

Research Article

Short-Term Reactive Power Forecasting Based on Real Power Demand Using Holt-Winters' Model Ensemble by Global Flower Pollination Algorithm for Microgrid

Rajeswari Bhol ¹, Sarat Chandra Swain ¹, Ritesh Dash ²,
Kalvakurthi Jyotheeswara Reddy ², C. Dhanamjayulu ³ and Baseem Khan ⁴

¹School of Electrical Engineering, KIIT Deemed to Be University, Bhubaneswar, India

²School of Electrical and Electronics Engineering, REVA University, Bangalore, India

³School of Electrical Engineering, Vellore Institute of Technology, Vellore, India

⁴Department of Electrical and Computer Engineering, Hawassa University, Hawassa 05, Ethiopia

Correspondence should be addressed to C. Dhanamjayulu; dhanush403@gmail.com and Baseem Khan; baseemkh@hu.edu.et

Received 1 February 2023; Revised 8 October 2023; Accepted 30 November 2023; Published 27 December 2023

Academic Editor: Hamza Faraji

Copyright © 2023 Rajeswari Bhol 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.

Load forecasting is an integral part of the energy study unit to schedule the generating unit by the load demand. Many studies were conducted on load forecasting based on real power demand; however, very few papers were published on reactive power demand. In this research work, an attempt has been made to predict the requirement of reactive power as a function of demand for real power. Household loads are considered for evaluating the demand for reactive power as a critical load. The attempt has been made based on the data collected from the laboratory experimental setup for one year. The load forecasting requires time series analysis of the data set along with error minimization between predicted values and actual value; therefore, the global flower pollination algorithm along with Holt-Winters' exponential model has been used to predict the reactive power. Autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA) models have been used as benchmarking models to evaluate the effectiveness of the model under various conditions. Python-developed model has been used to predict the demand for reactive power, and a MATLAB model has been developed to optimize the cost function. A detailed comparative analysis of the proposed model along with some well-established optimized models such as GA, PSO, and FPA has been presented related to evening peak demand for a microgrid architecture in Conclusion. The analysis includes median values of different quantities such as nMBE, nMAE, nRMSE, and RMSE. Normalized MBE, indicating underestimation and overestimation, is negative for ARIMA but 0.42 for HW-GFPA during validation and 0.43 for the testing data set. Normalized RMSE, measuring the variance between actual and forecasted values, is lowest at 0.803 for proposed HW-GFPA during validation and 0.799 for testing.

1. Introduction

The term “load” in load forecasting applications usually refers to active power (kW) or energy (kWh). There are few papers or articles on reactive power prediction. Many important processes are based on the correct understanding of reactive power such as voltage/VAR optimization, power quality improvement, frequency control, and steady-state power flow analysis [1]. Therefore, accurate reactive power forecasting is necessary for power engineers to understand future reactive

power profiles and improve the quality of grid operations at the transmission and distribution level [2, 3].

In recent years, the commercialization of smart grids and the liberalization of energy markets have created additional demands for accurate reactive power prediction [4, 5]. In literature, the day-ahead and hour-ahead planning of distribution energy resources requires reactive power forecasts to ensure compliance with grid limits [6]. These forecasts are necessary at the aggregate level and point load level to achieve good planning. The impact of reactive power on

optimal power flow solutions in smart grids and microgrids was discussed [7]. Reactive power forecasting is also relevant to the power market. In many countries, power markets only manage active power trades at regular time intervals. Nonetheless, reactive power projections test the technical feasibility of energy pathways from power plants to consumers by transmission system operators reconfiguring market systems when technical boundary conditions are violated [8, 9]. Moreover, reactive power forecasting is expected to become a fundamental input for the reactive power market shortly [10, 11].

As already mentioned above, the current practice of reactive power prediction consists of multiplying the active power prediction by a factor related to the average power factor or using a simple method that relies on the experience of grid operators [12, 13]. Several scientific papers deal with the prediction of reactive power. Artificial neural networks have been proposed to predict reactive power in houses and substations. Neural networks have also been presented in a probabilistic framework to predict reactive power in bulk utility buses [14, 15]. Another approach based on linear and piecewise linear relationships between active and reactive power was used to build a regression model of reactive power [16]. The Takagi-Sugano fuzzy method is applied to the ultrashort-term prediction of active and reactive power [17].

Reactive power forecasting based on fuzzy logic requires framework verification for each stage of approval. Again, rationalizing each step is also not feasible for analyzing the boundary condition [18]. This also affects the accuracy setting and enrolling capacity of the system. It also lacks hypothesis testing, which is required to validate the model under partial variation conditions. Furthermore, ANFIS involves a very complex structure and gradient learning strategies for its realization [19]. The ANFIS model becomes more complicated and, at the same time, also requires more computational time. The design of the membership function dynamically also requires a very complex analysis [20]. It is required to handle seven parameters at the input to predict the demand for reactive power in a microgrid. The seven parameters include voltage, frequency, time stamp, number of nonlinear load connected, temperature, real power drawn, and power factor. To address such a large number of data sets, it is required to go for regression analysis, which will combine all the input to predict the output at any instant of time for 5 minutes or 300 sec. Thus, short-term forecasting can be best modeled by the autoregressive integrated moving average (ARIMA) model. Here, in this research work, a seasonal variation to ARIMA has been added to make it simpler and practically feasible. These ARIMA and SARIMA models will act as base models or benchmarking models. The research paper has the following contributions and outcomes:

- (i) The dynamic optimization model for reactive power prediction based on real power demand and six parameters has been modeled using a global flower pollination algorithm
- (ii) The developed framework (algorithm) is accurately forecasting the short-term demand of reactive power

The rest of the research article has been organized as follows: Section 2 represents the problem formulation based

on the proposed scheme as mentioned above. Two benchmarking models such as ARIMA and SARIMA have been discussed under the benchmarking model of Section 4. A detailed result analysis has been described in Section 5 followed by the conclusion discussed in Section 6.

