $\{\mathbf{X}_0^i\}_{i=1}^N$ is produced by a priori probability $\rho(\mathbf{X}_0)$, with $\{\mathbf{Z}_0^i\}_{i=1}^N = (\{\mathbf{X}_0^i\}_{i=1}^N, \{\theta_0^i\}_{i=1}^N)$, and all particles have a weight of $1/N$, $k = 1$.
For $k = 1, \ldots, T$
For $i = 1, \ldots, N$
Prediction: Use equation of state to predict the unknown state $\mathbf{Z}_k^i$
Calculate particle weights:
$w_k^i = w_{k-1}^i \rho(Y_k \mid \mathbf{Z}_k) = w_{k-1}^i \rho_{ek}(Y_k - H(Z_k^i))$, $i = 1, \ldots, N$
End For
Normalized weight,
For $i = 1, \ldots, N$
$w_k^i = w_k^i / \sum_{i=1}^N w_k^i$.
End For
Then, the minimum mean square estimate of the state vector at time k is obtained as $\widehat{Z}_k \approx \sum_{i=1}^N w_k^i Z_k^i$. The priori condition probability density of system states is chosen as the important density function ρ_{ek}.
Resampling: $\{Z_k^i\}_{i=1}^N = \mathrm{resample}(\{X_k^i\}_{i=1}^N, \{\theta_k^i\}_{i=1}^N)$
End For
}

Algorithm 2: The SOSS-based PF algorithm.

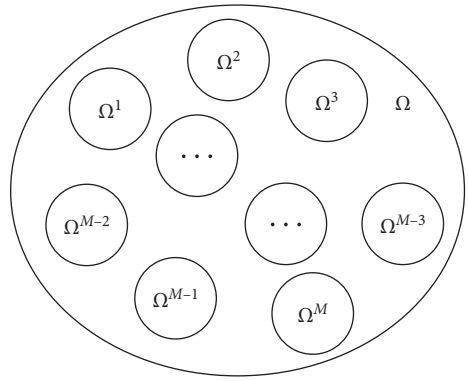


Figure 2: The relationship between the original space and subspaces.

$[P] = \mathrm{ESOSS}(\{\mathbf{X}_0^i\}_{i=1}^N, \{\theta_0^i\}_{i=1}^N, \Omega)$
{
For each subset Ω^q, $(q = 1, \ldots, M)$, initialize the particle set $\theta_0^{q,i} \sim U(c^q - d, c^q + d)$. Particle $\{\mathbf{X}_0^i\}_{i=1}^N$ is produced by a Priori Probability $\rho(X_0)$. $\{\mathbf{Z}_0^{q,i}\}_{i=1}^N = (\{\mathbf{X}_0^i\}_{i=1}^N, \{\theta_0^i\}_{i=1}^N, \Omega^q)$, All the particles have a weight of $1/N$, let $k = 1$.
For $k = 1, \ldots, T$
For $\Omega^q = \Omega^1, \ldots, \Omega^M$
For $i = 1, \ldots, N$
PREDICTION: Using Equation of state (17) to predict unknown states $\mathbf{Z}_k^{q,i}$.
Calculate particle weights:
$w_k^i = w_{k-1}^i \rho(Y_k \mid Z_k) = w_{k-1}^i \rho_{ek}(Y_k - H(Z_k^{q,i}))$, $i = 1, \ldots, N$
End For
Normalized weight,
For $i = 1, \ldots, N$
$w_k^i = w_k^i / \sum_{i=1}^N w_k^i$.
End For

Algorithm 3: The SOSS-based AFSA-PF algorithm.

3. Results

The results and discussion may be presented separately, or in one combined section, and may optionally be divided into headed subsections.

To verify its adaptability, the SOSS-based AFSA-PF algorithm was applied to estimate the state and identify the parameters of an IIoT-based process control system with unknown parameters. During the simulation, the error level of each sensor was adjusted: the observed noise was changed to $\varepsilon_{\omega} \sim N(0.04^2)$, and that of the odometer was kept at $\varepsilon_s \sim t(1)$. The initial state X_0 and initial value P_0 of the error variance matrix of the system were also kept stable. The SOSS-based AFSA-PF algorithm was first adopted to identify the unknown parameters δ_1, δ_2, and ξ (Figure 3).

The system states can be described by the following nonlinear mathematical model:

$$\begin{aligned}
\dot{x}_1(t) &= x_2(t),\\
\dot{x}_2(t) &= \frac{9.6\cos(x_1(t)) - \sin(x_1(t))\left(au(t) + bx_2^2(t)\sin(x_1(t)) - acx_4(t)\right.}{d - b\cos^2(x_1(t))},\\
\dot{x}_3(t) &= x_4((t)),\\
\dot{x}_4(t) &= au(t) + b\left(x_2^2(t)\sin(x_1(t)) - \dot{x}_2(t)\cos(x_1(t))\right) - acx_4(t),
\end{aligned} \tag{17}$$

where $x_1(t)$, $x_2(t)$, $x_3(t)$, and $x_4(t)$ are the system state; $u(t)$ is the control quantity of the system; and $a = 0.6$, $b = 0.020$, $c = 0.1$, and $d = 0.1865$ are the parameter variables of the system.

In each subset, the particle θ^i, $\{i = 1, \ldots, N\}$ of the unknown parameter θ is assumed to satisfy uniformly distributed sampling, that is, $\theta^i - U(c^q - d, c^q + d)$.

Based on the abovementioned artificial intelligence (AI) model, the proposed algorithm can be implemented in the following steps.

When the unknown parameter belongs to space Ω^q, the least mean square estimate of the state vector at time k can be given as

$$\widehat{Z}_k^q \approx \sum_{i=1}^N w_k^i Z_k^{q,i}. \tag{18}$$

The objective function can be expressed as

$$\psi = \rho_{ek}\left(Y_k - H\left(\widehat{Z}_k^q\right)\right), \tag{19}$$

where ρ_{ek} is the priori condition probability density of system states. The estimated value of unknown parameter θ_k^q in vector $\widehat{Z}_k^q$ can be adopted as the new centre of space Ω^q, i. e., $c^q = \theta_k^q$.

3.1. Foraging. Let c^q ($v_{q,l} <$ visual) be the current state of an AFs. Then, the AF chooses a state c^l in its perception range. If $\psi^q < \psi^l$, the AF will move one step in that direction; otherwise, the AF will move one step in a random direction:

$$\begin{bmatrix} \text{If } \psi^q < \psi^l, \ c_k^{q\text{next}} = c^q + \text{rand}(\cdot)\text{Step}\dfrac{c_k^l - c_k^q}{\left\|c_k^l - c_k^q\right\|}, \\ \text{Else}, \qquad c_k^{q\text{next}} = c^q + \text{rand}(\cdot)\text{Step}, \end{bmatrix}, \tag{20}$$

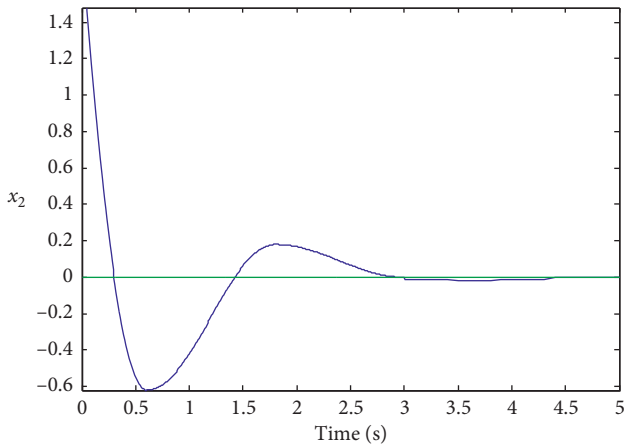


Figure 3: The unknown parameters identified by the SOSS-based AFSA-PF algorithm.

where Step is the maximum step length of the AF; rand $(\cdot)$ is a random number between 0 and 1.

3.2. Clustering. Let c^q be the current state of an AF and n_f be the number of AFs in other clusters in the perception range of the AF. Then, the centre of the neighbourhood of the AF can be described as $c^c = 1/n_f \sum_{l\neq q}^{n_f} c^l$. Then, it is possible to obtain the new unknown parameter Range $\theta^c - U(c^c - d, c^c + d)$, and its distribution $\{Z_{k-1}^{c,i}\}_{i=1}^N = \text{sample}(\{X_{k-1}^{c,i}\}_{i=1}^N, \{\theta_{k-1}^{c,i}\}_{i=1}^N, \Omega^c)$. Through particle weight prediction and calculation, we have $\widehat{Z}_k^c \approx \sum_{i=1}^N w_k^i Z_k^i$, and thus, $\psi^c = \rho_{ek}(Y_k - H(\widehat{Z}_k^c))$. If $\psi^c/n_f > \delta \cdot \psi^q$ (δ is the crowding factor), the centre of the sampling space Ω^q for the particle set θ^i in the next time $(\theta_k^{q,i} - U(c^{q\text{next}} - d, c^{q\text{next}} + d))$ can be calculated by

$$c_k^{q\text{next}} = c_q + \text{rand}(\cdot)\text{Step}\frac{c_k^c - c_c^q}{\left\|c_k^c - c_c^q\right\|}. \tag{21}$$

Otherwise, the AF will move as it does in the foraging process.

3.3. Bulletin Board. In our algorithm, a bulletin board is set up to record the optimal value of the current objective function. After each move, the AF compares its objective function Ω^q with the objective function Ω^o of the bulletin board. If the former is better, the objective function Ω^o of the bulletin board will be replaced with the objective function value of the AF. The state vector of the optimal value is regarded as the optimal value of $\widehat{Z}_k$.

Resampling:
$\{Z_k^{q,i}\}_{i=1}^N = \text{resample}(\{X_k^i\}_{i=1}^N, \{\theta_k^i\}_{i=1}^N, \Omega^q)$
End For
End For

The local linearization of the abovementioned nonlinear system is equivalent to the model of the original nonlinear system:

Rule 1 : If $x_1(t)$ is about 0
$$\dot{x}(t) = (A_1 + D\Delta E_{a1})x(t) + (B_1 + D\Delta E_{b1})u(t)$$
Rule 2 : If $x_1(t)$ is about $\pm$ 1.57
$$\dot{x}(t) = (A_2 + D\Delta E_{a2})x(t) + (B_2 + D\Delta E_{b2})u(t)$$

The parameters of the correlation matrix are as follows:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 30.174 & 0 & 0 & 0.2326 \\ 0 & 0 & 0 & 1 \\ -0.5957 & 0 & 0 & -0.0842 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0 \\ -2.1268 \\ 0 \\ 0.7533 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 20.1702 & 0 & 0 & 0.1709 \\ 0 & 0 & 0 & 1 \\ 0.0165 & 0 & 0 & -0.051 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 0 \\ -1.6278 \\ 0 \\ 0.6214 \end{bmatrix},$$

$$E_{a1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 99 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 6 \end{bmatrix}, \tag{22}$$

$$E_{b1} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 0 \end{bmatrix},$$

$$E_{a2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 6 \end{bmatrix},$$

$$E_{b2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

$$D = 0.01 \times I,$$

where I is a fourth-order identity matrix.

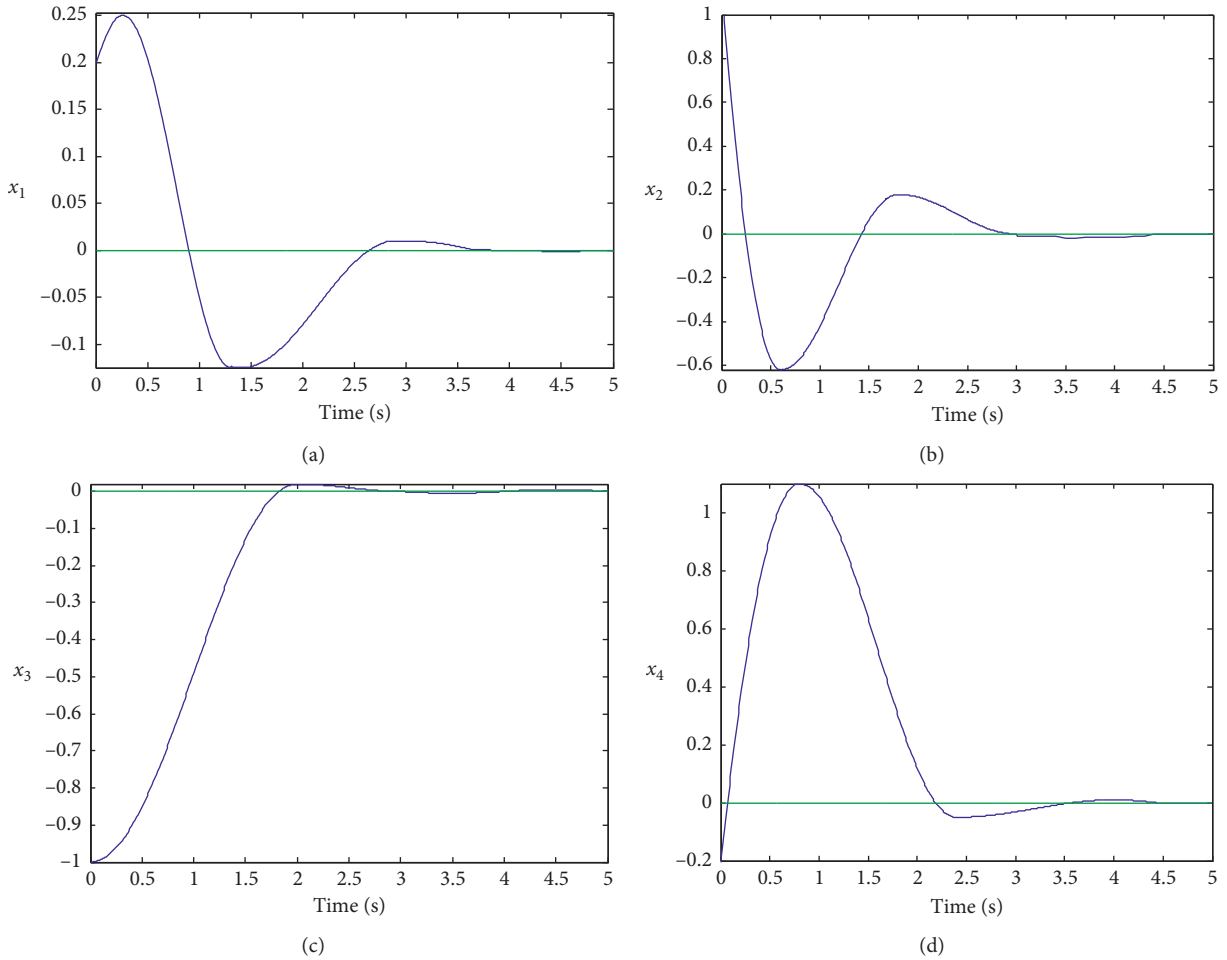


Figure 4: The actual state response of the system.

