

## Research Article

# Multistep-Ahead Prediction of Urban Traffic Flow Using *GaTS* Model

Benchao Wang,<sup>1,2</sup> Pan Qin ,<sup>1</sup> and Hong Gu<sup>1</sup>

<sup>1</sup>Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

<sup>2</sup>Department of Administration and Supervision, Liaoning Police College, Dalian 116036, China

Correspondence should be addressed to Pan Qin; qp112cn@dlut.edu.cn

Received 27 April 2020; Revised 3 June 2020; Accepted 24 June 2020; Published 17 July 2020

Academic Editor: Wei Wang

Copyright © 2020 Benchao Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mathematical models for traffic flow have been widely investigated for a lot of application, like planning transportation and easing traffic pressure by using statistics and machine learning methods. However, there remains a lot of challenging problems for various reasons. In this research, we mainly focused on three issues: (a) the data of traffic flow are nonnegative, and hereby, finding a proper probability distribution is essential; (b) the complex stochastic property of the traffic flow leads to the nonstationary variance, i.e., heteroscedasticity; and (c) the multistep-ahead prediction of the traffic flow is often of poor performance. To this end, we developed a Gamma distribution-based time series (*GaTS*) model. First, we transformed the original traffic flow observations into nonnegative real-valued data by using the Box-Cox transformation. Then, by specifying the generalized linear model with the Gamma distribution, the mean and variance of the distribution are regressed by the past data and homochronous terms, respectively. A Bayesian information criterion is used to select the proper Box-Cox transformation coefficients and the optimal model structures. Finally, the proposed model is applied to the urban traffic flow data achieved from Dalian city in China. The results show that the proposed *GaTS* has an excellent prediction performance and can represent the nonstationary stochastic property well.

## 1. Introduction

As the main driving force of development, traffic has significant effects on the flow of production factors and the daily life of the urban system. The intelligent transportation system (ITS) can effectively provide innovative services relating to different modes of transport and traffic management [1, 2], such as transportation planning [3], traffic pressure easing [4], and traffic accident evaluation [5]. It enables transport networks to be more informed, more coordinated, and more efficient for various users. ITS requires a reliable prediction of traffic information in real time. Thus, how to accurately and timely predict traffic is a challenging task, which has gained more and more attention.

The traffic flow is full of complex dynamics and is stochastic [6], which make analysis and prediction mainly depend on timely or historical traffic data. As we illustrate in the latter, the urban traffic flow data is a nonstationary

stochastic process with heterogeneous variance. Thus, the time series models are preferred for the prediction of traffic flow. On the other hand, the control operations, like variable speed limits (VSLs), are always embedded into ITS [7]. This fact suggests that the prediction models in ITS should be of concise structures natural to conduct the control and operation design. Thus, our studies focus on developing the data-driven time series model, which can predict the nonstationary distribution of the traffic flow and is of concise structures for the stochastic control design for ITS.

To construct the statistical model for the nonnegative traffic flow data, we first investigate which probability distribution is feasible for describing the uncertainty of the traffic flow. By detecting the change point of the traffic flow in 24 hours, we divide the traffic flow into four groups, whose distributions are separable from each other. According to the characters of the four groups, we proposed

Gamma distribution-based time series (*GaTS*) models motivated by the generalized linear model [8].

We take the original observations and their Box-Cox transformation [9, 10] as the response variable, which can be considered as random variables generated by a Gamma distribution-based stochastic process. Moreover, we extracted the homochronous term from the historical observations used as explanatory variables.

We use the Bayesian information criterion (BIC) to select the optimal model structure. By these means, the proposed model can predict not only the mathematical expectation but also the nonstationary variance from the past observations. Furthermore, using the homochronous term makes our model of outstanding accuracy in multistep-ahead prediction. Finally, the real data collected from Dalian city in China are used to validate the performance of the proposed *GaTS*. The computational results from the real-world data indicate that the homochronous term is helpful to enhance multistep-ahead prediction performance. Meanwhile, *GaTS* is of the linear structure. Thus, *GaTS* is efficient and convenient for further control design in ITS.

The rest of this paper is organized as follows: In Section 2, we review the studies on short-term traffic flow prediction. In Section 3, we present the *GaTS* methodology for generating traffic flow data as building blocks for prediction. In Section 4, we discuss the experimental results. At last, concluding remarks are described in Section 5.

## 2. Literature Review

Over the past few decades, a lot of mathematical models have been developed by using statistics and machine learning methods for traffic flow prediction. The regressive type model, including the autoregressive models and the support vector regression (SVR), has been used as the parametric methods. The nonlinear model, like the artificial neural network (ANN) model, has also been applied to the prediction of the traffic flow. Besides the parametric methods, the non-parametric models, including the  $k$ -nearest neighbour (KNN) model, were also constructed.

In the family of the regression-type parametric models, the autoregressive integrated moving average (ARIMA) models were widely used for predicting the traffic flow [11–17]. Besides, the extensions of ARIMA have been studied for the prediction of traffic flow. The space-time autoregressive integrated moving-average model was proposed to fulfil the internal relationships of the links [18]. Stathopoulos and Karlaftis [19] designed a model for predicting the traffic congestion on the basis of a multivariate time-series state-space model. Meanwhile, SVR is also used for traffic prediction [20–23]. These regressive models mainly focused on the prediction of the tendency (mathematical expectation) of the traffic flow and ignored the statistics for the dispersion (variance).

The ANN model is one of the most commonly used nonlinear models in artificial intelligence methods. Most of the researchers proposed to apply ANN for the traffic prediction problem using a different architecture of the ANN models or to treat the ANN model as a baseline for comparing a wide variety of classification methods [7, 24–28]. It has

been observed recently that many researchers in the field have proposed to integrate ANN with different preprocessing methods like fuzzy methods to improve its performances [29–32]. Wang et al. [33] proposed to design a prediction model for traffic flow by integrating a fuzzy ANN using the Taguchi method. This work employed the Taguchi method to fix up a count for sensors along the roadside. They proved the benefits of the information collected through the detectors. On similar lines, Quek et al. [34] proposed to utilize a fuzzy-based ANN to the problem of estimation of short-term traffic flow. The reported results indicated that the performance of the proposed model was promising in comparison to the backpropagation-based trained feed-forward (FF) ANN. The ANN and deep learning networks have achieved excellent performance on the traffic flow prediction problem [35–37]. However, they are relatively difficult to develop further traffic control designs for their complex structures.