2. Problem Formulation

In order to analyze the reactive power forecasting based on the real power and consequent price prediction, a time series model can be designed. The flow of reactive power (Q) is a function of voltage and real power drawn over a period of time. Therefore,

$$Q_t \sim T(0, \delta^2) \forall t \in Q, \quad (1)$$

where Q_t represents the reactive power drawn over time t and δ represents the functional parameter of voltage. Taking convolution of eq. (1) must satisfy

$$\text{conv}(Q_t, Q_u) = \delta^2 \zeta_{tv} \forall (t, u) \in Q. \quad (2)$$

The convolution method applied over eq. (2) results in the establishment of the white noise signal $\{Q_t\}_{t \in Q} \sim \text{WN}(0, \delta^2)$. Now, in order to analyze standard independent time series data, the system can be modeled in terms of the linear-independent model. Let us compare and model a linear time series signal $\{R_T\}_{t \in Q}$, representing real-time power transferred between two identified buses as a function of reactive power flow data. Therefore,

$$R_t - \sum_{i=1}^T (\psi_i R_{t-i}) = Q_t + \sum_{i=1}^{\tilde{T}} (\theta_i Q_{t-i}) \forall t \in Q. \quad (3)$$

Equation (3) shows that the flow of both real and reactive power including losses can be made the same with the rest of the white noise data. Both the time and its time inverse are on the same plane of operation; therefore, the series can be modeled in terms of the ARIMA model of the two different (p, q) processes. Figure 1 shows the contour for reactive power energy for the scaled area x .

So, defining for a set of periods of data by equation $\theta_i = 0$, the ARIMA function can be written as

$$R_t - \sum_{i=1}^T \psi_i R_{t-i} = \theta_t. \quad (4)$$

Equation (4) represents the series analysis of reactive power modeled on influenced as a function of real power and constant ψ_i over a period of time. Again, from a regression analysis point of view, this model is termed an ARIMA(p) process. Similarly, the ARIMA(q) process becomes

$$R_t = Q_t + \sum_{i=1}^{\tilde{T}} (\theta_i Q_{t-i}) \forall t \in Q. \quad (5)$$

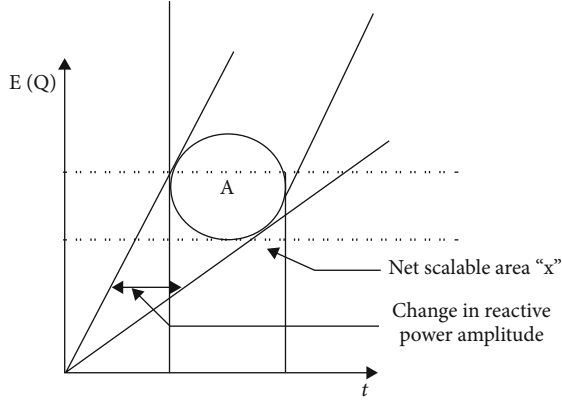


FIGURE 1: Contour for reactive power energy.

Again, from eq. (4), for $|\psi_i| < 1$, eq. (4) can be written as

$$R_t = \sum_{j=0}^N \psi^j Q_{t-j} + \psi^{N+1} R_{t-(N+1)}. \quad (6)$$

Applying energy to eq. (6), it becomes

$$E \left[\left(R_T - \sum_{i=1}^T \psi_i R_{T-i} \right)^2 \right] = \psi_i^{(2N+2)} E \left[R_{T-(N+1)}^2 \right] \rightarrow 0 \quad \forall N \rightarrow \infty. \quad (7)$$

Now, solving eq. (7) for $T \rightarrow \infty$, it becomes

$$\lim_{T \rightarrow \infty} E \left[\left(R_T - \sum_{i=1}^T \psi_i R_{T-i} \right)^2 \right] = \sum_{i=1}^N \psi_i^{2N+2} E \left[R_{T-(N+1)}^2 \right] \rightarrow 0 \quad \forall N \rightarrow \infty, \quad (8)$$

$$\lim_{T \rightarrow 0} E \left[\frac{2!}{T} R_t - \sum_{k=0}^{\bar{T}} R_t \right] = \sum_{i=1}^N \frac{\psi_i^{2N+2}}{3!} [R_t \rightarrow 0], \quad (9)$$

$$\lim_{T \rightarrow 0} E \left[2 \frac{\partial R_t}{\partial t} - \sum_{k=0}^{\bar{T}} \sum_{L=0}^T \frac{\partial^2 R_t}{\partial t^2} \right] = \sum_{i=1}^N \frac{2!(2N+1)}{3!} \psi_i^{2N+1} [R_t \rightarrow 0], \quad (10)$$

or

$$E[R_t] = \sum_{m=0}^{\infty} \psi^m E(q_{n-m})_k = 0. \quad (11)$$

Here, eq. (11) represents the unique existence solution of eq. (10). Now, deriving the covariance function, it becomes

$$\gamma_k = E[R_{t+k} R_t] = \sum_{i=1}^N \sum_{m=0}^{\infty} \psi^{i+m} E(Q_{t+k} \oplus Q_{t+N}). \quad (12)$$

Again, further deriving eq. (12), the energy content of the system becomes

$$E[DR_{tn} - (T : St_0 - T : St_{-n})^2] \leq D_k \sum_{k=0}^{m-1} [t_{k+1}^n - t_k^n]^2. \quad (13)$$

Again, on simplification, eq. (13) becomes

$$E[DR_{tn} - (T : St_0 - T : St_{-n})^2] \leq D_k |Q_t^n(t, R)| [t_{k+1}^n - t_k^n]. \quad (14)$$

Therefore, the inline optimization becomes

$$E[DR_{tn} - (T : St_0 - T : St_{-n})^2] = \sum_{k=0}^{m-1} W(N_k^n) (Q_k^{n+1} - Q_k^n), \quad (15)$$

$$\text{s.t.} = \begin{cases} \lim_{n \rightarrow \infty} E[DR_{tn} - (T : St_0 - T : St_{-n})^2] = \frac{WT^2}{2} + \left(N - \frac{1}{2}\right)T \\ Q_k^n \leq R_k^n \leq Q_{k+1}^n. \end{cases} \quad (16)$$

The energy density equation presented in eq. (16) shows the Gaussian distribution based on probability theory. The empirical confidence can be taken into consideration for evaluating the interpolation. Keeping in view the spikes present in the waveform, in this work, Gaussian flower pollination algorithm (GFPA) is used instead of the flower pollination algorithm (FPA). The merits of GFPA over FPA are listed below.