Assuming that the initial state is $x(0) = [0.1, 1, -1, -0.1]^T$, the weight matrix $Q = \text{diag}\{2, 0.02, 2, 0.02\}$ of system performance indices and $R = 0.02$ can be established. Finally, the state feedback matrix of the controller can be obtained as

$$F_1 = [\,332.1445 \quad 89.3544 \quad 76.3622 \quad 145.9341\,],$$
$$F_2 = [\,50.5252 \quad 163.3102 \quad 145.1502 \quad 136.9229\,]. \tag{23}$$

The upper bound of the corresponding system performance index can be obtained as

$$J^* = x_0^T P x_0 = 22.3245. \tag{24}$$

The designed controller was applied to the model of the original nonlinear system, and the simulation time was set to 5 s. The simulation results are displayed in Figures 4 and 5, where the blue solid lines and green dotted line stand for the actual state and control quantity of the system, respectively.

To verify its antijamming performance, the designed controller was simulated after the state vectors $x_1(t)$ and $x_3(t)$ of the original system were added pulse jamming signals at 1.28 s and 2.79 s, respectively. The simulated results on state variables and control variables are presented in Figures 6 and 7, respectively.

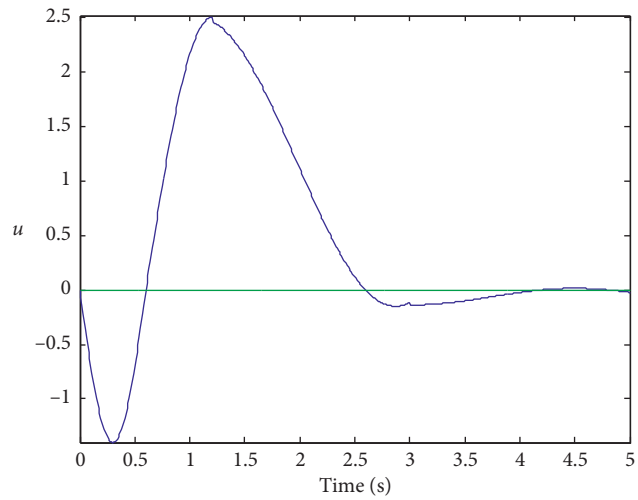


Figure 5: The controlled state response of the system.

To verify its robustness to variable system parameters, the designed controller was further simulated after changing system parameters from $b = 0.036$, $c = 0.1$, and $d = 0.365$. The simulated results on state variables and control variables are presented in Figures 8 and 9, respectively.

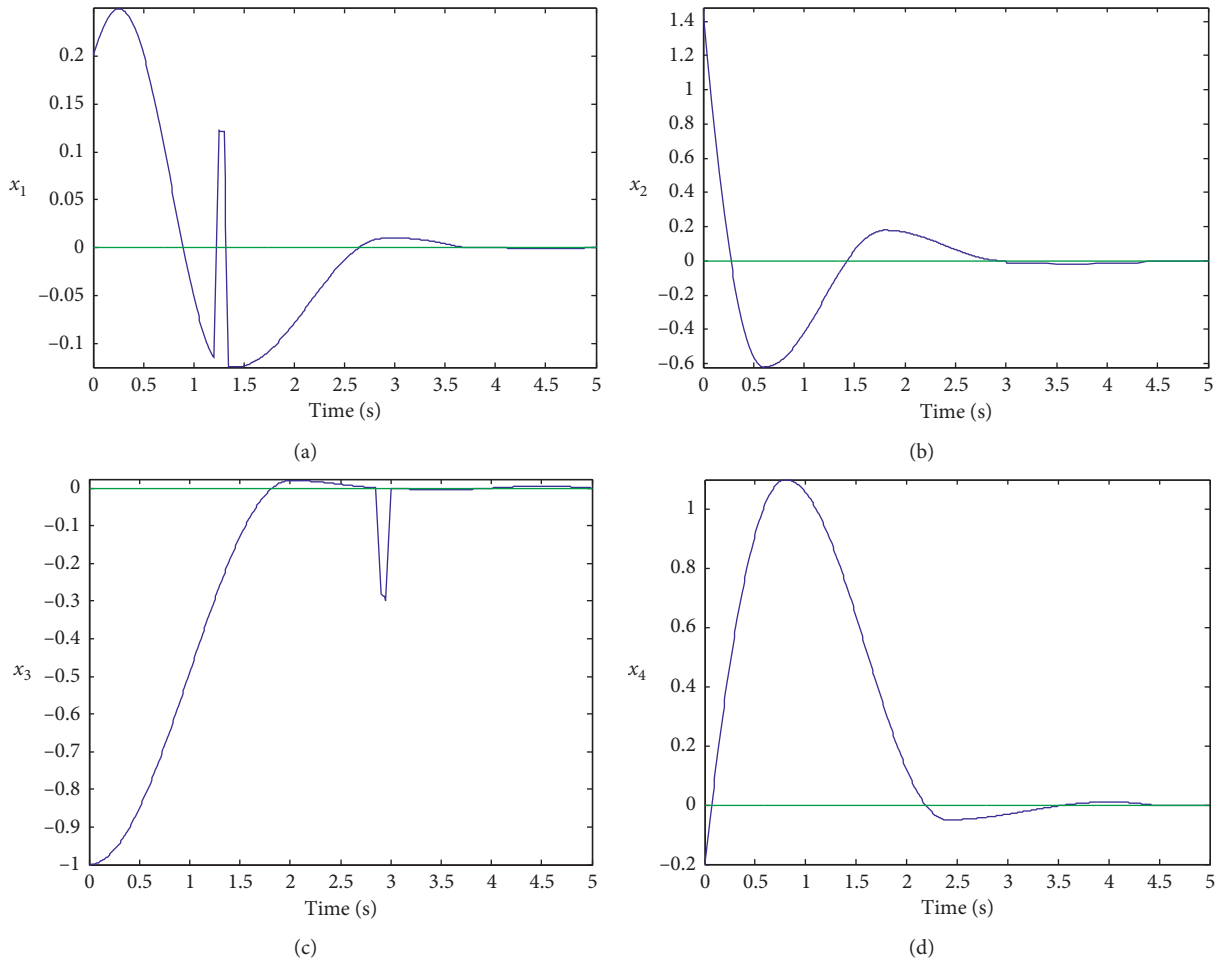


Figure 6: The actual state response of the system under jamming signals.

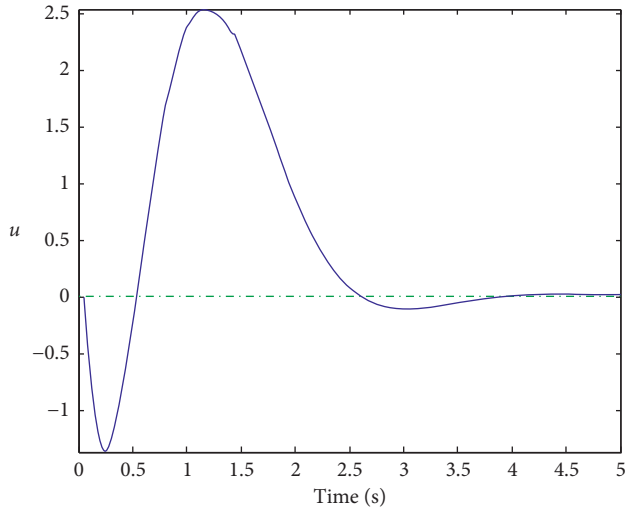


Figure 7: The controlled state response of the system under jamming signals.

From the abovementioned simulation results, the designed controller was found to keep the nonlinear system with unknown parameters stable and bring the system back to equilibrium state, despite the addition of jamming signals and the variation in system parameters. Changing the error level of sensors could affect the system error and also influence the

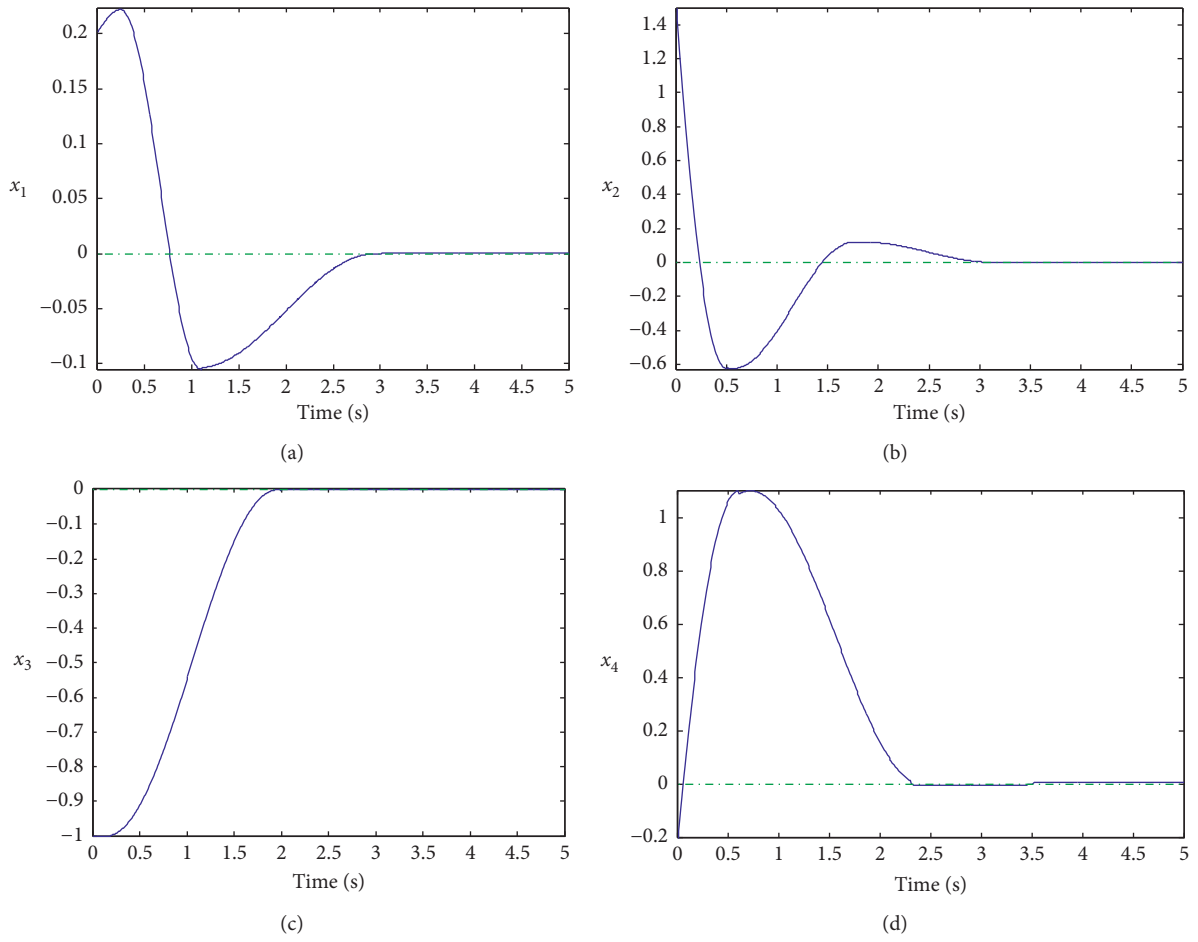


Figure 8: The actual state response of the system at $b = 0.035$, $c = 0.1$, and $d = 0.365$.