Effectively modelling traffic flow variance can produce more accurate confidence intervals for short-term traffic flow forecasts and thus improves prediction reliability. Because the generalized autoregressive conditional heteroscedasticity (GARCH) model can be used to describe the time-varying volatility structure of the time series data, it was used by Kamarianakis et al. [38] for prediction of the conditional variance of speed with the mean equation of the ARIMA model. Similarly, GARCH was used for dependent variance prediction to 15 min volume based on a seasonal ARIMA model [39, 40]. Furthermore, Tsekeris and Stathopoulos used a fractionally integrated asymmetric power GARCH model with the mean equation of an autoregressive fractionally integrated moving average model for traffic volatility prediction and found that the combined model outperformed the ARIMA-GARCH model [41]. Because of the stochastic characteristics in traffic flow series, another volatility model, the stochastic volatility model, was proposed by Tsekeris and Stathopoulos [42] for urban traffic variability prediction. The evaluation results showed that the stochastic volatility model could produce a more accurate forecast speed variance than GARCH.

KNN is an essential method in the family of nonparametric methods. KNN has the ability to predict the sampled data based on a number  $k$  without formulating a model [43, 44]. KNN can most benefit the situation with little prior knowledge. Keeping its simplicity and better performance into consideration, the popularity of the applied algorithm is increasing in the field of traffic prediction. Yu et al. [45] proposed a KNN model for regression of estimation of multiple-time-step prediction. The parameters are measured for each minute by a loop detector. Hou et al. [46] presented a model for determining the flow of short-term traffic based on KNN. The major limitation of the algorithm is that it requires tremendous computational resources for a massive amount of historical data. The algorithm also suffers a limitation of the sensitivity of the outliers of archival data.

Besides the works mentioned above, the Kalman filtering method [47, 48], advanced techniques for kernel regression [49, 50], and mixtures of multivariate Gaussian processes [51] were also used to the prediction of the traffic flow.

### 3. Methodology

**3.1. Data Description and Transformation.** This research focused on urban traffic flow data. The data under study were collected at Donglian road, which is an important business centre of Dalian city of China. The microwave radar vehicle detectors were used to record the traffic flow every 15 minutes from Jan. 5, 2016, at 00:15 to Jan. 15, 2016, at 00:00, including a total of 960 samples (Figure 1). As shown in Figure 1, the data are of apparent periodicity, which suggests that the information in the corresponding period of the past days can be helpful for the prediction. This fact motivates us to improve the multistep-ahead prediction by using the homochronous term.

Figure 2 summarizes the boxplots for the corresponding sampling points of 10 days. The ranges of the boxes at each corresponding sampling point show that the variance of traffic flow is small at night and is large in the daytime. Thus, the traffic data is of time-varying variance, i.e., heteroscedasticity. Furthermore, we divide the data into four segments by using the change point detection method for the periodic time series [52]. Figure 3 illustrates the histograms for the four segments. Figures 3(a)–3(c) suggest that the distribution for the positive random variable, like the log-normal distribution and the Gamma distribution, can be used. However, Figure 3(d) is the histogram of the data collected from midnight to early morning and is of a single right tail, which cannot be approximated by the density function of the log-normal distribution. Thus, we use the Gamma distribution to build the time series model.

Furthermore, we use the Box-Cox transformation to find a proper positive real-valued time series data as the following:

$$y_t = \begin{cases} \frac{(x_t + 1)^\lambda - 1}{\lambda}, & \text{for } \lambda \neq 0, \\ \log(x_t + 1), & \text{for } \lambda = 0. \end{cases} \quad (1)$$

We use BIC to select the proper  $\lambda$ .

**3.2. Model Structure.** Because the transformed  $y_t$  is nonnegative real-valued, we assume that  $y_t$  obeys a Gamma distribution. Let  $f(y_t | \mu_t, \sigma_t)$  denote the probability density function, with  $\mu_t$  and  $\sigma_t$  being, respectively, the location and scale parameters. Consequently, the conditional probability density of  $y_{t+j}$  can be formulated as the following:

$$f(y_{t+j} | \mu_t, \sigma_t) = \frac{1}{(\sigma_{t+j}^2 \mu_{t+j})^{1/\sigma_{t+j}} \Gamma(1/\sigma_{t+j}^2)} y_{t+j}^{(1/\sigma_{t+j}^2)-1} e^{-(y_{t+j}/\sigma_{t+j}^2 \mu_{t+j})}, \quad (2)$$

with  $\Gamma(\cdot)$  being the Gamma function.

The time-varying  $\sigma_t$  implies that the stochastic process generating  $y_t$  is nonstationary. To predict such nonstation-

ary on the basis of historical data,  $\mu_t$  and  $\sigma_t$  are regressed as follows:

$$\begin{cases} \mu_{t+j} = \beta_{10} + \beta_1^T u_{1t}, \\ \log(\sigma_{t+j}) = \beta_{20} + \beta_2^T u_{2t}, \end{cases} \quad (3)$$

where  $u_{it}$  is the explanatory vector given by  $[y_t, \dots, y_{t-l_y}, c_{t+j}, \dots, c_{t+j-l_{is}}]^T$  with  $l_{iy}$  and  $l_{is}$  being the maximum time lags of each variable for  $i = 1, 2$ .  $j$  means that (3) is used for  $j$ -step-head prediction.  $c_t$  is called the homochronous term, which is the mean of the observations at time  $t$  in the past five days up to the day containing  $y_t$ . Then, for the data set  $\{(y_t, c_t) | t = 1, 2, \dots, T\}$ , the likelihood can be formulated as the following:

$$\begin{aligned} L(\mathcal{B} |) &= f(y_0, y_1, \dots, y_t | \mathcal{B}, c_0, c_1, \dots, c_t) \\ &= f(y_0, y_1, \dots, y_{l-1}) \\ &\quad \times \prod_{t=l}^{T-j} f(y_{t+j} | \mathcal{B}, y_0, y_1, \dots, y_{l-1}, c_0, c_1, \dots, c_{t+j}) \\ &= f(y_0, y_1, \dots, y_{l-1}) \times \prod_{t=l}^{T-j} f(y_{t+j} | \mathcal{B}, u_{it}), \end{aligned} \quad (4)$$

where  $f(y_0, y_1, \dots, y_{l-1})$  is the initial joint distribution and  $\mathcal{B} = \{\beta_{i0}, \beta_i | i = 1, 2\}$  is the set of unknown parameters. Note that the initial joint distribution is not the function about  $\mathcal{B}$ . The parameter set  $\mathcal{B}$  can be estimated by solving the following maximum likelihood estimation problem:

$$\hat{\mathcal{B}} = \arg \max L(\mathcal{B}). \quad (5)$$

**3.3. Evaluation Criteria for Prediction Performance.** To comprehensively test the prediction performance of the models, several evaluation criteria are calculated. We use the mean absolute error (MAE) [53] and root mean square error (RMSE) [54] to show the scale of the prediction error:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n}, \\ \text{RMSE} &= \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}}. \end{aligned} \quad (6)$$