- (i) Easy in segregating the collateral spike
- (ii) Step size can be varied in accordance with spike length
- (iii) Analysis of periodic exponential term with respect to seasonal variation in reactive power demand

Now, the modified optimization equation becomes

$$DR_{tn}(Q_t^{\text{iter}+1}) = DR_{tn}(Q_t^{\text{iter}}) + L(DR_{tn}\{Q_t^{\text{iter}} - Q_t^{\text{iter}-1}\} + \exp\left(-\frac{d(Q_t^{\text{iter}}, Q_t^{\text{iter}+1})^2}{2e^2}\right)). \quad (17)$$

Being a seasonal regression on time series data, the exponential decaying part presented in eq. (1) converges itself to a solution. A detailed process of flowchart is presented in Figure 2.

3. Data Collection and Analysis

Reactive power analysis over a time period requires a rigorous evaluation of load performance for both linear type and nonlinear type of loads. Therefore, two simulation test

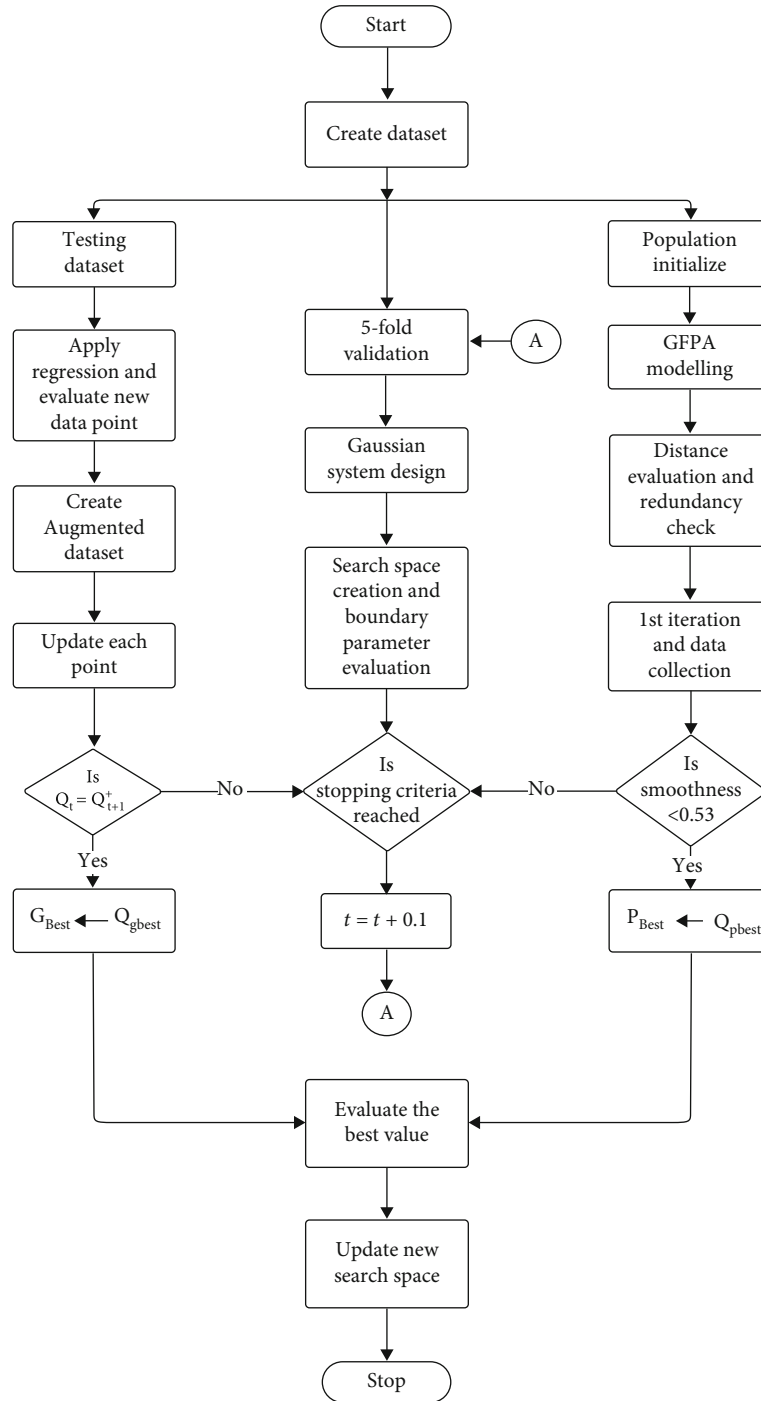


FIGURE 2: Process flowchart for GFPA-enabled prediction system.

models have been created inside the lab one with 20 kVAR load and 30 kVAR load. Both the loads are connected to the microgrid through two different sources such as solar PV and battery. The 20 kVAR load consists of high-speed blowers, heaters, and 7 induction motors (each 2 hp). In order to capture the line power data such as time stamp, reactive power, real power, current, and energy consumption, a phasor measurement unit has been installed at each nodal point. Here, the nodal point represents the load termination point. All the loads were connected to the main grid

through the line scheduler (TPS2301 and SLVS277) with a preloaded time setting. This enables the operator to create an artificial transition between loads, thereby enabling the PMU to capture the required power quality. Thus, the data collected from PMU was saved into a CSV file for further analysis, testing, and evaluation.

The data were collected every 15 seconds for three months. After the collection of the data depending upon the type of load, the data were segregated into 5 different clusters based on their nominal, ordinal, interval, and ratio.

All the data have undergone a redundancy check for incomplete data sets like inserting the unavailable data wherever necessary or deleting a particular set of data. Table 1 shows the basic statistical analysis of PMU data.

The data has been separated from each other based on the null hypothesis. As observed in Table 1, the mean value is lowest for 45 days of data, i.e., -0.223. This signifies that the data points as collected from PMU are not far away from each other. This also signifies that, at no condition, the null hypothesis is rejected. The average standard deviation by consolidating all the data is limited to 1.01%, which is much smaller than the allowed statistical deviation of 5%. Again, as observed, the kurtosis is lightly tailed, which means that the data are not highly deviating from its normal distribution.

Table 2 represents the null hypothesis analysis of the cluster. As seen, the *R*-square error is gradually decreasing from cluster C-15 (days) to C-90. This means that as we approach a large data set, the data set becomes more viable for a time series analysis because the goodness of fit has also increased. The average standard error by considering all the clusters becomes 0.698 (average). This signifies that 69.8% of data can take part in the regression analysis, as they represent how distinct the data are inside the data set. The variation in the *df*1 and *df*2 lies in the range of 0.99 and 1.

This section of the data analysis part represents that all the clusters of data are arranged in proper order as desired by the time series algorithm. All the statistical analysis as shown in Tables 1 and 2 has been utilized by the benchmarking model for further analysis in the next section.