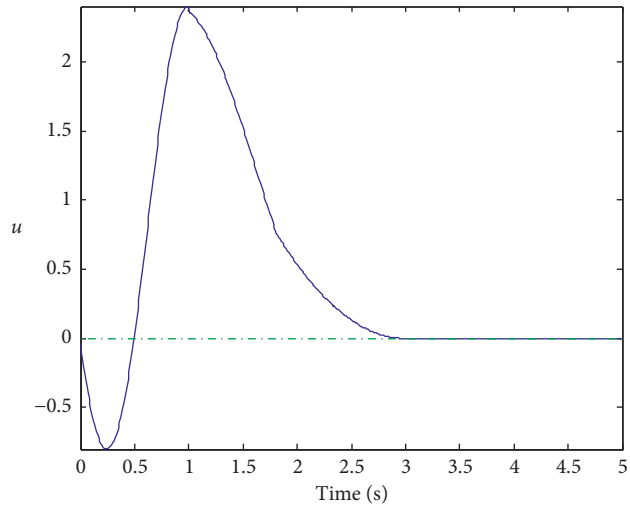


Figure 9: The controlled state response of the system at $b = 0.038$, $c = 0.1$, and $d = 0.364$.

identification of unknown parameters. The stronger the sensor noise, the greater the system error. However, the influence on unknown parameters is way smaller than the location error. As a result, the proposed algorithm could accurately estimate the estimation parameters.

4. Conclusions

The existing filtering algorithms for the IIoT-based process control system only work if the system model is deterministic. In real-world scenarios, it is very difficult to find an

IIoT-based process control system that is truly deterministic. The model parameters of the system are often unknown, though their ranges might be identifiable. To eliminate the effect of the unknown parameters on the IIoT-based control of industrial processes, this paper develops a novel AFSA-PF algorithm based on the SOSS model. The SOSS-based AFSA-PF algorithm can effectively estimate system states, approximate the actual distribution of unknown parameters, and disclose the true value of the unknown parameters. With the aid of the proposed algorithm, the system model could reflect the actual states of the IIoT-based process control system. The research findings provide a good reference for applying AI techniques in industrial process control.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by High-Level Talents Foundation of Ankang University (Grant No. 2019AYQDZR01) and Ankang Science and Technology Research and Development Project (Grant No. AK2019SF-09).

References

[1] R. Fedele, M. Merenda, F. Giammaria, R. Carotenuto, and F. G. D. Corte, "Energy harvesting for IoT road monitoring systems," *Instrumentation Mesure Métrologie*, vol. 18, no. 4, pp. 605–623, 2018.

[2] A. Rahaman, M. Islam, M. Islam, M. Sadi, and S. Nooruddin, "Developing IoT based smart health monitoring systems: a review," *Revue d'Intelligence Artificielle*, vol. 33, no. 6, pp. 435–440, 2019.

[3] D. Krčmařík, M. Petrů, and R. Moezzi, "Innovative IoT sensing and communication unit in agriculture," *European Journal of Electrical Engineering*, vol. 21, no. 3, pp. 273–278, 2019.

[4] W.-T. Sung and Y.-C. Hsu, "Designing an industrial real-time measurement and monitoring system based on embedded system and ZigBee," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4522–4529, 2011.

[5] N. G. Polson, J. R. Stroud, and P. Müller, "Practical filtering with sequential parameter learning," University of Pennsylvania Working Paper, Philadelphia, PA, USA, 2006.

[6] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[7] A. Ataei and Q. Wang, "Non-linear control of an uncertain hypersonic aircraft model using robust sum-of-squares method," *IET Control Theory & Applications*, vol. 6, no. 2, pp. 203–215, 2012.

[8] J. Chen, C. P. Pan, and H. C. Zhao, "One near space hypersonic aircraft attack angle observer and controller design," in *Proceedings of the 2011 Chinese Control and Decision Conference (CCDC)*, Mianyang, China, May 2011.

[9] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[10] S. K. Zhou, R. Chellappa, and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1491–1506, 2004.

[11] H. H. Fu, J. Xu, H. Zhang, M. Zhang, and X. X. Xu, "A novel video target tracking method based on lie group manifold," *Traitement du Signal*, vol. 35, no. 3-4, pp. 331–340, 2018.

[12] G. Kitagawa, "Theory and methods," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1203–1215, 1998.

[13] M. Hürzeler and H. R. Künsch, "Approximating and maximising the likelihood for a general state-space model," *Sequential Monte Carlo Methods in Practice*, Springer, New York, NY, USA, 2001.

[14] K. Yano, "A self-organizing state space model and simplex initial distribution search," *Computational Statistics*, vol. 23, no. 2, pp. 197–216, 2008.

[15] M. Jiang and K. Zhu, "Multiobjective optimization by artificial fish swarm algorithm," in *Proceedings of the 2011 IEEE International Conference on Computer Science and Automation Engineering*, Shanghai, China, June 2011.

[16] D. Li, F. Qian, and J. Gao, "Performance-first control for discrete-time LQG problems," *IEEE Transactions on Automatic Control*, vol. 54, no. 9, pp. 2225–2230, 2009.

[17] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, vol. 96, no. 2, pp. 293–320, 2003.

[18] D. Zhang, J. Tan, H. Tian, Z. Wang, and W. Guo, "Aquifer parameter inversion by artificial fish swarm algorithm based on quantum theory," *Ingénierie des Systèmes d Information*, vol. 24, no. 1, pp. 29–33, 2019.

[19] P. Guo, X. Wang, and Y. Han, "The enhanced genetic algorithms for the optimization design," in *Proceedings of the 2010 3rd International Conference on Biomedical Engineering and Informatics*, Yantai, China, October 2010.

[20] C.-T. Cheng, K. Fallahi, H. Leung, and C. K. Tse, "A genetic algorithm-inspired UUV path planner based on dynamic programming," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1128–1134, 2012.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 4715082, 16 pages
https://doi.org/10.1155/2020/4715082

Research Article

Mathematical Analysis of Heat Transfer in Peristaltic Transport through a Rough Nonuniform Inclined Channel

R. Shukla,[1] **A. Medhavi**,[1] **S. S. Bhatt**,[2] **and R. Kumar**[3]

[1]*Department of Mechanical Engineering, Kamla Nehru Institute of Technology, Sultanpur, India*
[2]*Department of Applied Sciences and Humanities, Faculty of Engineering & Technology, Lucknow University, Lucknow, India*
[3]*B. N. College of Engineering and Technology, Lucknow, India*

Correspondence should be addressed to S. S. Bhatt; shankarbhatt56@gmail.com

Received 14 July 2020; Revised 4 August 2020; Accepted 8 August 2020; Published 24 August 2020

Guest Editor: Hussein Abulkasim

A mathematical study for investigating heat transfer for the peristaltic flow of a Newtonian fluid through a nonuniform inclined channel while considering surface roughness of the inner wall has been conducted. Special attention is focused to analyze the physical behaviour of various flow characteristics at different magnitudes of surface roughness parameter. The results obtained for temperature have been characterized by two boundary conditions: uniform surface temperature and uniform heat flux. Effect of inner wall surface roughness and angle of inclination on the heat transfer for the peristaltic flow through a nonuniform channel has been discussed in detail. The expression for pressure gradient, velocity, and the temperature was obtained. The effect of inner wall surface roughness and the angle of inclination, on pressure rise and temperature, has also been elaborated in detail. The problem formulation developed has been solved using the long wavelength and low Reynolds number assumption.

1. Introduction

Peristaltic transport has gained the attention of researchers during the last five decades due to its applications in industry and physiology. Peristaltic motion involves a continuous progressive wave of area shrinking and enlargement, along the length of an expandable duct or tube. Peristaltic motion is an involuntary neuromuscular function, which occurs in several body organs of living beings, like the movement of fodder in the esophagus, chyme transport, urine deliver to the bladder through the ureter, transport mechanism in the cervical canal, ova transportation from the ovary to the uterus, fluid flow in ductus efferent, blood movement from small blood arteries, and lymph transportation through the lymphatic vessels. The principle involved in this type of motion is used for designing various contemporary medical equipments such as roller-finger pumps and pumping of blood in the heart-lung machine. This mechanism is applied to the transport of toxic liquids in nuclear installations to prevent the mixing of radioactive elements in the environment. Apart from the above, this phenomenon finds applications in slurry transport, corrosive chemicals, and sanitary liquids from one point to the other.

Profound knowledge of the subject of peristaltic transport can be obtained from the vast literature available on the same. The first study in this area was conducted by Latham [1], which was further elaborated by Shapiro et al. [2] in detail. An effect of nonuniformity of the tube as well as the channel was observed in detail by Gupta and Seshadri [3] taking vas deferens as a reference. Heat transfer analysis of a toxic liquid subjected to peristaltic transport through a channel in nuclear operating power plants was carried out by Radhakrishnamacharya and Murty [4]. Excellent investigations have been done by El Koumy et al. [5] and Mekheimer et al. [6] to study the effect of constant magnetic field on peristaltic transport of electrically conducting Maxwell fluid, through a two-dimensional channel and microchannel, having porosity, with the consideration of Hall effect. An

important study to find the effect of rotation on couple stress fluid flowing through a channel by peristaltic mechanism was done by Abd Elmaboud et al. [7]. Tripathi and Anwar Bég [8] studied peristaltic transport (chyme dynamics) through a diseased intestine which acts as a nonuniform porous medium. Two eminent investigations, for compressible, electrically conducting Jeffrey fluid and non-Newtonian nanofluid in a microfluidic and nonuniform channel considering the combined influence of the magnetic field and rheological properties with the consideration of slip velocity were conducted by Abdelsalam and Vafai [9] and Abdelsalam and Bhatti [10]. The collective impact of magnetic field, heat transfer, and porosity on the magneto-hydrodynamic peristaltic motion of the fluid through a channel has been investigated by Eldesoky et al. [11], and a decrement in temperature profile was found due to inclusion of fluid suspension. Victor and Shah [12, 13] found a solution for the energy equation of blood flow in a tube that is subjected to uniform wall temperature and heat flux in a fully developed region of velocity and temperature. The peristaltic mechanism exists in several body organs. The complications of blood flow using the peristaltic flow mechanism in a uniform and nonuniform channel as well as in tube were analyzed by Srivastava and Srivastava [14], supposing blood as a Casson fluid. Mekheimer [15] performed an analysis of blood flow during peristaltic motion through a nonuniform channel subjected to a magnetic field. An outstanding observation in the literature [16] is done to find the behavioural change in a small blood artery in which fluid with the suspended particle is transported by the peristaltic mechanism. Later, Abd Elmaboud et al. [17] conducted a study for the electromagnetic flow of two immiscible fluids, one permeable matrix soaked in a Newtonian fluid and the other pure fluid passes through the inclined channel, and it was found that electric field enhances velocity profile. Admirable investigations were done in studies by Abdelsalam et al. [18], Sohail et al. [19], and Abdelsalam and Sohail [20] to find the impact of heat and mass transfer during the transportation in the effect of the magnetic field through a two-dimensional channel and 3D flow. Sohail et al. [21] addressed the effect of momentum, heat, and mass transfer on the mixed convection boundary layer of Casson liquid for a porous medium. Sadaf and Abdelsalam [22] revealed the hostile effect of nanofluid in a wavy nonuniform annulus with flexible boundaries. Some other worthy research studies by Abdelsalam and Bhatti [23], Abdelsalam and Bhatti [24], Abdelsalam and Bhatti [25], and Eldesoky et al. [26] have been performed with magnetic field effect in a tube for blood flow. Biological organs in which biofluids flow have a small degree of roughness or uncertain (random) spatial variability. Park et al. [27] calculated the effective blood viscosity, with the observation of endothelium roughness of coronary arteries. The degree of roughness within the tissues is very helpful in recognizing the diseases. Evaluation of surface roughness of coronary arteries, in which blood flow takes place through the peristaltic mechanism, helps in developing a remedy for coronary heart disease. Burton and Espino [28] investigated the consequences of intense physiological loading on surface roughness; from the observations, it was revealed that initially, mechanical overloading has no significant effect on roughness; however, when a chemical processing was applied on the surface, an increase in surface roughness was noticed in the longitudinal and circumferential directions. This intense overloading results in damage to the arteries, which is notable after dehydration of the tissue.

Several analyses have been performed in the present area for mini and microchannels. Bhatti et al. [29] conducted a study on the swimming pattern of gyrotactic microorganisms in a blood-based nanofluid flowing through the anisotropically narrowing artery. Veatch [30] carried out experiments on eleven rectangular passages of different cross sections, sizes, and roughness; experiments were performed using four different fluids, and it was concluded that, for laminar flow, friction factor and Reynolds number are independent of roughness. Wall roughness has a significant impact on the heat transfer rate during fluid motion. Study of the impact of wall roughness while considering the roughness orientation, roughness structure, and shape of the roughness element has become a key area of research due to its vast applications in engineering. Excellent studies in the field of heat transfer for the roughed wall have been performed by Han et al. [31] and Taslim et al. [32]. Taylor et al. [33] found the influence of roughness on the fluid motion for surfaces having up to 14% relative roughness, and constricted flow diameters were not considered for analysis during the study; however, till that time, the studies were limited to 5% of relative roughness, which violated micro- and macrofluidic flow systems. Wagner and Kandlikar [34] and Dharaiya and Kandlikar [35] investigated experimentally the influence of sinusoidal structured roughness on heat transfer rate during liquid transport through micro- and macrochannels; experiments were performed for different pitches and roughness heights for a correct understanding of frictional losses in fully developed laminar flow. Kharati-Koopaee and Zare [36] and Attalla et al. [37] performed a study to find the heat transfer effect during fluid flow in microchannels having a rectangular cross section with aligned and offset structured sinusoidal roughness. From

the above study, it was concluded that, to lower the pressure loss for the fluid under consideration, high roughness height or low channel height gave favourable results.