Because the above two criteria cannot be used to evaluate the models crossing the data sets, the coefficient of determination  $R^2$  calculated from the observation  $y_t$  and estimated value  $\hat{y}_t$  is used as the following:

$$R^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2}{\sum_t (y_t - \bar{y})^2}. \quad (7)$$

Here,  $\bar{y}$  is the sample mean of  $y_t$ . From (7), we can know that  $R^2$  does not consider the time-varying variance. Thus, it

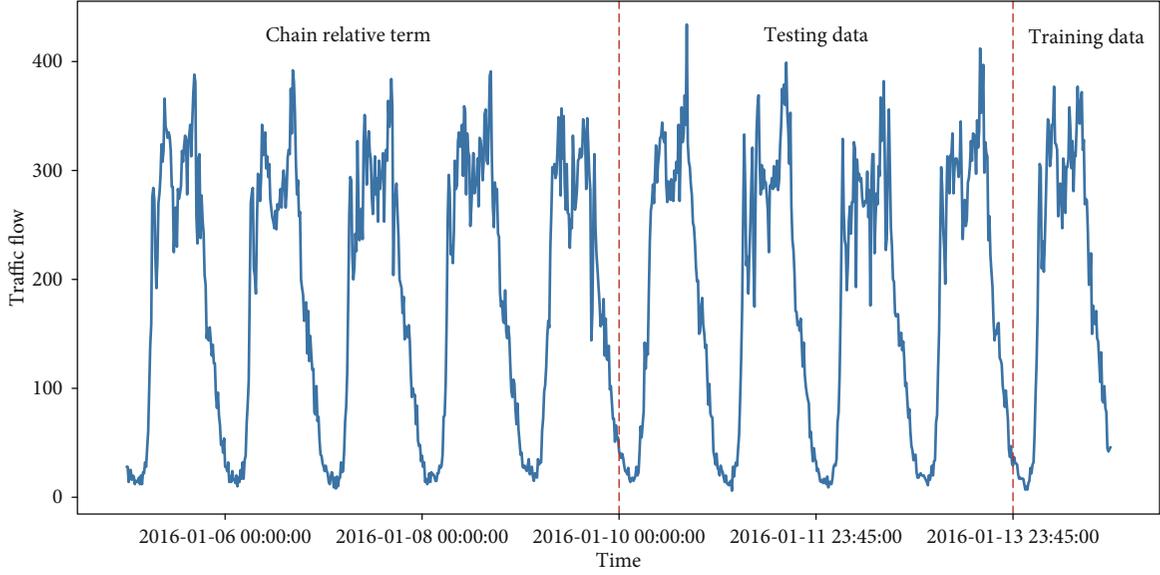


FIGURE 1: Time series of traffic flow in ten days. (Jan. 6, 2016, to Jan. 15, 2016).

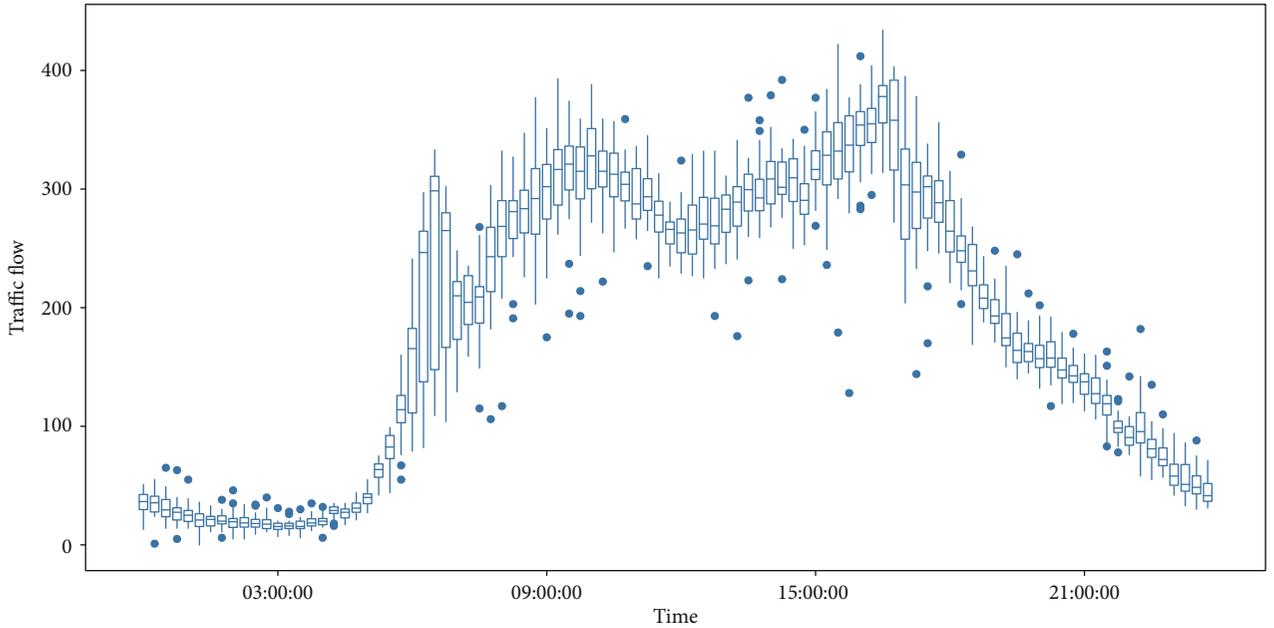


FIGURE 2: Boxplots for each sampling point in 24 hours summarized by 10-day time series.

is not proper for evaluating *GaTS* models with a time-varying variation. To solve this problem, we use the adjusted coefficient of determination  $R_H^2$  [55] as the following:

$$R_H^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2 / \hat{\sigma}_t^2}{\sum_t (y_t - \bar{y}_\omega)^2 / \hat{\sigma}_t^2}, \quad (8)$$

with

$$\bar{y}_\omega = \frac{\sum_t y_t / \hat{\sigma}_t^2}{\sum_t 1 / \hat{\sigma}_t^2} \quad (9)$$

being weighted mean. Note that  $R^2 = R_H^2$  if the scale parameter  $\sigma_t$  is estimated as constant.