4. Benchmarking Model

4.1. ARIMA. The autoregressive integrated moving average time series model is an amalgamation of differenced autoregressive and moving average models, i.e., a combination of past output and some random noise. The AR in ARIMA represents the autoregression aspect of time series analysis on its data, MA signifies the moving average of errors, and *I* stands for integration, indicating that the data has been differenced with its previous values. Therefore, modifying eq. (15) in terms of ARIMA,

$$\begin{aligned} DR_{in} = & I + w_1 T : St_{t-1} + w_2 T : St_{t-2} + \dots + w_n T \\ & : St_{t-n} + E_t + N_1 E_{t-1} + N_2 E_{t-2} + \dots + N_m E_{t-m}. \end{aligned} \quad (18)$$

Equation (18) represents the ARIMA model reactive power forecasting. Here, *n* represents the autoregressive, and *m* represents the moving average error. *T* : *St*_{*t*-1} represents the random noise. Now, based on eq. (18), two ADF models have been modeled with two different *P* values. The critical parameters for the order of (1, 1, 2) have been evaluated. Table 3 shows the ADF statistics for different *P* values.

Table 4 shows the ARIMA model error analysis for the order (1, 1, 2). Here, in Table 4, the minimum standard deviation has been observed for Ar.L1.D value of 0.213 and that of -0.455 for 0.975 accuracy level. Similarly, Table 5 represents the ARIMA modulus status for the order (1, 1, 2).

Tables 6 and 7 represent the ARIMA model error analysis for the order (2, 1, 2) and modulus status. As shown in Table 6, the minimum standard error of 0.116 was recorded for Ar.L2.D value. Similarly, the corresponding probability value is 0.772, and that of the upper boundary is 0.262. The corresponding moving average value is 0.307 (std. error). This signifies that the system becomes valid under the upper limit of -0.211 to 0.262.

Figure 3 shows the autocorrelation analysis with 1st-order difference and 2nd-order difference of the original reactive power demand series. The higher-order difference analysis has been conducted to find any possible good solution. It is observed that the autocorrelation error for the second difference is very small as compared to the first-order difference. Again, for 91 number of observation, the AIC is 1306.99 and BIC is of 1322.058. Similarly, the HQIC becomes 1313.07.

Figure 4 represents the forecasted reactive power as a function of real power demand for 20 days. Here, the analysis has been carried out for 95% of the confidence interval. Similarly, Figure 5 shows the rolling mean and standard deviation analysis for the original forecasted data. The rolling standard deviation shows the deviation from the average of the reactive curve. Here, it can be found that the maximum deviation is 0.07×10^4 .

4.2. SARIMA. Seasonal ARIMA is one of the most widely used forecasted models. The direct modeling of the SARIMA model can be modeled by using SARIMA (*p, d, q*) × (*P, D, Q*). Here, (*p, d, q*) represents the nonseasonal component, and (*P, D, Q*) represents the seasonal parameter.

Figure 6 represents the autocorrelation function plot for the SARIMA model. As observed, the initial spike = 1 and seasonal spike = 10. It clearly depicts that autoregression is having order 1 and seasonal AR is of order 1.

Similarly, Figure 7 represents the partial autocorrelation for SARIMA. As observed from the figure, the initial lag spike is at 1, and that of the seasonal spike lies at 10. It means that it has a moving average of 1, and that of the seasonal dependent average order is also 1.

Based on the order and class as obtained from Figure 7, Table 8 represents the SARIMA fit model with respect to reactive power time series data.

The autocorrelation has a direct impact on the partial correlation with an order of 50.37% as reported by Ljung-Box. Similarly, a probability of 87% dependency has been observed from the result analysis. The heteroskedasticity factor has been found to be 0.58, which determines the monotonicity involved in the prediction process.

The standard residual plot along with the histogram is plotted in Figure 8. As observed here, the mean residual plot is zero and all the residuals are uncorrelated. This states that the model is well fitted with the data and (*p, q, d*) values.

4.3. Extreme Gradient Boosting (XGBoost). The XGBoost is the application-oriented model of random forest (RF). XGBoost adjusts the weight of weak learners, thereby improving the forecasting rate of each tree. XGBoost improves the generation of complex data sets by avoiding

TABLE 1: Basic statistical analysis of PMU data.

Parameters	15 days	30 days	45 days	60 days	75 days	90 days
Mean	0.073	0.016	-0.223	0.115	0.083	0.067
Standard error	0.0089	0.0019	0.0026	0.00138	0.0099	0.008
Median	0.17	0.21	0.16	0.16	0.19	0.23
Mode	0.07	0.09	0.10	0.26	0.34	0.31
Standard Deviation	0.0411	0.052	0.0816	0.0922	0.161	0.183
Kurtosis	-1.06	-1.13	-1.17	-1.23	-1.27	-1.09
Skewness	-0.279	-0.288	-0.363	-0.419	-0.421	-0.393

TABLE 2: Null hypothesis analysis of cluster.

Cluster (days)	R	R-square	Adjusted R-square	Std. error	df1	df2
C-15	-0.88	-0.22	-0.346	0.831	1	1
C-30	-0.93	-0.34	-0.408	0.866	1	1
C-45	-0.96	-0.42	-0.456	0.785	1	0.99
C-60	-1.08	-0.48	-0.462	0.911	0.99	0.99
C-75	-1.13	-0.53	-0.346	0.785	1	0.99
C-90	-1.17	-0.57	-0.471	0.803	0.99	0.99

TABLE 3: ADF statistics analysis for two different P values.

Model	P value	Critical parameter		
		1%	5%	10%
ADF-1	$1.20 \cdot 10^{-6}$	-3.50	-2.89	-2.58
ADF-2	2.21 times 10^{-10}	-3.50	-2.57	-2.37

TABLE 4: ARIMA model error analysis for the order (1, 1, 2).

Parameter	Coefficient	Std. error	Z	$P > z $	0.025	0.975
Constant	0.2171	1.116	0.195	0.846	-1.970	2.404
Ar.L1.D value	-0.8928	0.213	-3.996	0.000	-1.331	-0.455
ma.L1.D value	-0.1489	0.260	-0.572	0.569	-0.659	0.361
ma.L2.D value	-0.8511	0.260	-3.276	0.002	-1.360	-0.342

TABLE 5: ARIMA modulus status for the order (1, 1, 2).