From the above, it is evident that, though many investigations have been undertaken for explaining the variation in heat transfer rate due to surface roughness during fluid flow, the variation in heat transfer rate due to wall roughness for peristalsis mechanism during transport of fluid has yet to be explored. The current analysis involves the study of heat transfer, while pondering the wall roughness for a two-dimensional nonuniform inclined channel when a Newtonian fluid flows through it under the influence of peristaltic waves.

2. Problem Formulation

Figure 1 shows the flow geometry, for the laboratory frame of reference. It is supposed that the walls of the nonuniform channel have sinusoidal-shaped roughness. The study has been carried out in the rectangular coordinate system (X, Y).

The wall geometry is

$$H(X,t) = d(X) + b\sin\left(\frac{(2\pi(X-ct))}{\lambda}\right) - b_1\cos^4\left(\frac{\pi X}{\lambda_1}\right), \tag{1}$$

with

$$d(X) = a + KX, \tag{2}$$

where a represents the half channel width at entry, b is the amplitude, λ is the wavelength, c is the wave propagation velocity, K is the nonuniformity parameter, λ_1 is the pitch, t is the time, X is the axial variable, and b_1 is the roughness height.

Continuity equation is as follows:

$$\frac{\partial U}{\partial X} + \frac{\partial U}{\partial Y} = 0. \tag{3}$$

Momentum equation is as follows:

$$\rho\left(\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial X} + V\frac{\partial U}{\partial Y}\right) = -\frac{\partial p}{\partial X} + \mu\left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2}\right) + \rho g\sin\alpha, \tag{4}$$

$$\rho\left(\frac{\partial V}{\partial t} + U\frac{\partial V}{\partial X} + V\frac{\partial V}{\partial Y}\right) = -\frac{\partial p}{\partial Y} + \mu\left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2}\right) - \rho g\cos\alpha, \tag{5}$$

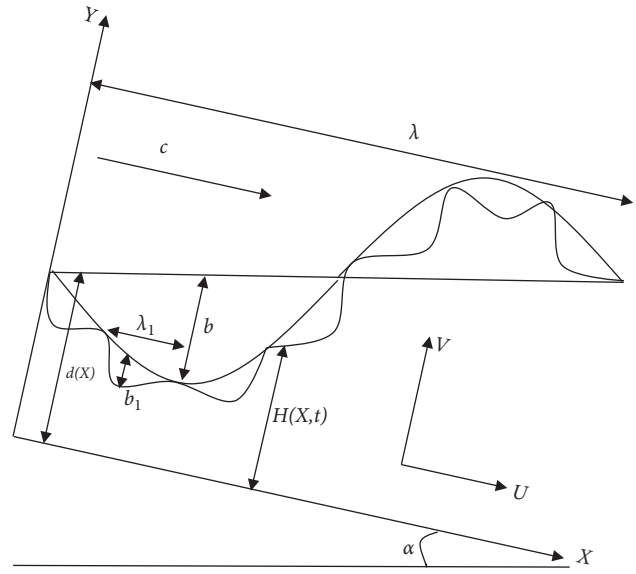


Figure 1: Geometry of the problem.

$$C_p\left(\frac{\partial T}{\partial t} + U\frac{\partial T}{\partial X} + V\frac{\partial T}{\partial Y}\right) = \frac{k}{\rho}\left(\frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2}\right) + \nu\left(2\left(\left(\frac{\partial U}{\partial X}\right)^2 + \left(\frac{\partial V}{\partial Y}\right)^2\right) + \left(\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y}\right)^2\right), \tag{6}$$

with boundary conditions

$$\begin{cases} U = 0, & \text{at } Y = H, \\ \dfrac{\partial U}{\partial Y} = 0, & \text{at } Y = 0, \\ T = T_c, & \text{at } Y = 0, \\ T = T_w, & \text{at } Y = H, \end{cases} \tag{7}$$

where U and V are the velocity components for X and Y directions, respectively, in the fixed frame of reference, p is the pressure, ρ is the density, ϑ is the kinematic coefficient of viscosity, c_p is the specific heat at constant pressure, T is the temperature of fluid, k is the thermal conductivity, T_c is the temperature at the centre of channel, T_w is the temperature at the wall, and α is the angle of inclination.

3. Analysis

The nondimensional parameters are as follows:

$$
\begin{aligned}
U' &= \frac{U}{c},\\
V' &= \frac{V}{c\delta},\\
X' &= \frac{X}{\lambda},\\
Y' &= \frac{Y}{a},\\
\theta_H &= \frac{T - T_c}{T_w - T_c},\\
\theta_T &= \frac{T - T_w}{T_c - T_w},\\
t' &= \frac{ct}{\lambda},\\
p' &= \frac{a^2}{c\mu\lambda}p,\\
\delta &= \frac{a}{\lambda},\\
\phi &= \frac{b}{a},\\
\mathrm{Pr} &= \frac{\mu C_p}{k},\\
\mathrm{Ec}_H &= \frac{c^2}{(T_w - T_c)C_p},\\
\phi_1 &= \frac{b_1}{a},\\
\mathrm{Re} &= \frac{\rho a c}{\mu},\\
\lambda_1' &= \frac{\lambda_1}{\lambda},\\
\mathrm{Ec}_T &= \frac{c^2}{(T_c - T_w)C_p},\\
\mathrm{Fr} &= \frac{c^2}{ga},
\end{aligned}
\tag{8}
$$

where Re is the Reynolds Number, δ is the wavenumber, θ_H is the dimensionless temperature in the case of uniform heat flux, θ_T is the dimensionless temperature in the case of uniform surface temperature, Ec_H is the Eckert number in the case of uniform heat flux, Ec_T is the Eckert number in the case of uniform surface temperature, and Pr is the Prandtl number.

Using equation (8) in equations (3)–(6) and dropping primes with the supposition of long wavelength and low Reynolds number,

$$\frac{\partial U}{\partial X} + \frac{\partial U}{\partial Y} = 0, \tag{9}$$

$$\frac{\partial^2 U}{\partial Y^2} + \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha = \frac{\partial p}{\partial X}, \tag{10}$$

$$\frac{\partial p}{\partial Y} = 0, \tag{11}$$

$$\frac{\partial^2 \theta}{\partial Y^2} + \mathrm{EcPr}\left(\frac{\partial U}{\partial Y}\right)^2 = 0, \tag{12}$$

with boundary conditions

$$
\begin{cases}
U = 0, & \text{at } Y = h = \dfrac{H}{a} = 1 + \dfrac{K\lambda X}{a} + \phi\sin(2\pi(X - t)) - \phi_1\cos^4\left(\dfrac{\pi X}{\lambda_1}\right),\\
\dfrac{\partial U}{\partial Y} = 0, & \text{at } Y = 0,\\
\theta_H = 0,\ \theta_T = 1, & \text{at } Y = 0,\\
\theta_H = 1,\ \theta_T = 0, & \text{at } Y = h,
\end{cases}
\tag{13}
$$

where K is the nonuniformity of the wall, ϕ is the amplitude ratio, ϕ_1 is the roughness parameter, and λ_1 is the dimensionless pitch.

On solving equations (9)–(12) with given boundary conditions equation (13), we get

$$U = \left(\frac{Y^2 - h^2}{2}\right)\left(\frac{dp}{dX} - \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha\right),$$

$$\theta_H = \mathrm{Ec}_H \mathrm{Pr}\left(\frac{dp}{dX} - \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha\right)^2\left(\frac{h^3 y - y^4}{12}\right) + \frac{y}{h}, \tag{14}$$

$$\theta_T = \mathrm{Ec}_T \mathrm{Pr}\left(\frac{dp}{dX} - \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha\right)^2\left(\frac{h^3 y - y^4}{12}\right) - \frac{y}{h} + 1.$$

The flow flux $q(X,t)$ is found as

$$q(X,t) = \int_0^h U\,dY = -\frac{h^3}{3}\left(\frac{dp}{dX} - \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha\right), \tag{15}$$

or

$$\frac{dp}{dX} = \frac{-3q(X,t)}{h^3} + \frac{\mathrm{Re}}{\mathrm{Fr}}\sin\alpha. \tag{16}$$

As the roughness parameter $\phi_1 \longrightarrow 0$ and angle of inclination $\alpha \longrightarrow 0$, the result obtained from equation (16) is the same as that found by Gupta and Seshadri [3].

The pressure rise Δp across the channel of L length in dimensionless forms is

$$\Delta p = \int_0^{L/\lambda}\left(\frac{dp}{dX}\right)dX. \tag{17}$$

4. Results and Discussion

In the present study, graphs have been drawn using MATHEMATICA software. The instantaneous flow rate $q(x, t)$, periodic in $(x - t)$, is as follows [3]:

$$q(X,t) = Q + \phi \sin 2\pi(X - t), \tag{18}$$

and magnitude of various parameters is taken as $a = 0.01$ cm, $L = \lambda = 10$ cm, $K = 0.5(a/L) = 0.0005$ [14, 15], and $\phi_1 = (b_1/a)$; the magnitude of b_1 is from [28].

Figure 2 depicts changes in the pressure rise as a function of time over one period of the wave, at different surface roughness parameters and channel inclinations, at a particular value of all other parameters like K, Q, λ_1, ϕ_1, Re, and Fr. It is obvious from Figure 2 that, during one period of the wave, an increase in the angle of inclination enhances the pressure rise. The effect of surface roughness is most significant at the time $t \approx 0.3$, which corresponds to the point where maximum occlusion occurs from the entry of the channel. It is seen that pressure rise increases with an increase in the roughness parameter for a particular choice of all other parameters. This can be interpreted that, with increasing surface roughness at any particular section, the cross section area will reduce due to which the pressure rise increases with increasing surface roughness at that section.

Figure 3 displays the changes in pressure rise against flow rate for various surface roughness parameters at different angles of inclinations, at the particular value of all other parameters $(K, t, \lambda_1, \phi,$ Re, and Fr). The magnitude of the pressure rise decreases linearly with an increase in flow rate for all magnitude of the surface roughness parameter and angle of inclination. In a noninclined channel, the maximum flow rate occurs up to $Q = 0.64$, and increase in inclination from 0 to $(\pi/2)$ enhances the flow rate. At any particular value of flow rate in-between $Q = 0$ and 0.64, pressure rise increases continuously; when the surface roughness parameter increases from $Q = 0.64$ to $Q = 1$, reverse impact on the value of pressure rise is seen with increasing surface roughness parameter. The pressure rise increases also with an increase in the angle of inclination from 0 to $(\pi/2)$. The pressure rise is maximum at no flow condition, and this value rises with the surface roughness parameter and angle of inclination. It can be observed that, in the peristaltic pumping zone, the maximum flow rate is obtained at zero pressure rise. From Figure 3, it is also observed that the magnitude of the maximum flow rate can be enhanced by increasing the angle of inclination for a particular choice of all other parameters.

The amplitude ratio has a significant impact on the pressure rise. Figure 4 shows that, at higher magnitudes of amplitude ratio $(\phi > 0.8)$, a rapid augment is visualised in the magnitude of pressure rise. With an increase in inclination from 0 to $(\pi/2)$, the pressure rise increases at any particular value of amplitude ratio.

Figures 5 and 6 depict the variation in pressure rise against flow rate at different values of Re and Fr, respectively, for different surface roughness parameters. It can be noticed from Figures 5 and 6 that the pressure rise increases with an increase in surface roughness parameter at zero flow rate, and this difference goes on reducing as flow rate increases.

From Figure 5, it is also observed that, for different values of surface roughness parameter, the magnitude of maximum pressure rise and maximum flow rate increases with an increase in Re. Again from Figure 6, it is seen that the value of maximum pressure rise and maximum flow rate decreases with an increase in Fr for different values of surface roughness parameter. Variation in flow rate and pressure rise with Fr has opposite nature to that of Re.

Figures 7(a) and 7(b) and Table 1 describe the effect of the angle of inclination on the temperature distribution for different values of surface roughness parameter; similarly, Figures 8(a) and 8(b), Table 2, Figures 9(a) and 9(b), and Table 3 describe the effect of Re and Fr on the temperature distribution at different values of surface roughness parameter. The entire analysis performed above has been carried out under two conditions, uniform heat flux and uniform surface temperature condition.