The appropriate model structure determined by the time lags is crucial for prediction. Note that both  $R^2$  and  $R_H^2$  are monotonic increasing with the complexity of the model. Therefore, they cannot be used for model structure selection. Instead,

$$BIC = -2 \sum_{t=l}^{T-j} \log f(y_{t+j} | \hat{\mathcal{B}}, u_{it}) + |\hat{\mathcal{B}}| \log(T-j-l+1) \quad (10)$$

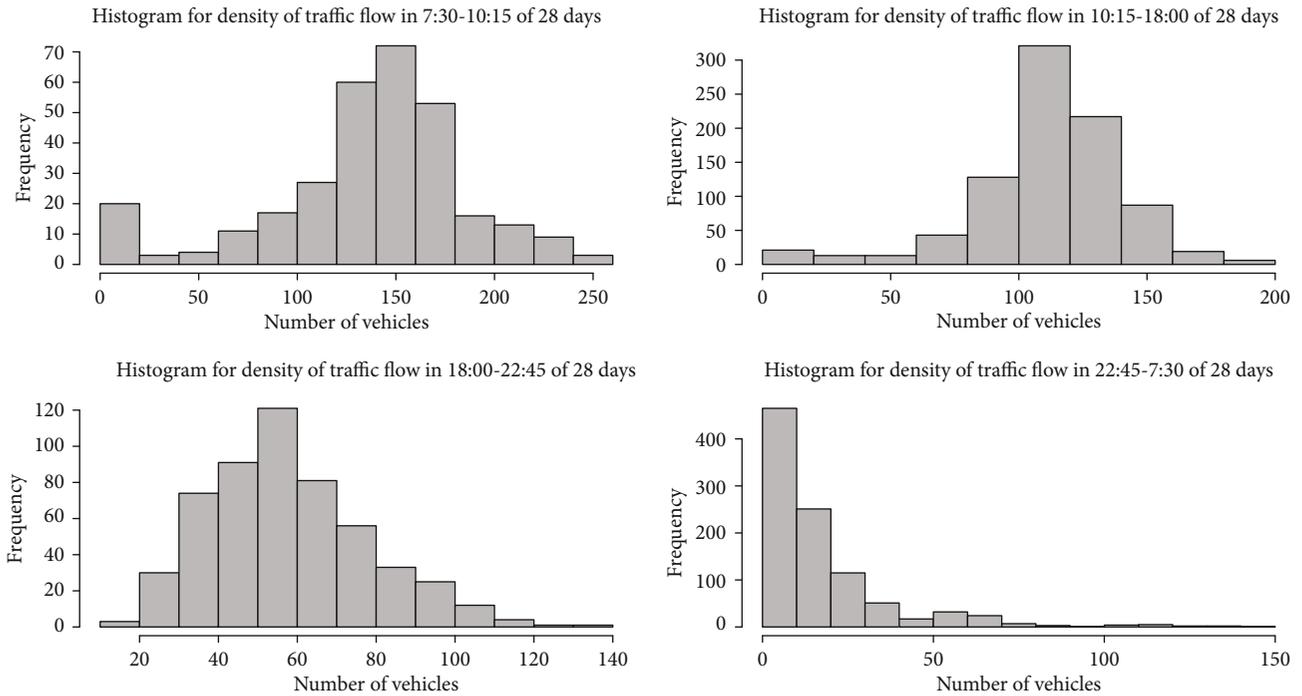


FIGURE 3: Histograms of four segments divided by changing points.

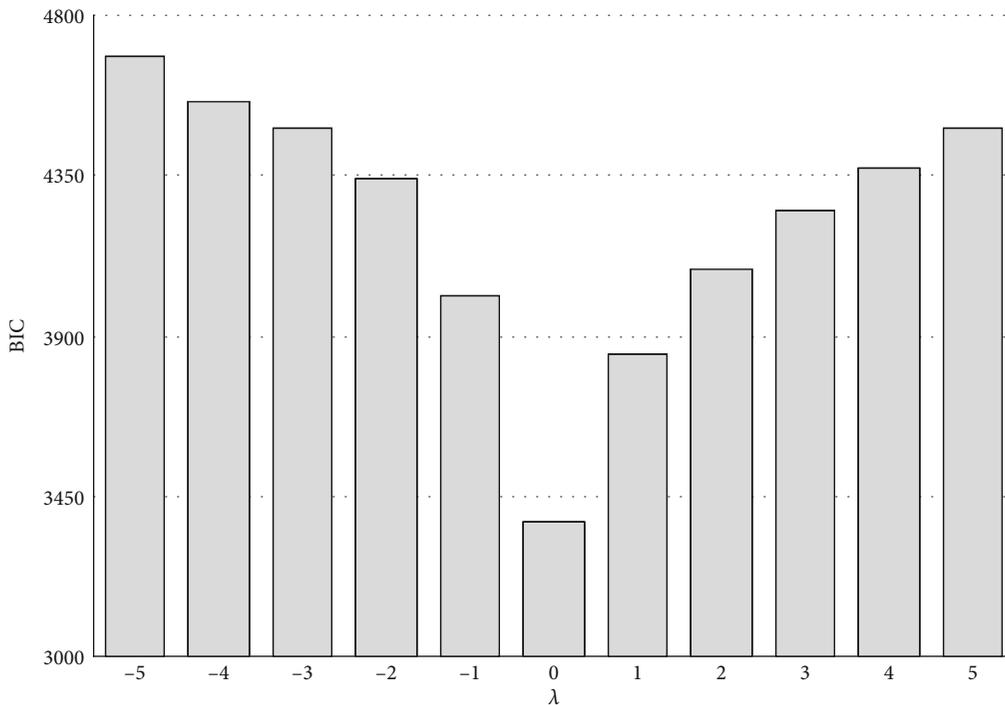


FIGURE 4: BIC values corresponding to several  $\lambda$ 's for Box-Cox transformation.

is used to evaluate  $GaTS$ , where  $|\widehat{\mathcal{B}}|$  means the number of the estimated parameters. We computed BIC to evaluate different models. The optimal  $\lambda$  in the Box-Cox transformation illustrated in Figure 4 and the model structures with the minimum BIC are selected. For fairness, all the

evaluations for various changes are conducted by using the original data. That is, we hold inverse transformation of the Box-Cox transformation (1) for the prediction of the transformed data, and all the evaluation criteria are computed by using the original data.

TABLE 1: Results of six models  $\mathcal{M}_i$ ,  $i = 1, 2, \dots, 6$ , from one- to four-step-ahead.