Model name	Real	Imaginary	Modulus	Frequency
Ar.1	-1.12	+0.00j	1.12	0.5
Ma.1	1.00	+0.00j	1.00	0.0
Ma.2	-1.17	+0.00j	1.17	0.54

TABLE 6: ARIMA model error analysis for the order (2, 1, 2).

Parameter	Coefficient	Std. error	Z	$P > z $	0.025	0.975
Constant	0.2193	1.150	0.191	0.849	-2.035	2.473
Ar.L1.D value	-0.830	0.323	-2.569	0.012	-1.463	-0.197
Ar.L2.D value	0.0338	0.116	0.290	0.772	-0.194	0.262
ma.L1.D value	-0.1873	0.308	-0.609	0.544	-0.790	0.416
ma.L2.D value	-0.8125	0.307	-2.647	0.010	-1.414	-0.211

TABLE 7: ARIMA modulus status for the order (2, 1, 2).

Model name	Real	Imaginary	Modulus	Frequency (kHz)
Ar.1	-1.1508	+0.00j	1.1508	0.5
Ar.2	25.70	+0.00j	25.70	0.00
mA.1	1.001	+0.00j	1.001	0.00
mA.2	-1.2306	+0.00j	+1.2306	0.5

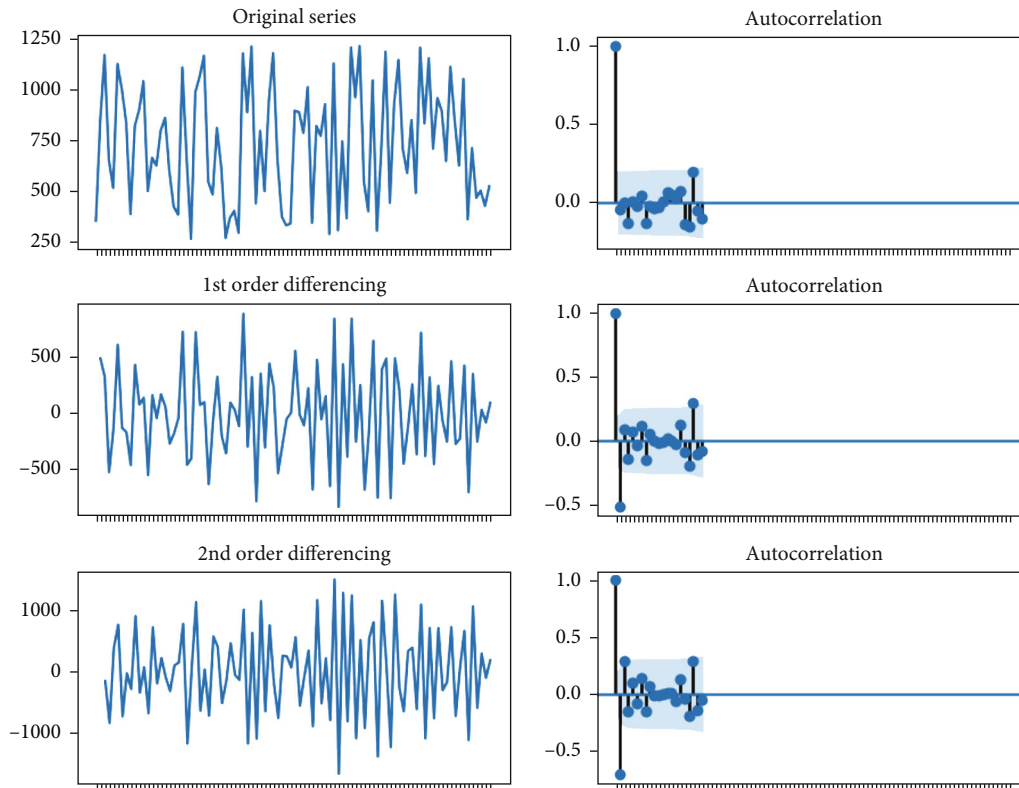


FIGURE 3: Autocorrelation analysis with 1st-order difference and 2nd-order difference.

overfitting of the curve against the data set. The nonlinearity relationship among the complex data set can be handled by providing additive training on each data point, where each previous state of iteration solution ($n - 1$) is considered as an actual forecast. This enables high-dimensional feature space analysis. Mathematically,

$$\hat{y}^{(i)} = \sum_{i=1}^t \left(\frac{P(\hat{y}|_x)^{(i-1)} + P(\hat{y}|_x)^{(i)}}{t} \right). \quad (19)$$

In eq. (19), i represents the iteration level, and t represents the time series data of each level. \hat{y}_t represents state of data, and that of $\hat{y}|_x^{(i-1)}$ represents the previous state of data. Equation (19) offers a wide range of hyperparameter selection, which enables the operator to optimize the model performance.

XGBoost parameter analysis for regression study has been presented in Table 9. Figure 9(a) represents the

predicted vs. actual graph of reactive power demand with a learning rate (L-Rate) of 0.032. As observed, the sample has an associate rule of 0.11. The associate rule here represents the probability of data points inside the data set or how closely they are related to the data of different data sets. The result shows that (Figures 9(b) and 9(c)) the maximum probability of 0.14 can be achieved with the data set of L-Rate and that of 0.34 for tree analysis. Figures 9(d)–9(f) represent feature importance vs. feature index for three different learning rates. As observed in Figure 9(d), the feature importance is positive for sample-3 only. This means that the accuracy of the model in determining the feature prediction is 0.02 which is very low. In Figure 9(e), the feature importance becomes positive for two samples with the highest value of 0.2 and with a learning rate of 0.067. Therefore, it can be assumed that there exists a scope for further improvement. In the 3rd trial in Figure 9(f), all 3 samples have shown positive feature importance with a maximum accuracy level of 0.43.