Figures 7(a) and 7(b) and Table 1 describe the effect of the angle of inclination on the temperature distribution for different values of surface roughness parameter; similarly, Figures 8(a) and 8(b), Table 2, Figures 9(a) and 9(b), and Table 3 describe the effect of Re and Fr on the temperature

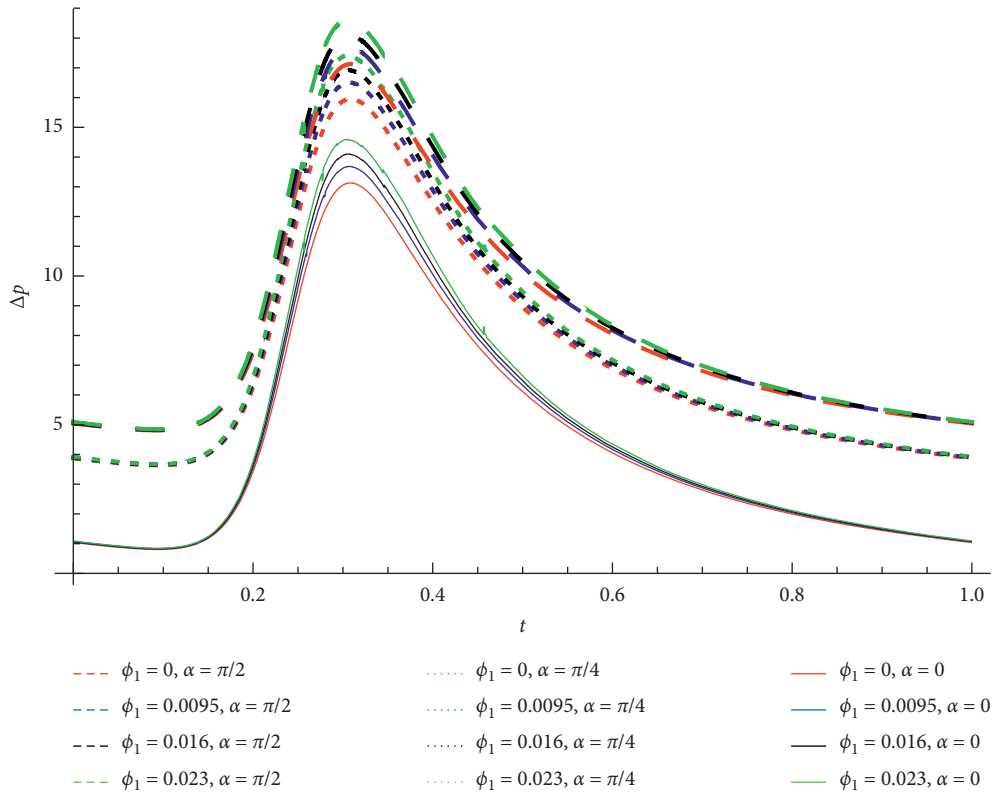


FIGURE 2: Pressure rise versus time for various values of ϕ_1, α at $K = 0.0005, Q = 0.3, \lambda_1 = 0.00011, \phi = 0.8$, Re $= 0.4$, and Fr $= 0.1$.

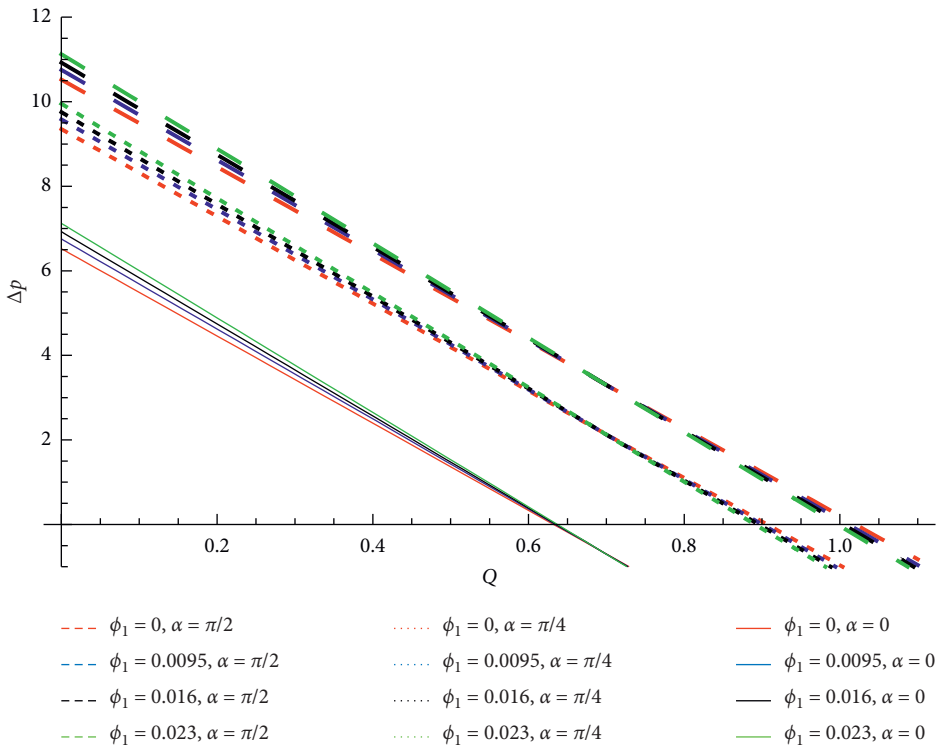


FIGURE 3: Pressure rise versus flow rate for different values of ϕ_1, α at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, \phi = 0.8$, Re $= 0.4$, and Fr $= 0.1$.

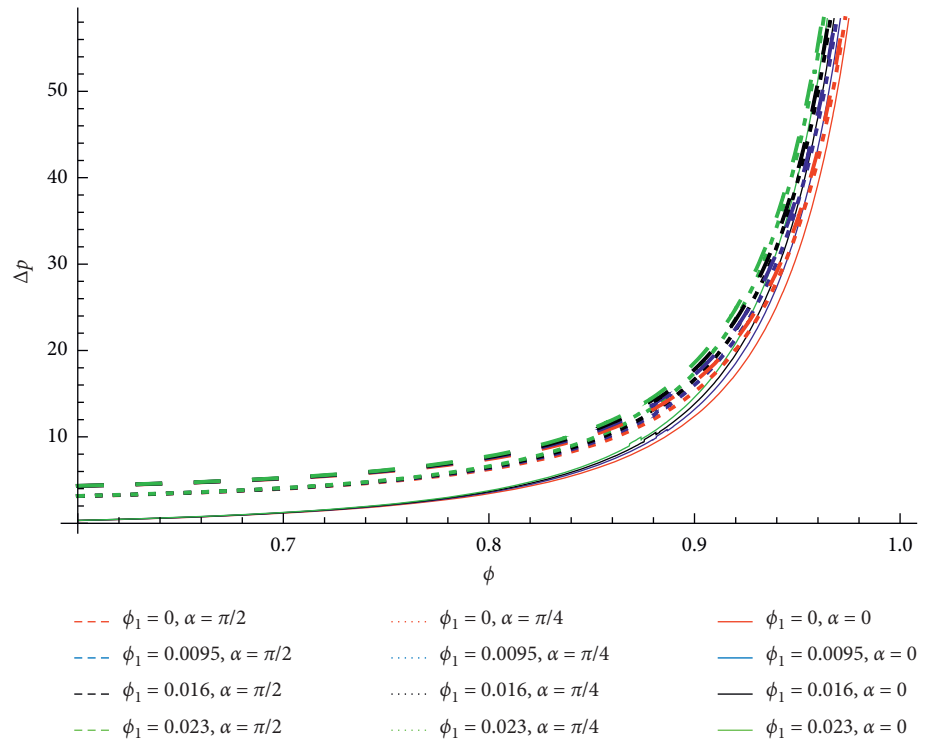


FIGURE 4: Pressure rise versus amplitude ratio for different values of ϕ_1, α at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, Q = 0.3, \text{Re} = 0.4,$ and $\text{Fr} = 0.1$.

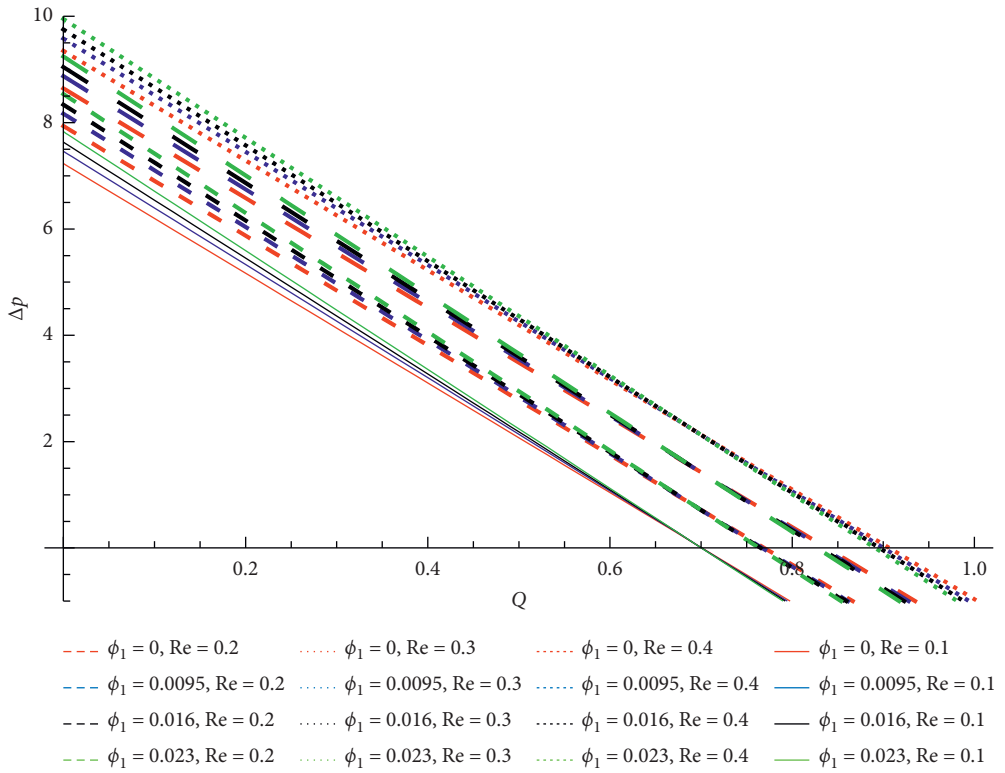


FIGURE 5: Pressure rise versus flow rate for different values of ϕ_1, Re at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, \phi = 0.8, \alpha = (\pi/4),$ and $\text{Fr} = 0.1$.

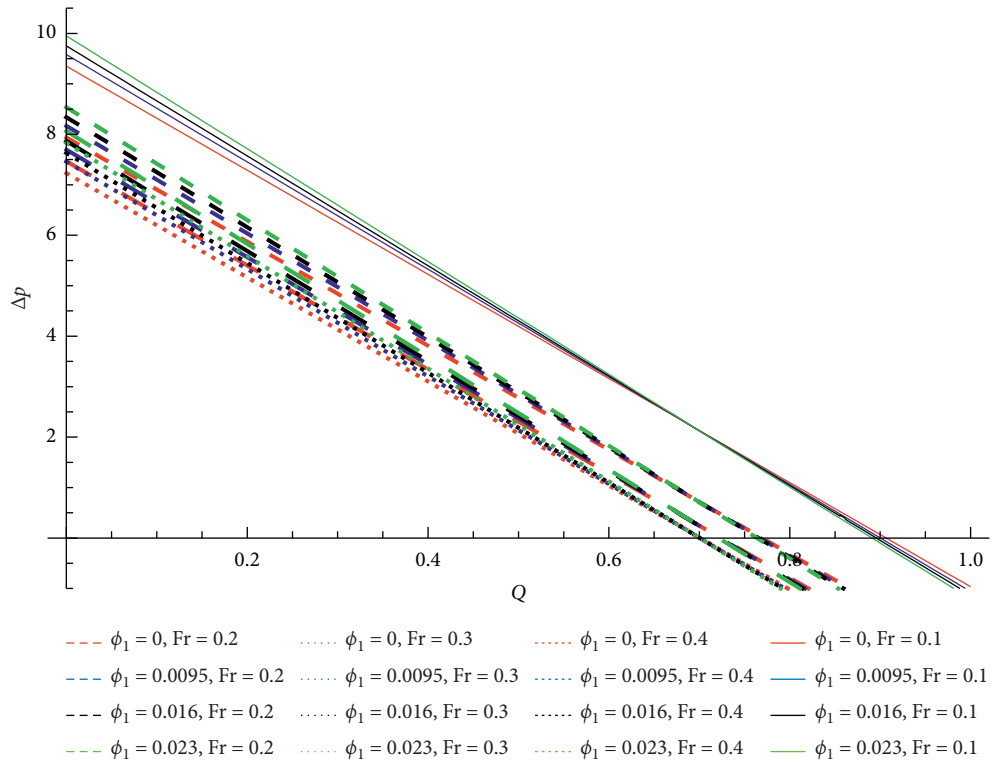


FIGURE 6: Pressure rise versus flow rate for different values of ϕ_1, Fr at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, \phi = 0.8, \alpha = (\pi/4)$, and Re = 0.4.

distribution at different values of surface roughness parameter. The entire analysis performed above has been carried out under two conditions: uniform heat flux and uniform surface temperature condition.

Figure 7(a) shows the variation in θ_H against y for several values of channel inclination and surface roughness. As y increases, θ_H increases from a minimum value at the centre of the channel to a maximum value at the channel wall, for a channel having zero inclination. For a noninclined channel, it is visible from Table 1 that θ_H increases and θ_T decreases at any particular value of y, as the surface roughness parameter is increased. It is also concluded from Table 1 that, for a particular value of surface roughness parameter in a noninclined channel, when y increases, θ_H increases and θ_T decreases.