Prediction	Location $\mu_t$		Scale $\sigma_t$		$R_{Htr}^2$	$R_{Hte}^2$	MAE	RMSE	BIC
	$y_t$	$c_t$	$y_t$	$c_t$					
<i>1-step ahead</i>									
$\mathcal{M}_1$	9	×	×	×	0.913	0.9209	23.1128	35.2989	3624.776
$\mathcal{M}_2$	9	×	1	×	0.9758	0.9816	21.2866	30.8052	3553.057
$\mathcal{M}_3$	4	1	1	0	<b>0.9856</b>	<b>0.9906</b>	15.5009	<b>20.4802</b>	<b>3356.326</b>
$\mathcal{M}_4$	5	1	1	0	0.9794	0.9853	16.1826	21.9583	3466.493
$\mathcal{M}_5$	4	1	1	0	0.9804	0.9869	15.9534	21.5957	3443.754
$\mathcal{M}_6$	5	1	1	0	0.9828	0.9892	<b>15.5007</b>	20.7218	3370.927
<i>2-step ahead</i>									
$\mathcal{M}_1$	9	×	×	×	0.8054	0.7984	33.1687	56.364	3790.313
$\mathcal{M}_2$	9	×	1	×	0.9584	0.966	31.569	47.9216	3767.335
$\mathcal{M}_3$	4	1	1	0	<b>0.9835</b>	0.9884	<b>16.7219</b>	<b>22.9739</b>	<b>3416.928</b>
$\mathcal{M}_4$	5	1	1	0	0.9753	0.9795	18.6151	26.5485	3578.103
$\mathcal{M}_5$	4	1	1	0	0.9817	0.9845	17.1965	24.5677	3493.702
$\mathcal{M}_6$	4	1	1	0	0.9834	<b>0.9888</b>	16.8457	23.5447	3423.929
<i>3-step ahead</i>									
$\mathcal{M}_1$	9	×	×	×	0.6415	0.5814	41.5642	81.2282	3944.636
$\mathcal{M}_2$	9	×	1	×	0.7395	0.7138	36.8849	63.09	3901.592
$\mathcal{M}_3$	3	1	1	0	<b>0.9823</b>	<b>0.9869</b>	17.5917	<b>24.303</b>	<b>3422.813</b>
$\mathcal{M}_4$	4	1	1	0	0.9705	0.9748	19.194	27.9442	3598.335
$\mathcal{M}_5$	4	1	1	0	0.9778	0.9816	18.2077	25.8022	3505.349
$\mathcal{M}_6$	3	1	1	0	0.9822	0.9863	<b>17.3481</b>	24.7383	3431.816
<i>4-step ahead</i>									
$\mathcal{M}_1$	9	×	×	×	0.5572	0.5127	54.2951	115.5521	4042.217
$\mathcal{M}_2$	9	×	1	×	0.6141	0.5618	45.6247	77.9117	3993.805
$\mathcal{M}_3$	3	1	1	0	0.9856	<b>0.9859</b>	<b>17.9045</b>	<b>25.4079</b>	<b>3425.344</b>
$\mathcal{M}_4$	4	1	1	0	0.9769	0.9656	20.0474	28.9373	3603.713
$\mathcal{M}_5$	4	1	1	0	0.984	0.9779	18.8877	26.9388	3521.464
$\mathcal{M}_6$	3	1	1	0	<b>0.9864</b>	0.9858	18.0882	25.606	3434.907

## 4. Experiments

*4.1. Model Structure Setting Up.* Table 1 shows the one- to four-step-ahead ( $15 \times 1$  minutes to  $15 \times 4$  minutes) prediction results obtained by three kinds of models. ( $\{\mathcal{M}_i \mid i = 1, 2, 3\}$  defined by the autoregressive variable  $y_t$  and homochronous term  $c_t$ , where  $\mathcal{M}_1$  is of time-invariant scale variable  $\sigma_t = \sigma$ ,  $\mathcal{M}_2$  and  $\mathcal{M}_3$  are of time-varying scale parameter  $\sigma_t$ , and the homochronous term is applied to  $\mathcal{M}_3$ . For  $\mathcal{M}_1$  to  $\mathcal{M}_3$ , the proper  $\lambda$ 's for the Box-Cox transformation are selected by BIC. Figure 4 shows the BIC values for the  $\lambda$  candidates.  $\mathcal{M}_4$  is the normal distribution-based model.  $\mathcal{M}_5$  is GaTS by using the original data without the Cox-Box transformation.  $\mathcal{M}_6$  is the log-normal distribution-based time series model.) The 2nd to 5th columns are the maximum lags of each variable selected by BIC. Here,  $\times$  means that the variables are unnecessary to be estimated; 0 means

that the variables are not chosen by BIC.  $R_{Htr}^2$  is calculated from the training data, and  $R_{Hte}^2$  is derived from the testing data.

*4.2. Expectation Prediction.* The difference between  $R_{Htr}^2$  values and  $R_{Hte}^2$  values is minimal, which suggests that all the models have been well estimated without overfitting or underfitting. Bold numbers are the best values for each evaluation criterion. They indicate that  $\mathcal{M}_3$ 's with time-varying scale parameters and homochronous terms are the optimal models for one- to four-step-ahead predictions. Meanwhile,  $R_{Hte}^2$  values show that  $\mathcal{M}_3$ 's are of the best prediction performance. In  $\mathcal{M}_3$ 's, the selected regression structures for  $\sigma_t$  are more straightforward than those for  $\mu_t$ . This is similar to the results in work [55]. The homochronous terms  $c_t$  represent the periodicity of  $\mu_t$  rather than that of  $\sigma_t$ , which are