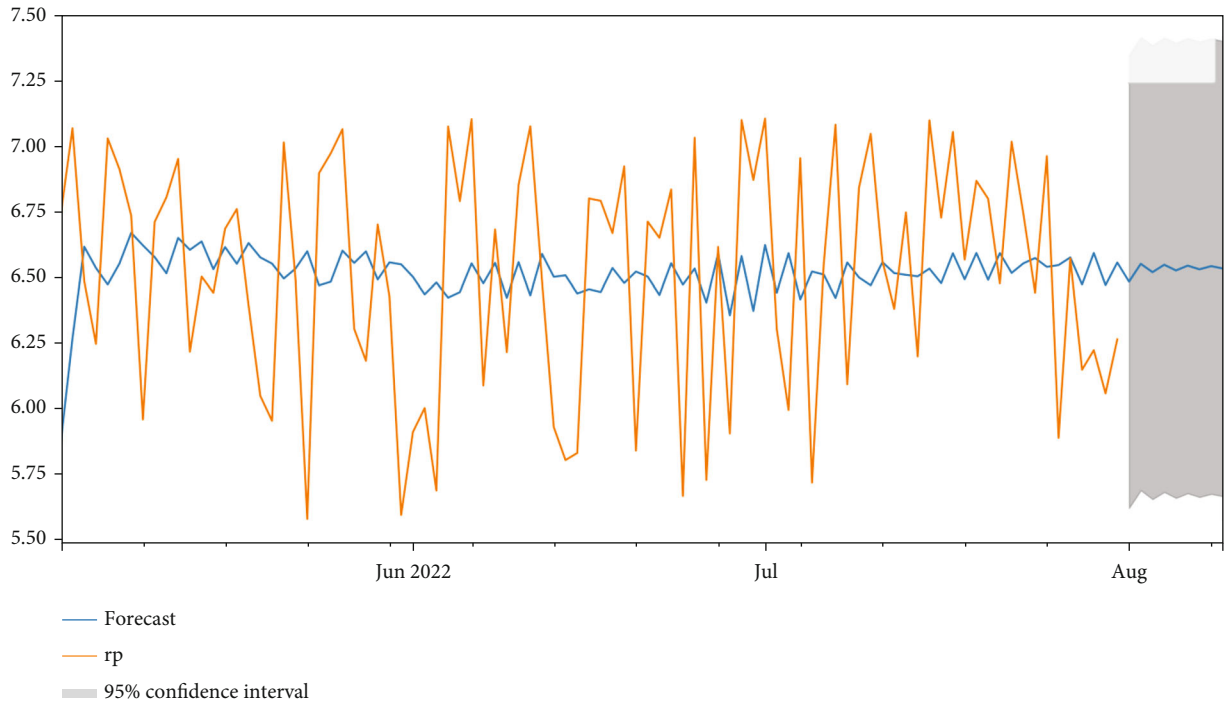


FIGURE 4: Forecasted reactive power as a function of real power demand for 20 days.

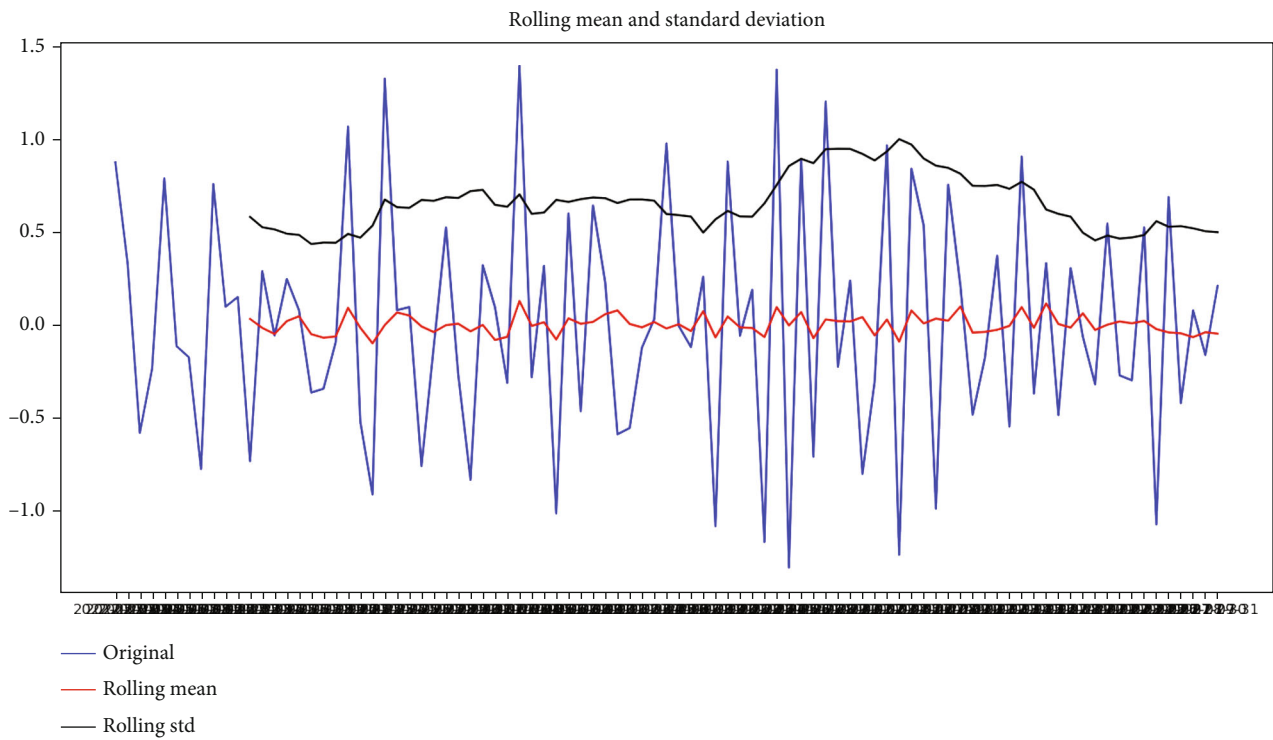


FIGURE 5: Rolling mean and standard deviation analysis for the original forecasted data.

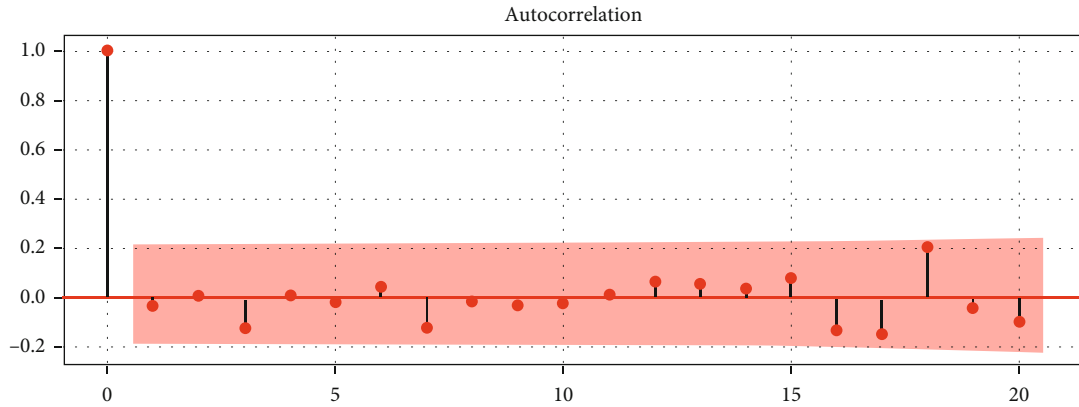


FIGURE 6: Autocorrelation function plot for SARIMA.