With the variation in θ_H against y, for channels inclined at $(\pi/4)$ and $(\pi/2)$ inclination, initially θ_H increases rapidly to a maximum value after which it starts decreasing; finally, it coincides with the maximum value of θ_H for a channel having zero inclination, at the wall of the channel. Furthermore, for channels inclined at $(\pi/4)$ and $(\pi/2)$, from Table 1, it is found that the value θ_H and θ_T decreases at any particular value of y, as the value of surface roughness parameter increases. For any particular value of surface roughness, with an increase in y, θ_H increases and θ_T decreases for a channel having $(\pi/4)$ inclination. It is also clear from Table 1 that, for any particular value of surface roughness, θ_H increases as y increases from 0 to 0.8, after which it starts decreasing and θ_T increases as y increases from 0 to 0.4, after which it starts decreasing for a channel having $(\pi/2)$ inclination.

Figure 7(b) shows the variation in θ_T against y for different values of channel inclination and surface roughness. The value of θ_T decreases with increase in y and at $y = 0$; θ_T is maximum, for channels having zero and $(\pi/4)$ inclination. But in the case of channels having $(\pi/2)$ inclination, with increase in y, θ_T increases up to a maximum value, and then, it decreases to zero at the channel wall.

Figures 8(a) and 8(b) show the temperature profile for different values of Re and surface roughness parameters for both boundary conditions mentioned above, i.e., uniform heat flux and uniform surface temperature condition. The above is also described numerically through data presented in Table 2. Figures 9(a) and 9(b) show the temperature profile for different values of Fr and surface roughness parameters, and the same is described numerically through the data presented in Table 3. From the above two figures, it is clear that temperature variation at different values of Re and Fr follows a similar pattern to that depicted in Figures 7(a) and 7(b) except in the case of a channel having $(\pi/2)$ inclination, wherein θ_T initially increases and then it decreases. The variation in θ_T for different values of Re and Fr is similar in nature; θ_T decreases with increasing y for all values of Re and Fr. From Table 2, it is clear that, at Re = 0.1 and 0.2, for any particular value of y, θ_H goes on increasing, while θ_T goes on decreasing upon increasing surface roughness parameter. As the value of the surface roughness parameter is incremented, at Re = 0.3, for a particular value of y, θ_H decreases for all values of y from 0 to 0.8, and θ_T also decreases for all values of y. Excluding the value of θ_H at $y = 0.4$, the variation in θ_H and θ_T, for any particular value of y, with increasing values of surface roughness parameter

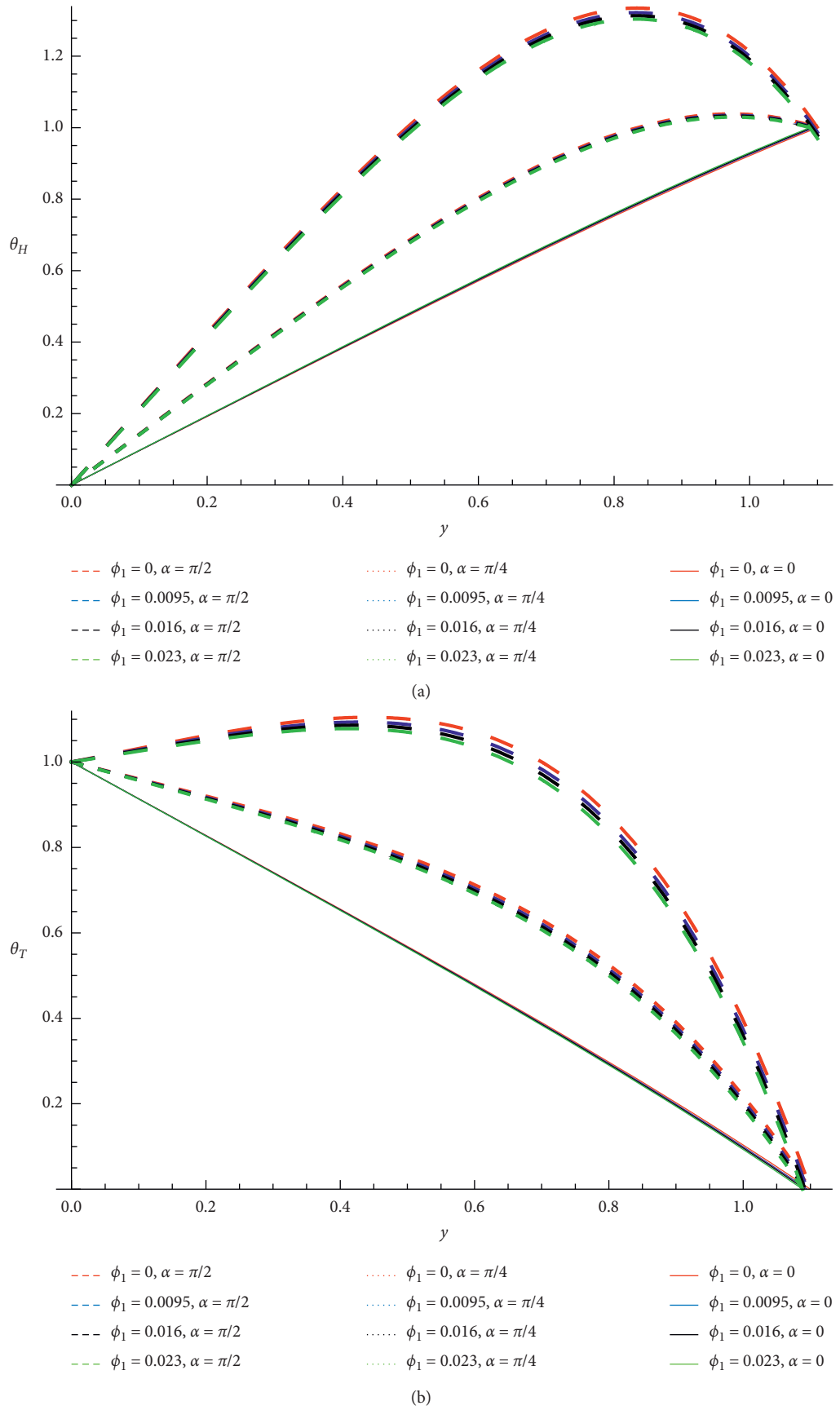
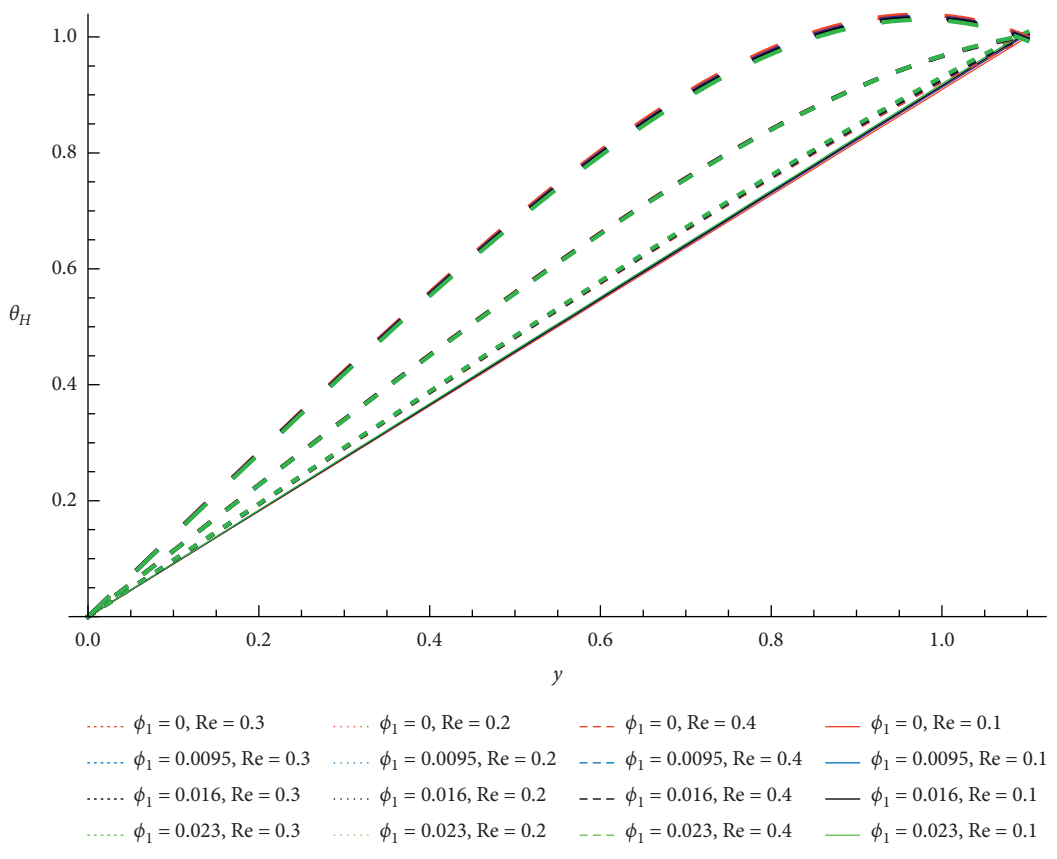


FIGURE 7: Temperature vs y for different values of ϕ_1, α at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, x = 0.2, \phi = 0.8$, Re $= 0.4$, Fr $= 0.1$, $Q = 0.3$, Pr $= 1$, $\mathrm{Ec}_H = 1$, and $\mathrm{Ec}_T = 1$: (a) uniform heat flux condition; (b) uniform surface temperature.

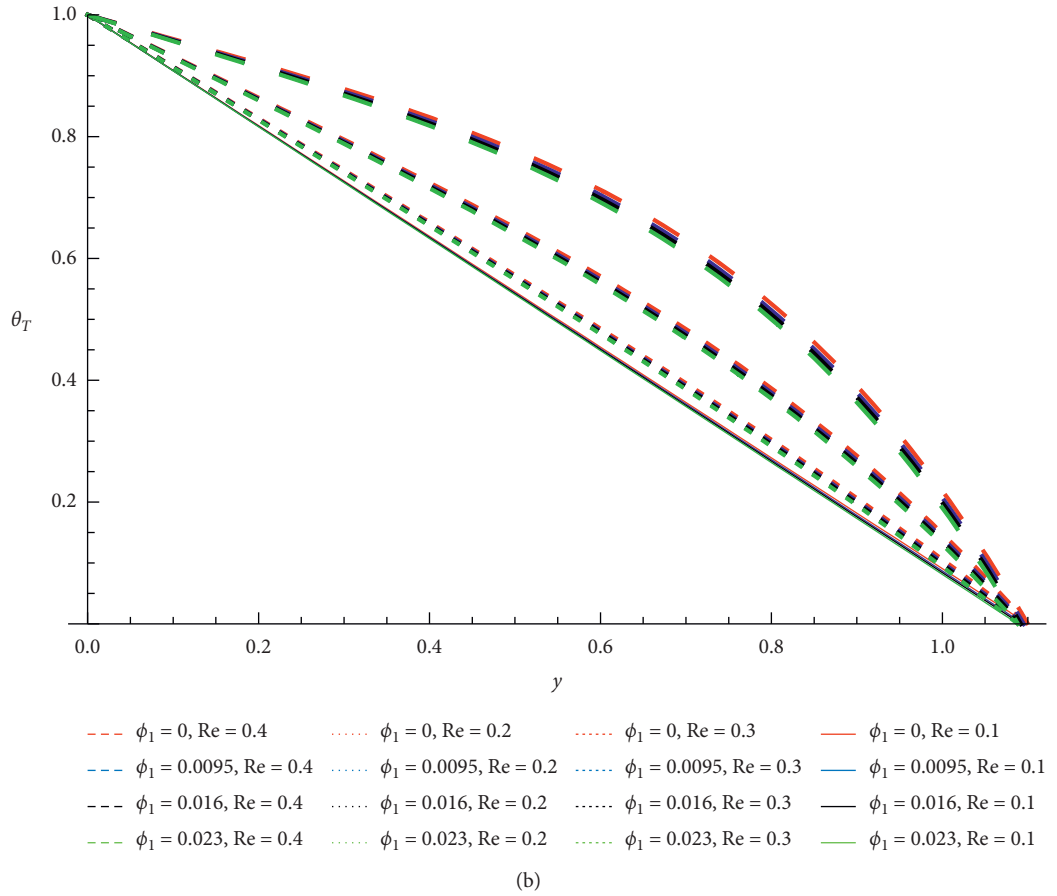
Table 1: Temperature distribution at different values of angle of inclination and surface roughness parameter.