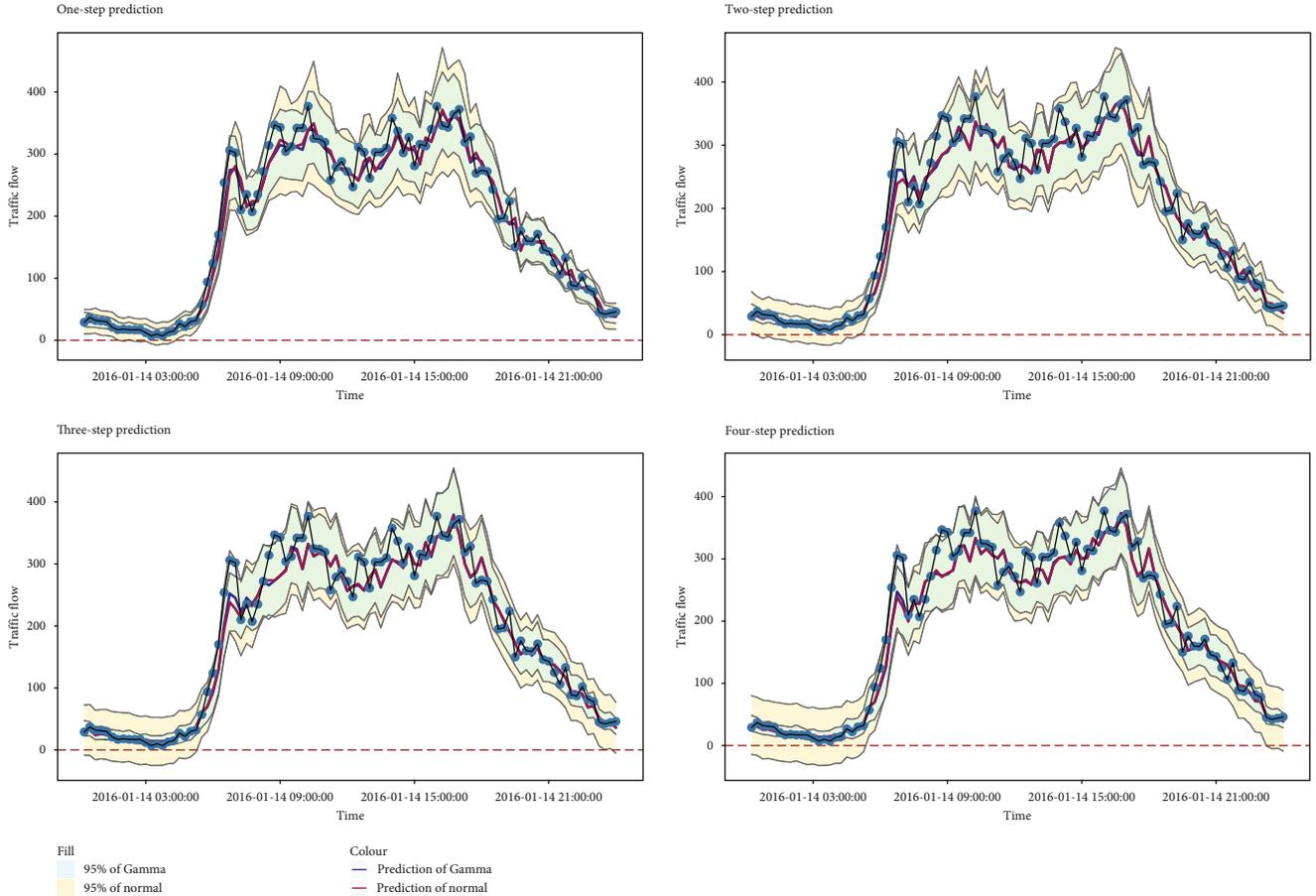


FIGURE 5: Range estimation by using Gamma and normal distribution-based models from one- to four-step-ahead.

the mean of the traffic flow at the same time of 5 successive days before the day containing prediction time  $t$ . Therefore, the homochronous terms have not been selected to regress to  $\sigma_t$  by BIC, as we expected.

We also construct normal distribution, *GaTS* without the Cox-Box transformation and log-normal distribution-based models, denoted by  $\mathcal{M}_4$ 's,  $\mathcal{M}_5$ , and  $\mathcal{M}_6$  in Table 1. By comparing  $\mathcal{M}_i$ 's for  $i = 3, 4, 5, 6$ , we find that  $\mathcal{M}_3$ 's are optimal models according to the minimum BIC values. This suggests that *GaTS* are more optimal than the other distribution-based ones.

**4.3. Range Estimates.** Figures 5 and 6 illustrate the range estimation results for one day on Jan. 14, 2016, in which 95% confidence intervals (CI) are obtained by the predicted  $\hat{\mu}_t$  and  $\hat{\sigma}_t$  along with  $t$ . In Figure 5, the green fields are collected by four  $\mathcal{M}_3$ 's in Table 1, and the yellow areas are obtained by the corresponding normal distribution-based models. Figure 5(a) shows that *GaTS* have a similar CI prediction performance, compared with the normal distribution-based models, for one-step-ahead prediction. However, Figures 5(b)–5(d) indicate that the *GaTS*-based models are of a more narrow range than the normal distribution-based models for multistep-ahead prediction. Furthermore, the normal distribution-based models for multistep-ahead prediction even obtained the negative lower bound of CI, when the traffic

flow values are small. This is contradicting to the fact that the traffic flow is positive-valued.

In Figure 6, the green fields are also collected by four  $\mathcal{M}_3$ 's in Table 1, and the yellow areas are obtained by the corresponding log-normal distribution-based models with the identical structures of four  $\mathcal{M}_3$ 's. From Figures 5(a)–5(d), we can see that both the two models of exponential transformation are more stabilized than using normal distribution directly from one- to four-step. Therefore, the variance of the log-normal distribution is a wider range than the Gamma distribution in the daytime, which is contrary to the real traffic flow state. That will influence the estimation and prediction of the traffic state by the ITS. Furthermore, the CI obtained by *GaTS* can well approximate the heteroscedasticity shown in Figure 2. Thus, *GaTS* can achieve more rational CI predictions.

## 5. Conclusion

This research mainly focused on the prediction issue on the urban traffic flow. By specifying GLM with the Gamma distribution, we proposed *GaTS* to predict the nonstationary stochastic process of the traffic flow. The objective of *GaTS* is to predict the probability distribution of the traffic flow in real time. To this end, the Gamma distribution presents the stochastic properties of nonnegative-valued traffic flow