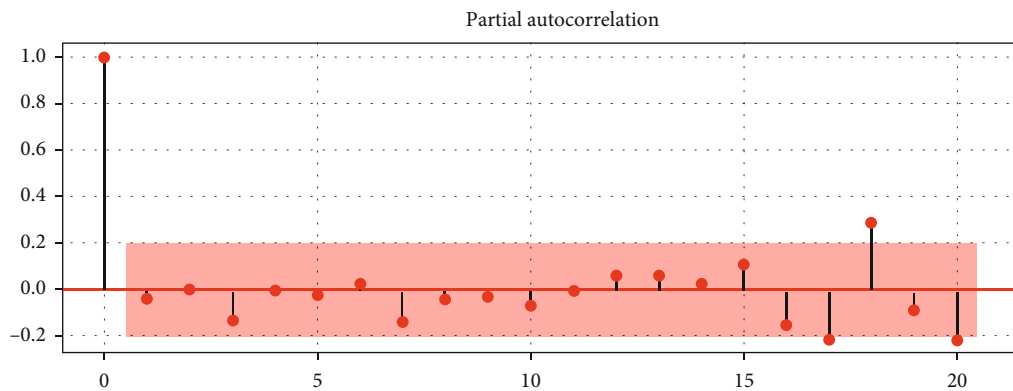


FIGURE 7: Partial autocorrelation for SARIMA.

TABLE 8: SARIMA fit model with respect to reactive power time series data.

Parameter	Coefficient	Std. error	Z	$P > z $	0.025	0.975
ar.L1	-1.1571	0.264	-5.25	0.00	-1.974	-0.940
ar.L2	-0.4891	0.137	-3.56	0.00	-0.758	-0.220
ma.L1	-0.0046	2751.87	-1.67	1.00	-5393.57	5393.56
ma.L2	-0.9954	2739.04	-0.00	1.00	-5369.41	5397.42
ma.S.L4	-0.9122	8.393	-0.109	0.913	-7.36	15.53
ma.S.L8	-0.0863	0.783	-0.110	0.912	-1.62	1.44
Sigma 2	0.2848	783.49	0.000	1.00	-1535.33	1535.90

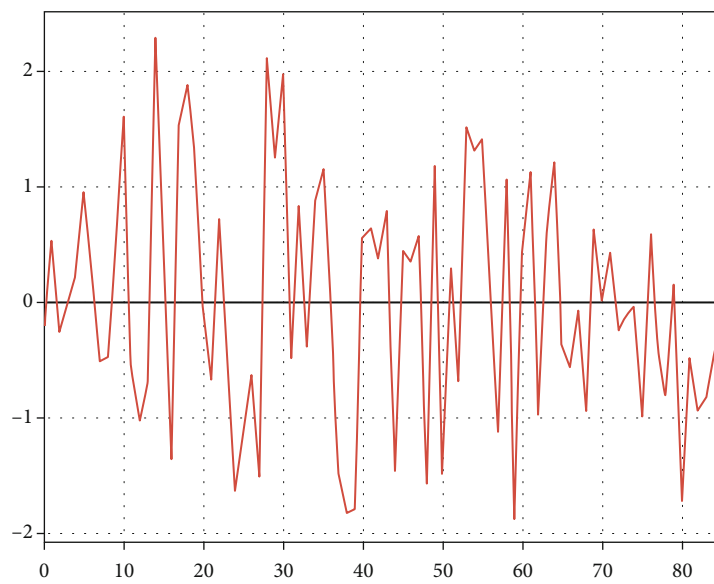
5. Result Analysis

The experimental setup for the proposed work has been started with data collection for reactive power from the laboratory-operated critical load. Three months of continuous data have been collected, i.e., May 2022 to July 2022, through the phasor measurement unit (PMU) from various sensitive points. The data has been tested for redundancy check using Python programming, and wherever the data was missing is filled with mean data from that series.

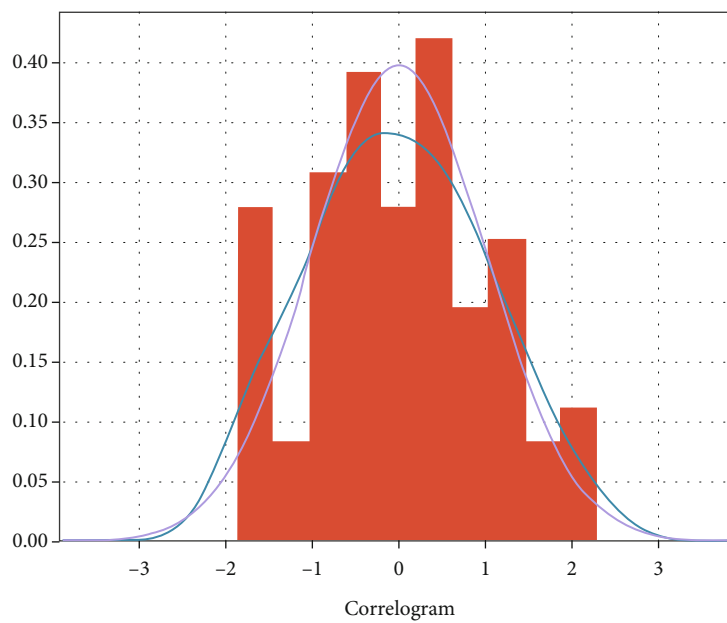
In order to create a proper mathematical optimization model for its validation with respect to the problem-

formulated modeling, a curve fitting analysis has been carried out over the data set. Four different methods of curve fitting have been applied to the data set such as sine order-8, interpolant, polynomial, and smoothing order-4. The detailed statistical analysis of the method is shown in Table 10.