	θ_T				θ_H			
	$\alpha = 0$				$\alpha = 0$			
y	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$
0	1	1	1	1	0	0	0	0
0.2	0.8282	0.8276	0.8271	0.8266	0.1919	0.1928	0.1934	0.1941
0.4	0.6556	0.6543	0.6534	0.6524	0.3829	0.3847	0.3860	0.3873
0.6	0.4800	0.4779	0.4764	0.4749	0.5709	0.5735	0.5754	0.5773
0.8	0.2976	0.2946	0.2925	0.2902	0.7522	0.7555	0.7577	0.7602
1.0	0.1035	0.0992	0.0962	0.0930	0.9217	0.9253	0.9277	0.9304
	$\alpha = (\pi/4)$				$\alpha = (\pi/4)$			
0	1	1	1	1	0	0	0	0
0.2	0.9203	0.9173	0.9153	0.9132	0.2839	0.2825	0.2816	0.2807
0.4	0.8319	0.8261	0.8222	0.8179	0.5592	0.5566	0.5548	0.5529
0.6	0.7127	0.7043	0.6986	0.6924	0.8036	0.7999	0.7975	0.7948
0.8	0.5256	0.5151	0.5080	0.5003	0.9801	0.9760	0.9732	0.9702
1.0	0.2186	0.2070	0.1992	0.1908	1.0368	1.0331	1.0307	1.0282
	$\alpha = (\pi/2)$				$\alpha = (\pi/2)$			
0	1	1	1	1	0	0	0	0
0.2	1.0617	1.0565	1.0529	1.0491	0.4254	0.4217	0.4192	0.4166
0.4	1.1029	1.0925	1.0855	1.0779	0.8302	0.8230	0.8181	0.8129
0.6	1.0704	1.0553	1.0451	1.0341	1.1613	1.1510	1.1440	1.1365
0.8	0.8759	0.8569	0.8440	0.8303	1.3304	1.3178	1.3093	1.3002
1.0	0.3956	0.3743	0.3598	0.3444	1.2138	1.2004	1.1914	1.1818



(a)

Figure 8: Continued.



(b)

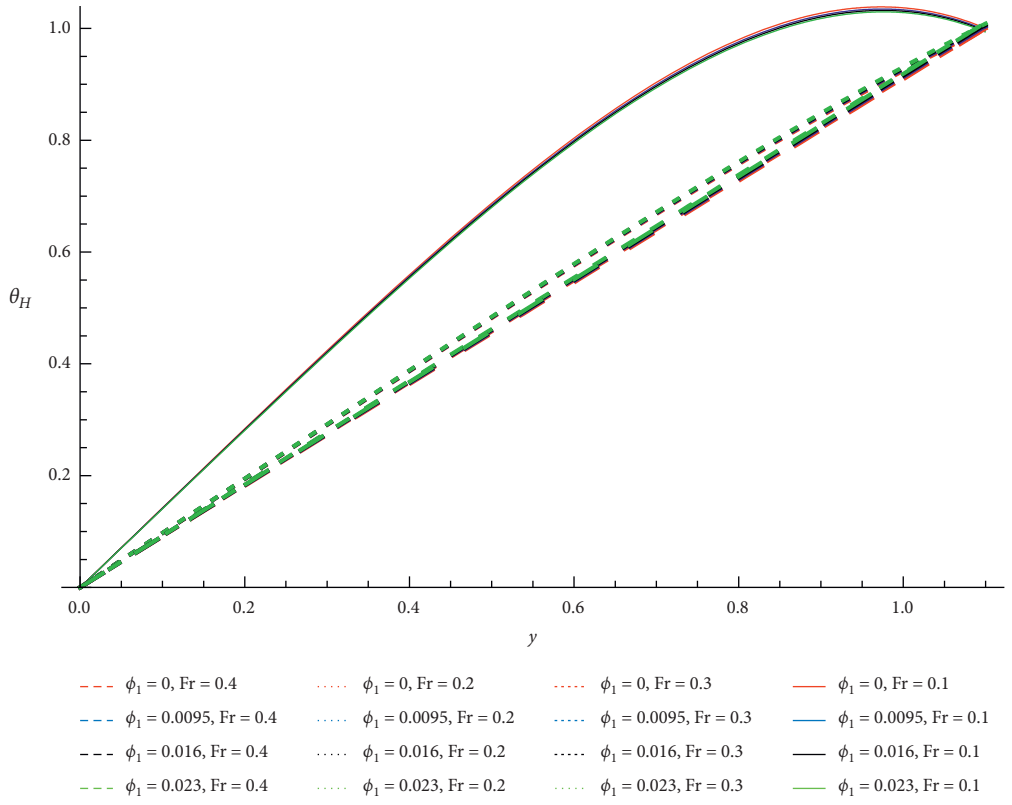
FIGURE 8: Temperature vs y for different values of ϕ_1, Re at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, x = 0.2, \phi = 0.8$, $\mathrm{Fr} = 0.1, Q = 0.3, \alpha = (\pi/4)\mathrm{Pr} = 1, \mathrm{Ec}_H = 1$, and $\mathrm{Ec}_T = 1$: (a) uniform heat flux condition; (b) uniform surface temperature.

TABLE 2: Temperature distribution at different values of Re and surface roughness parameter.

	θ_T				θ_H			
	Re = 0.1				Re = 0.1			
y	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$
0	1	1	1	1	0	0	0	0
0.2	0.8182	0.8174	0.8168	0.8162	0.1818	0.1826	0.1831	0.1837
0.4	0.6364	0.6348	0.6337	0.6325	0.3636	0.3652	0.3663	0.3674
0.6	0.4545	0.4522	0.4505	0.4487	0.5455	0.5478	0.5494	0.5512
0.8	0.2727	0.2695	0.2674	0.2650	0.7273	0.7304	0.7326	0.7449
1.0	0.0909	0.0869	0.0842	0.0812	0.9091	0.9130	0.9157	0.9187
	Re = 0.2				Re = 0.2			
0	1	1	1	1	0	0	0	0
0.2	0.8301	0.8289	0.8281	0.8272	0.1938	0.1941	0.1944	0.1947
0.4	0.6593	0.6569	0.6552	0.6534	0.3866	0.3873	0.3878	0.3884
0.6	0.4849	0.4813	0.4789	0.4763	0.5758	0.5770	0.5778	0.5787
0.8	0.3024	0.2979	0.2949	0.2916	0.7570	0.7588	0.7601	0.7615
1.0	0.1059	0.1008	0.0973	0.0936	0.9241	0.9269	0.9289	0.9310
	Re = 0.3				Re = 0.3			
0	1	1	1	1	0	0	0	0
0.2	0.8642	0.8622	0.8609	0.8595	0.2278	0.2275	0.2272	0.2270
0.4	0.7245	0.7207	0.7181	0.7152	0.4518	0.4511	0.4507	0.4502
0.6	0.5709	0.5654	0.5616	0.5575	0.6618	0.6610	0.6605	0.6599
0.8	0.3867	0.3798	0.3751	0.3700	0.8413	0.8407	0.8403	0.8400
1.0	0.1485	0.1409	0.1357	0.1301	0.9666	0.9669	0.9672	0.9676

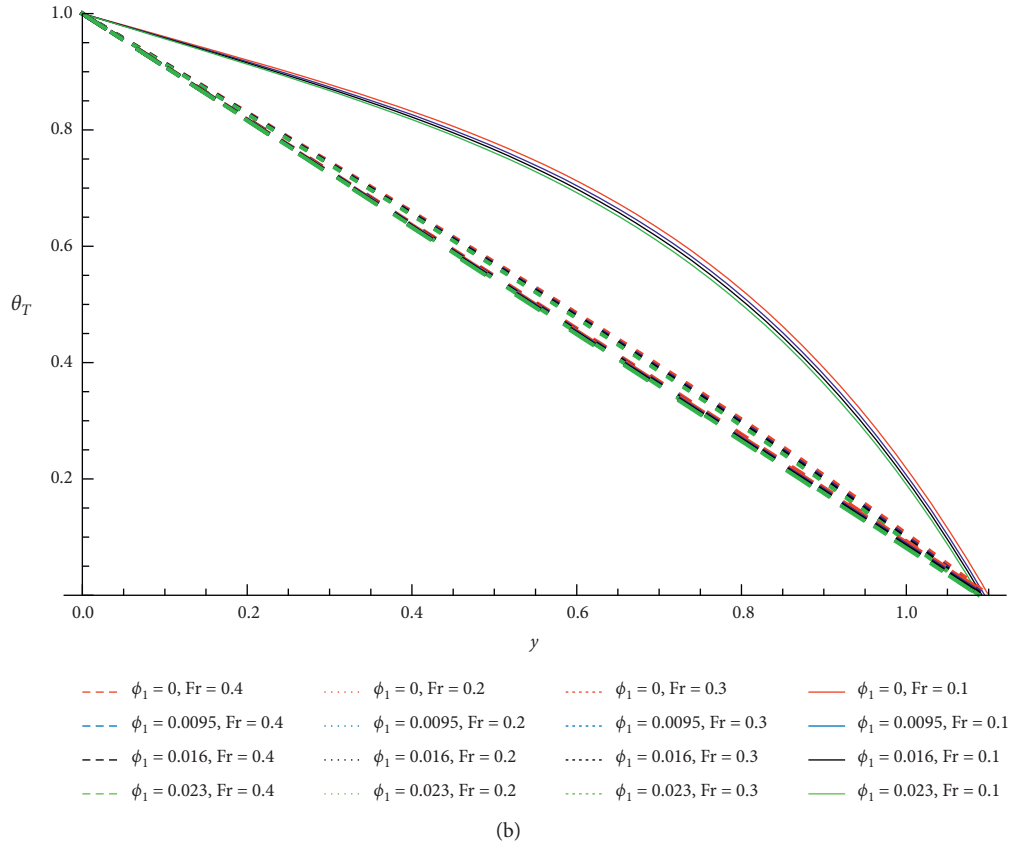
Table 2: Continued.

θ_T					θ_H			
Re = 0.4					Re = 0.4			
0	1	1	1	1	0	0	0	0
0.2	0.9203	0.9173	0.9153	0.9132	0.2839	0.2825	0.2816	0.2807
0.4	0.8319	0.8261	0.8222	0.8179	0.5592	0.5566	0.5448	0.5529
0.6	0.7127	0.7043	0.6946	0.6924	0.8036	0.7999	0.7975	0.7948
0.8	0.5256	0.5151	0.508	0.5003	0.9801	0.9760	0.9732	0.9702
1.0	0.2186	0.2070	0.1992	0.1908	1.0368	1.0331	1.0307	1.0282



(a)

Figure 9: Continued.



(b)

FIGURE 9: Temperature vs y for different values of ϕ_1, Fr at $K = 0.0005, t = 0.2, \lambda_1 = 0.00011, x = 0.2, \phi = 0.8$, Re $= 0.4$, $Q = 0.3, \alpha = (\pi/4)$Pr $= 1$, $\text{Ec}_H = 1$, and $\text{Ec}_T = 1$: (a) uniform heat flux condition; (b) uniform surface temperature.

TABLE 3: Temperature distribution at different values of Fr and surface roughness parameter.

	θ_T				θ_H			
	Fr = 0.1				Fr = 0.1			
y	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$	$\phi_1 = 0$	$\phi_1 = 0.0095$	$\phi_1 = 0.016$	$\phi_1 = 0.023$
0	1	1	1	1	0	0	0	0
0.2	0.9203	0.9173	0.9153	0.9132	0.2839	0.2825	0.2816	0.2807
0.4	0.8319	0.8261	0.8222	0.8179	0.5592	0.5566	0.5548	0.5529
0.6	0.7127	0.7043	0.6986	0.6924	0.8036	0.7999	0.7975	0.7948
0.8	0.5256	0.5151	0.508	0.5003	0.9801	0.9760	0.9732	0.9702
1.0	0.2186	0.2070	0.1992	0.1908	1.0368	1.0331	1.0307	1.0282
	Fr = 0.2				Fr = 0.2			
0	1	1	1	1	0	0	0	0
0.2	0.8301	0.8289	0.8281	0.8272	0.1938	0.1941	0.1944	0.1947
0.4	0.6593	0.6569	0.6552	0.6534	0.3866	0.3873	0.3878	0.3884
0.6	0.4849	0.4813	0.4789	0.4763	0.5758	0.5770	0.5778	0.5787
0.8	0.3024	0.2979	0.2949	0.2916	0.7570	0.7588	0.7601	0.7615
1.0	0.1059	0.1008	0.0973	0.0936	0.9241	0.9269	0.9289	0.9310
	Fr = 0.3				Fr = 0.3			
0	1	1	1	1	0	0	0	0
0.2	0.8197	0.8188	0.8182	0.8175	0.1833	0.1840	0.1845	0.1850
0.4	0.6393	0.6375	0.6363	0.6349	0.3666	0.3679	0.3689	0.3699
0.6	0.4585	0.4558	0.4539	0.4519	0.5494	0.5514	0.5528	0.5544
0.8	0.2766	0.2731	0.2707	0.2681	0.7311	0.7339	0.7359	0.7380
1.0	0.0928	0.0886	0.0858	0.0827	0.9110	0.9147	0.9173	0.9201

Table 3: Continued.

	θ_T				θ_H			
	Fr = 0.4				Fr = 0.4			
0	1	1	1	1	0	0	0	0
0.2	0.8182	0.8174	0.8168	0.8162	0.1818	0.1826	0.1831	0.1837
0.4	0.6364	0.6348	0.6337	0.6325	0.3636	0.3652	0.3663	0.3674
0.6	0.4545	0.4522	0.4505	0.4487	0.5455	0.5478	0.5494	0.5512
0.8	0.2727	0.2695	0.2674	0.2650	0.7273	0.7304	0.7326	0.7349
1.0	0.0909	0.0869	0.0842	0.0812	0.9091	0.9130	0.9157	0.9187

at Re = 0.4 depicts a similar nature to that observed for Re = 0.3. For all values of Re at any given value of surface roughness parameter, as y increases, θ_H increases while θ_T decreases.