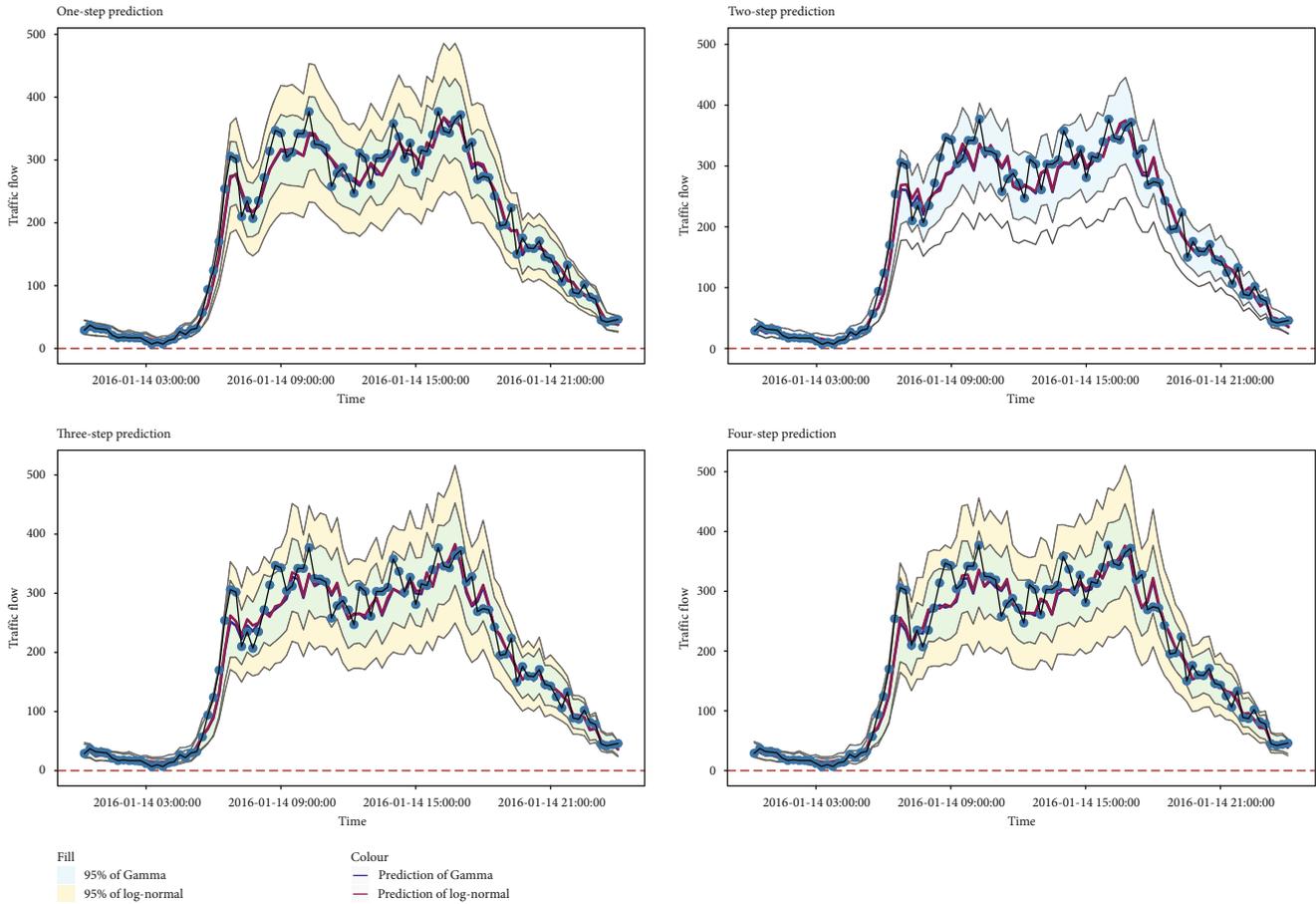


FIGURE 6: Range estimation by using Gamma and log-normal distribution-based models from one- to four-step-ahead.

data and one-side-tailed histogram for the midnight period. The GLM structure supports the heteroscedasticity of the traffic flow. The homochronous term improves the accuracy of the multistep-ahead prediction of the mathematical expectation.

The traffic flow data in this research were collected from Dalian, which is a large port in northern China, as well as a major destination for Chinese tourists [56]. The aggregation of a large number of different types of crowds not only brings traffic and environmental problems but also makes the state of the use of the region's urban public space become complicated and contradictory [57]. Because it is relatively difficult to improve transportation infrastructure, we are focusing on developing intelligent software control and management to relieve traffic congestion. ITS need more accurate historical information and future prediction of the road network [2]. Furthermore, to control and regularize the traffic flow, the model with precise accuracy should be of simple structures. Thus, our proposed *GaTS* is more proper than the research focus on the models of ANN and deep learning. Furthermore, a series of *GaTS* can be extended to model the joint distribution for the joint prediction of the multiple sensors.

Several research topics can be further considered. We successfully specified GLM by using the Gamma distribution

for the urban zone. However, the probability distributions for other zones, like the highways, should be further summarized. The potential external factors, which have a relation with the traffic flow and can be governed by ITS, should be investigated to be embedded into *GaTS*. On the basis of *GaTS* with the external factors, the cost function for the control and regularization of the traffic flow should be constructed, and the corresponding optimization solver should be developed.

## Data Availability

The data that support the findings of this study are available from the ITS database of the traffic police department in Dalian city of China. But restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the traffic police department.

## Conflicts of Interest

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

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (Grant Nos. U1560102, 61633006, and 61502074).

## References

- [1] D. Pavlyuk, "Feature selection and extraction in spatiotemporal traffic forecasting: a systematic literature review," *European Transport Research Review*, vol. 11, no. 1, 2019.
- [2] H. Li, Z. Li, R. White, and X. Wu, "A real-time transportation prediction system," *Applied Intelligence*, vol. 39, no. 4, pp. 793–804, 2013.
- [3] W. Ma and Z. Qian, "On the variance of recurrent traffic flow for statistical traffic assignment," *Transportation Research Part C: Emerging Technologies*, vol. 81, pp. 57–82, 2017.
- [4] S. Sarkar, S. Chawla, S. Ahmad et al., "Effective urban structure inference from traffic flow dynamics," *IEEE Transactions Big Data*, vol. 3, no. 2, pp. 181–193, 2017.
- [5] P. Greibe, "Accident prediction models for urban roads," *Accident Analysis & Prevention*, vol. 35, no. 2, pp. 273–285, 2003.
- [6] S. G. Wang, D. X. Yu, X. G. Ma, and X. Xing, "Analyzing urban traffic demand distribution and the correlation between traffic flow and the built environment based on detector data and POIs," *European Transport Research Review*, vol. 10, no. 2, 2018.
- [7] Z. H. Khan and T. A. Gulliver, "A macroscopic traffic model for traffic flow harmonization," *European Transport Research Review*, vol. 10, no. 2, 2018.
- [8] J. A. Nelder and R. W. Wedderburn, "Generalized linear models," *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, p. 370, 1972.
- [9] G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 26, no. 2, pp. 211–252, 1964.
- [10] R. M. Sakia, "The Box-Cox transformation technique: a review," *The Statistician*, vol. 41, no. 2, pp. 169–178, 1992.
- [11] S. Ishak and H. Al-Deek, "Performance evaluation of short-term time-series traffic prediction model," *Journal of Transportation Engineering*, vol. 128, no. 6, pp. 490–498, 2002.
- [12] 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*, vol. 1644, no. 1, pp. 132–141, 1998.
- [13] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.
- [14] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *European Transport Research Review*, vol. 7, no. 3, 2015.
- [15] H. Zhang, X. Wang, J. Cao, M. Tang, and Y. Guo, "A multivariate short-term traffic flow forecasting method based on wavelet analysis and seasonal time series," *Applied Intelligence*, vol. 48, no. 10, pp. 3827–3838, 2018.
- [16] X. Luo, L. Niu, and S. Zhang, "An algorithm for traffic flow prediction based on improved SARIMA and GA," *KSCE Journal of Civil Engineering*, vol. 22, no. 10, pp. 4107–4115, 2018.
- [17] H. Zhang, X. Wang, J. Cao, M. Tang, and Y. Guo, "A hybrid short-term traffic flow forecasting model based on time series multifractal characteristics," *Applied Intelligence*, vol. 48, no. 8, pp. 2429–2440, 2018.
- [18] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 606–616, 2011.
- [19] A. Stathopoulos and M. G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 121–135, 2003.
- [20] M. Castro-Neto, Y. S. Jeong, and M. K. Jeong, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6164–6173, 2009.
- [21] W. C. Hong, "Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm," *Neurocomputing*, vol. 74, no. 12–13, pp. 2096–2107, 2011.
- [22] M. W. Li, W. C. Hong, and H. G. Kang, "Urban traffic flow forecasting using Gauss-SVR with cat mapping, cloud model and PSO hybrid algorithm," *Neurocomputing*, vol. 99, pp. 230–240, 2013.
- [23] Y. S. Jeong, Y. J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013.
- [24] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.
- [25] J. Mackenzie, J. F. Roddick, and R. Zito, "An evaluation of htm and lstm for short-term arterial traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1847–1857, 2019.
- [26] G. G. Zhang, C. Zhang, and W. D. Zhang, "Evolutionary echo state network for long-term time series prediction: on the edge of chaos," *Applied Intelligence*, vol. 50, no. 3, pp. 893–904, 2020.
- [27] M. S. Dougherty and M. R. Cobbett, "Short-term inter-urban traffic forecasts using neural networks," *International Journal of Forecasting*, vol. 13, no. 1, pp. 21–31, 1997.
- [28] W. Wang, J. Chen, J. Wang, J. Chen, J. Liu, and Z. Gong, "Trust-enhanced collaborative filtering for personalized point of interests recommendation," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6124–6132, 2020.
- [29] T. Zhou, G. Han, X. Xu, C. Han, Y. Huang, and J. Qin, "A learning-based multimodel integrated framework for dynamic traffic flow forecasting," *Neural Processing Letters*, vol. 49, no. 1, pp. 407–430, 2019.
- [30] G. Shen, X. Kong, and X. Chen, "Short-term traffic flow intelligent hybrid forecasting model and its application," *Journal of Control Engineering and Applied Informatics*, vol. 13, no. 3, pp. 65–73, 2011.
- [31] J. Z. Zhu, J. X. Cao, and Y. Zhu, "Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections," *Transportation Research Part C: Emerging Technologies*, vol. 47, pp. 139–154, 2014.
- [32] A. Stathopoulos, L. Dimitriou, and T. Tsekeris, "Fuzzy modeling approach for combined forecasting of urban traffic flow," *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 7, pp. 521–535, 2008.

- [33] X. X. Wang, L. H. Xu, and K. X. Chen, "Data-driven short-term forecasting for urban road network traffic based on data processing and LSTM-RNN," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3043–3060, 2019.
- [34] C. Quek, M. Pasquier, and B. B. S. Lim, "POP-TRAFFIC: a novel fuzzy neural approach to road traffic analysis and prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 2, pp. 133–146, 2006.
- [35] J. Gao, Y. L. Murphey, and H. Zhu, "Multivariate time series prediction of lane changing behavior using deep neural network," *Applied Intelligence*, vol. 48, no. 10, pp. 3523–3537, 2018.
- [36] G. Yu and J. Liu, "A hybrid prediction approach for road tunnel traffic based on spatial-temporary data fusion," *Applied Intelligence*, vol. 49, no. 4, pp. 1421–1436, 2019.
- [37] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.
- [38] Y. Kamarianakis, A. Kanas, and P. Prastacos, "Modeling traffic volatility dynamics in an urban network," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1923, no. 1, pp. 18–27, 2005.
- [39] J. H. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, no. 1, pp. 50–64, 2014.
- [40] W. Huang, W. W. Jia, J. Guo et al., "Real-time prediction of seasonal heteroscedasticity in vehicular traffic flow series," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3170–3180, 2018.
- [41] C. Chen, J. Hu, and Q. Meng, "Short-time traffic flow prediction with ARIMA-GARCH model," in *2011 IEEE Intelligent Vehicles Symposium (IV)*, pp. 607–612, Baden-Baden, Germany, June 2011.
- [42] T. Tsekeris and A. Stathopoulos, "Short-term prediction of urban traffic variability: stochastic volatility modeling approach," *Journal of Transportation Engineering*, vol. 136, no. 7, pp. 606–613, 2010.
- [43] H. H. Xie, X. H. Dai, and Y. Qi, "Improved K nearest neighbor algorithm for short-term traffic flow prediction," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 14, no. 3, pp. 87–94, 2014.
- [44] Z. Zheng and D. Su, "Short-term traffic volume forecasting: A  $k$ -nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm," *Transportation Research Part C: Emerging Technologies*, vol. 43, no. SI, pp. 143–157, 2014.
- [45] B. Yu, X. Song, F. Guan, Z. Yang, and B. Yao, "K-nearest neighbor model for multiple-time-step prediction of short-term traffic condition," *Journal of Transportation Engineering*, vol. 142, no. 6, p. 04016018, 2016.
- [46] X. Hou, Y. Wang, and S. Hu, "Short-term traffic flow forecasting based on two-tier K-nearest neighbor algorithm," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 2529–2536, 2013.
- [47] W. He, C. Zhong, M. A. Sotelo, X. Chu, X. Liu, and Z. Li, "Short-term vessel traffic flow forecasting by using an improved Kalman model," *Cluster Computing*, vol. 22, no. S4, pp. 7907–7916, 2019.
- [48] A. Emami, M. Sarvi, and S. A. 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.
- [49] D. Ngoduy, "Kernel smoothing method applicable to the dynamic calibration of traffic flow models," *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 6, pp. 420–432, 2011.
- [50] R. Huang and S. Sun, "Kernel regression with sparse metric learning," *Journal of Intelligent Fuzzy Systems*, vol. 24, no. 4, pp. 775–787, 2013.
- [51] S. Sun, "Infinite mixtures of multivariate Gaussian processes," in *2013 International Conference on Machine Learning and Cybernetics*, pp. 1011–1016, Tianjin, China, July 2013.
- [52] R. Lund, X. L. Wang, Q. Q. Lu, J. Reeves, C. Gallagher, and Y. Feng, "Changepoint detection in periodic and autocorrelated time series," *Journal of Climate*, vol. 20, no. 20, pp. 5178–5190, 2007.
- [53] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate research*, vol. 30, no. 1, pp. 79–82, 2005.
- [54] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.
- [55] P. Qin and R. Nishii, "Statistical Prediction of Dst Index by Solar Wind Data and  $\text{StS}$  -Distributions," *IEEE Transactions on Plasma Science*, vol. 43, no. 11, pp. 3908–3915, 2015.
- [56] L. Guo and H. Li, "Effect of foreign trade and FDI on Dalian's upgrading of industrial structure-an analysis based on the empirical data in 1990-2010," *Journal of Lanzhou Commercial College*, vol. 5, p. 20, 2012.
- [57] N. He and S. C. Zhao, "Urban road traffic impedance function - Dalian city case study," *Journal of Highway and Transportation Research and Development (English Edition)*, vol. 8, no. 3, pp. 90–95, 2014.