From Table 3 it is found that, except at Fr = 0.1 for all other values of Fr, for any particular value of y, θ_H increases and θ_T decreases upon incrementing the value of surface roughness parameter. At Fr = 0.1, for any particular value of y, with an increase in surface roughness, both θ_H and θ_T decrease. For a particular choice of surface roughness parameter, as y increases, θ_H increases and θ_T decreases for all values of Fr.

It can be seen that temperature profiles show a reverse behaviour at uniform surface temperature and uniform heat flux condition. For better understanding, the effect of surface roughness on temperature for various values of channel inclination Re and Fr is shown below in the tabular form.

5. Conclusion

The objective of this analysis is to explain the variation occurring in heat transfer due to the inner wall roughness of a nonlinear inclined channel, for transportation of Newtonian Fluid through the peristaltic mechanism. From the literature survey conducted during the study, it has been found that, so far, no attempt has been made by researchers to address the above problem. The mathematical results obtained in the present study may apply to the transportation of liquids used in various industries. This study may be useful for analysis of heat conduction in many body organs in which physiological fluid transported. The conclusion drawn from the analysis is as follows:

(1) An increase in the maximum magnitude of pressure rise is observed when the surface roughness parameter increased.

(2) The effect of surface roughness on pressure rise is majorly seen at the time $t \approx 0.3$, which corresponds to the point where maximum occlusion occurs.

(3) The maximum pressure rise value increases, as the angle of inclination is increased from 0 to $(\pi/2)$.

(4) Pressure rise increases at all points for the period of a wave as the inclination of the channel is increased from 0 to $(\pi/2)$.

(5) For a particular choice of pressure rise value, the flow rate increases with angle of inclination from 0 to $(\pi/2)$.

(6) The significant effect of the angle of the inclination of the channel has been observed in the region $0.6 < \Phi < 0.94$.

(7) Maximum flow rate increases with an inclination of the channel from 0 to $(\pi/2)$.

(8) The value of maximum pressure rise and flow rate rises when Reynolds number is augmented.

(9) For any value of Reynolds number, the maximum pressure attains higher values when the surface roughness parameter is increased.

(10) The maximum pressure rise and maximum flow rate decrease as the Froude number is increased.

(11) For any particular Froude number, maximum pressure rise attains a higher magnitude at a higher surface roughness parameter.

(12) The temperature profiles show opposite behaviours at uniform heat flux conditions to that of uniform surface temperature conditions.

(13) Excluding the value of θ_H at zero inclination, the value of θ_H and θ_T decreases with an increase in surface roughness, at all angles of inclination for any particular value of y.

Nomenclature

H: Vertical displacement
b: Amplitude of the wave
a: Inlet half channel width
K: Nonuniformity parameter
c: Wave velocity
t: Time
λ: Wavelength
b_1: Roughness height
λ_1: Pitch
X: Axial variable
Y: Transverse variable
U: Axial velocity
V: Transverse velocity
ρ: Density
p: Pressure
μ: Coefficient of viscosity
g: Gravitational acceleration
α: Angle of inclination of channel
ϑ: Kinematic coefficient of viscosity
c_p: Specific heat at constant pressure
k: Thermal conductivity

T: Temperature of fluid
T_w: Temperature at the wall
T_c: Temperature at centre of the channel
θ_H: Dimensionless temperature in case of uniform heat flux
θ_T: Dimensionless temperature in case of uniform surface temperature
Ec_H: Eckert number in case of uniform heat flux
Ec_T: Eckert number in case of uniform surface temperature
Pr: Prandtl number
h: Dimensionless vertical displacement
t': Dimensionless time
ϕ_1: Roughness parameter
ϕ: Amplitude ratio
X': Dimensionless axial variable
Y': Dimensionless transverse variable
U': Dimensionless axial velocity
V': Dimensionless transverse velocity
p': Dimensionless pressure
Re: Reynolds number
Fr: Froude number
δ: Wavenumber
Δp: Pressure rise
F: Friction force at the wall of channel
L: Length of channel
q: Instantaneous flow rate
Q: Time mean flow rate.

Data Availability

The data used to support the findings of this study are available from the first author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] T. W. Latham, "*Fluid motion in a peristaltic pump*," M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1966.

[2] A. H. Shapiro, M. Y. Jaffrin, and S. L. Weinberg, "Peristaltic pumping with long wavelengths at low Reynolds number," *Journal of Fluid Mechanics*, vol. 37, no. 4, pp. 799–825, 1969.

[3] B. B. Gupta and V. Seshadri, "Peristaltic pumping in non-uniform tubes," *Journal of Biomechanics*, vol. 9, no. 2, pp. 105–109, 1976.

[4] G. Radhakrishnamacharya and V. R. Murty, "Heat transfer to peristaltic in a non uniform channel," *Defence Science Journal*, vol. 43, no. 3, pp. 275–280, 1993.

[5] S. R. El Koumy, E. S. I. Barakat, and S. I. Abdelsalam, "Hall and porous boundaries effects on peristaltic transport through porous medium of a Maxwell model," *Transport in Porous Media*, vol. 94, no. 3, pp. 643–658, 2012.

[6] K. S. Mekheimer, S. R. Komy, and S. I. Abdelsalam, "Simultaneous effects of magnetic field and space porosity on compressible Maxwell fluid transport induced by a surface acoustic wave in a microchannel," *Chinese Physics B*, vol. 22, no. 12, Article ID 124702, 2013.

[7] Y. Abd Elmaboud, S. I. Abdelsalam, and K. S. Mekheimer, "Couple stress fluid flow in a rotating channel with peristalsis," *Journal of Hydrodynamics*, vol. 30, no. 2, pp. 307–316, 2018.

[8] D. Tripathi and O. Anwar Bég, "Peristaltic propulsion of generalized burgers' fluids through a non-uniform porous medium: a study of chyme dynamics through the diseased intestine," *Mathematical Biosciences*, vol. 248, pp. 67–77, 2014.

[9] S. I. Abdelsalam and K. Vafai, "Combined effects of magnetic field and rheological properties on the peristaltic flow of a compressible fluid in a microfluidic channel," *European Journal of Mechanics—B/Fluids*, vol. 65, pp. 398–411, 2017.

[10] S. I. Abdelsalam and M. M. Bhatti, "The study of non-Newtonian nanofluid with hall and ion slip effects on peristaltically induced motion in a non-uniform channel," *RSC Advances*, vol. 8, no. 15, pp. 7904–7915, 2018.

[11] I. M. Eldesoky, S. I. Abdelsalam, W. A. El-Askary, and M. M. Ahmed, "Concurrent development of thermal energy with magnetic field on a particle-fluid suspension through a porous conduit," *BioNanoScience*, vol. 9, no. 1, pp. 186–202, 2019.

[12] S. A. Victor and V. L. Shah, "Heat transfer to blood flowing in a tube," *Biorheology*, vol. 12, no. 6, pp. 361–368, 1975.

[13] S. A. Victor and V. L. Shah, "Steady state heat transfer to blood flowing in the entrance region of a tube," *International Journal of Heat and Mass Transfer*, vol. 19, no. 7, pp. 777–783, 1976.

[14] L. M. Srivastava and V. P. Srivastava, "Peristaltic transport of blood: casson model-II," *Journal of Biomechanics*, vol. 17, no. 11, pp. 821–829, 1984.

[15] K. S. Mekheimer, "Peristaltic flow of blood under effect of a magnetic field in a non-uniform channels," *Applied Mathematics and Computation*, vol. 153, no. 3, pp. 763–777, 2004.

[16] S. I. Abdelsalam and K. Vafai, "Particulate suspension effect on peristaltically induced unsteady pulsatile flow in a narrow artery: blood flow model," *Mathematical Biosciences*, vol. 283, pp. 91–105, 2017.

[17] Y. Abd Elmaboud, S. I. Abdelsalam, K. S. Mekheimer, and K. Vafai, "Electromagnetic flow for two-layer immiscible fluids," *Engineering Science and Technology, an International Journal*, vol. 22, no. 1, pp. 237–248, 2019.

[18] S. I. Abdelsalam, M. M. Bhatti, A. Zeeshan, A. Riaz, and O. A. Bég, "Metachronal propulsion of a magnetised particle-fluid suspension in a ciliated channel with heat and mass transfer," *Physica Scripta*, vol. 94, no. 11, p. 115301, 2019.

[19] M. Sohail, R. Naz, and S. I. Abdelsalam, "On the onset of entropy generation for a nanofluid with thermal radiation and gyrotactic microorganisms through 3D flows," *Physica Scripta*, vol. 95, no. 4, Article ID 045206, 2020.

[20] S. I. Abdelsalam and M. Sohail, "Numerical approach of variable thermophysical features of dissipated viscous nanofluid comprising gyrotactic micro-organisms," *Pramana*, vol. 94, no. 1, p. 67, 2020.

[21] M. Sohail, R. Naz, and S. I. Abdelsalam, "Application of non-Fourier double diffusions theories to the boundary-layer flow of a yield stress exhibiting fluid model," *Physica A: Statistical Mechanics and Its Applications*, vol. 537, Article ID 122753, 2020.

[22] H. Sadaf and S. I. Abdelsalam, "Adverse effects of a hybrid nanofluid in a wavy non-uniform annulus with convective boundary conditions," *RSC Advances*, vol. 10, no. 26, pp. 15035–15043, 2020.

[23] S. I. Abdelsalam and M. M. Bhatti, "The impact of impinging TiO_2 nanoparticles in Prandtl nanofluid along with endoscopic and variable magnetic field effects on peristaltic blood flow," *Multidiscipline Modeling in Materials and Structures*, vol. 14, no. 3, pp. 530–548, 2018.

[24] S. I. Abdelsalam and M. M. Bhatti, "New insight into AuNP applications in tumour treatment and cosmetics through wavy annuli at the nanoscale," *Scientific Reports*, vol. 9, p. 260, 2019.

[25] S. I. Abdelsalam and M. M. Bhatti, "Anomalous reactivity of thermo-bioconvective nanofluid towards oxytactic microorganisms," *Applied Mathematics and Mechanics*, vol. 41, no. 5, pp. 711–724, 2020.

[26] I. M. Eldesoky, S. I. Abdelsalam, W. A. El-Askary, A. M. El-Refaey, and M. M. Ahmed, "Joint effect of magnetic field and heat transfer on particulate fluid suspension in a catheterized wavy tube," *BioNanoScience*, vol. 9, pp. 723–739, 2019.

[27] S. W. Park, M. Intaglietta, and D. M. Tartakovsky, "Impact of endothelium roughness on blood flow," *Journal of Theoretical Biology*, vol. 300, pp. 152–160, 2012.

[28] H. E. Burton and D. M. Espino, "The effect of mechanical overloading on surface roughness of the coronary arteries," *Applied Bionics and Biomechanics*, vol. 2019, Article ID 2784172, 8 pages, 2019.

[29] M. M. Bhatti, M. Marin, A. Zeeshan, R. Ellahi, and S. I. Abdelsalam, "Swimming of motile gyrotactic microorganisms and nanoparticles in blood flow through anisotropically tapered arteries," *Frontiers in Physics*, vol. 8, 2020.

[30] R. W. Veatch, *The Effect of Surface Roughness on Fluid Flow through Narrow Rectangular Passages*, Society of Petroleum Engineers, Richardson, TX, USA, 1965.

[31] J. C. Han, Y. M. Zhang, and C. P. Lee, "Augmented heat transfer in square channels with parallel, crossed, and V-shaped angled ribs," *Journal of Heat Transfer*, vol. 113, no. 3, pp. 590–596, 1991.

[32] M. E. Taslim, T. Li, and D. M. Kercher, "Experimental heat transfer and friction in channels roughened with angled, V-shaped and discrete ribs on two opposite walls," in *Proceedings of the International Gas Turbine and Aeroengine Congress and Exposition*, pp. 13–16, Hague, Netherlands, June 1994.

[33] J. B. Taylor, A. L. Carrano, and S. G. Kandlikar, "Characterization of the effect of surface roughness and texture on fluid flow-past, present, and future," *International Journal of Thermal Sciences*, vol. 45, no. 10, pp. 962–968, 2006.

[34] R. N. Wagner and S. G. Kandlikar, "Effects of structured roughness on fluid flow at the microscale level," *Heat Transfer Engineering*, vol. 33, no. 6, pp. 483–493, 2012.

[35] V. V. Dharaiya and S. G. Kandlikar, "A numerical study on the effects of 2D structured sinusoidal elements on fluid flow and heat transfer at microscale," *International Journal of Heat and Mass Transfer*, vol. 57, no. 1, pp. 190–201, 2013.

[36] M. Kharati-Koopaee and M. Zare, "Effect of aligned and offset roughness patterns on the fluid flow and heat transfer within microchannels consist of sinusoidal structured roughness," *International Journal of Thermal Sciences*, vol. 90, pp. 9–23, 2015.

[37] M. Attalla, H. M. Maghrabie, and E. Specht, "An experimental investigation on fluid flow and heat transfer of rough minichannel with rectangular cross section," *Experimental Thermal and Fluid Science*, vol. 75, pp. 199–210, 2016.