Figure 10(a) represents the sum of the sine curve fitting for different samples. An 8-order fitting model has been applied. As observed, 7.38×10^5 numbers of the sample have been used to predict the sample. Root mean square error (RMSE) has been observed for the smoothing spine curve (Figure 10(b)). As observed, only 112 numbers of the

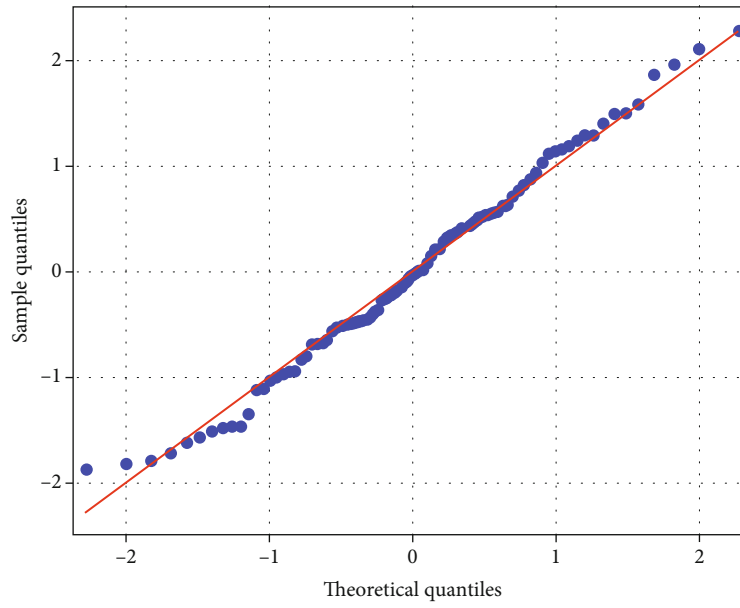


(a)

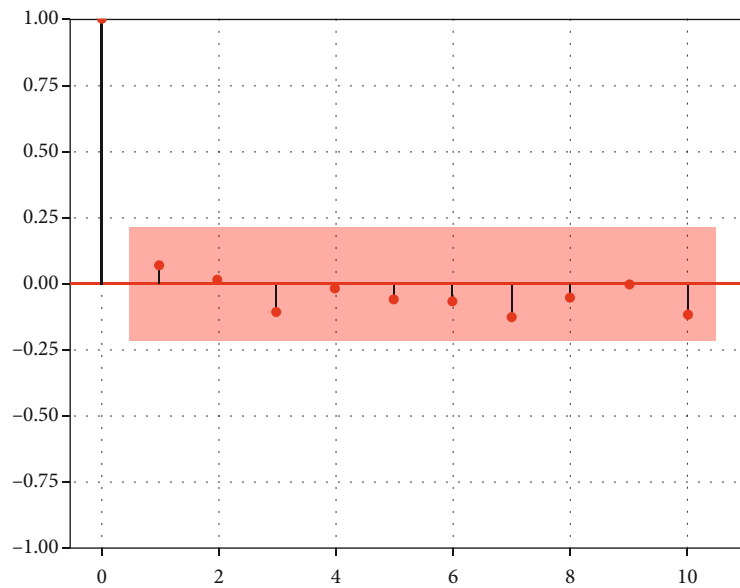


(b)

FIGURE 8: Continued.



(c)



(d)

FIGURE 8: (a) Standard residual plot. (b) Histogram plus estimated density. (c) Normal Q-Q. (d) Correlation plot.

samples have been outside the boundary of the fitted curve. Similarly, polynomial and interpolate analysis has been presented in Figures 10(c) and 10(d). Overfitting and underfitting of the curve have been found in Figures 10(a) and 10(d). Therefore, in this research, article curve fitting model of the smoothing spine has been taken into consideration for further analysis.

In order to optimize the local sample on the smoothing spine, curve global flower pollination-based optimization has been applied in accordance with the optimization equation as presented under eq. (27) and eq. (28). Algorithm 1, presents the pseudocode used in the pollination algorithm

for the global optimization of the curve, here referred to as the reactive power fitted curve.

Figures 11 and 12 represent the cost function optimization for two different cost functions of $c_1 = 0.17$ and $c_2 = 0.21$. It is observed that the error in tracking the trajectory path is limited to 0.026 for c_1 and 0.022 for c_2 . At about 10.00 sample time (Figure 12), the actual direction of optimization is the same as that of trajectory. Based on this, Tables 11 and 12 represent a detailed comparative analysis of different optimization algorithms such as PSO, GA, and FPA. Here, FPA shows better performance for 750 samples of observation with min. deviation of 0.47.

TABLE 9: XGBoost parameters for regression analysis.

Cluster level	Parameters	Range	SHAP	Associate rule
Sample-1	Objective	Squared error	—	—
	L-Rate	0.032	0.04	0.11
	Max.Depth	2.86	13	0.11
	Bytree	0.51	0.7	0.28
Sample-2	Objective	Squared error	—	—
	L-Rate	0.067	0.048	0.12
	Max.Depth	3.34	14.15	0.12
	Bytree	0.58	0.80	0.31
Sample-3	Objective	Squared error	—	—
	L-Rate	0.070	0.053	0.14
	Max.Depth	3.80	14.99	0.12
	Bytree	0.66	0.90	0.34

Figure 13(a) represents the fitness function analysis for FPA with an initial pollen level of 2.33, enabling reactive power optimization. It determines the closeness of the optimized solution to the set point. Here, 5-fold cross-validation has been applied to find the mean fitness value at 4603.58. Figure 13(b) shows the current best individual. Here, a sample pollination size of 10% has been used for the evaluation of the distance between two conjugate pollens. In order to find the best local optimized solution, the average distance between two pollen has been maintained at 0.03×10^{-3} as shown in Figure 13(c). Fitness scaling for each individual is presented in Figure 13(d). Here, expectation vs. the raw score is presented. As noticed for a score level of 4603, a maximum expectation of 4 has been found. Similarly, each individual's fitness is presented in Figure 13(e), where a fitness value of 4000 is found for every individual. The selection function for each individual is presented in Figure 13(f). Here, individuals 6 and 9 have been taken into consideration for further analysis because these individuals have the highest number of children. Similarly, Figure 14 represents the flower pollination optimization with an initial pollen level of 3.07.

In order to analyze the data using Holt-Winters' exponential smoothing, the following steps have been followed:

- (i) Prepare the data set to find about trend and seasonal patterns
- (ii) Linearity inside the elements of the data set is established
- (iii) Identification of seasonal multiplication or addition for prediction over the data set

In this paper, the data set has been tested for both seasonal multiplication and addition models evaluating the effectiveness of the system. Table 13 shows the seasonal multiplication method under different weight conditions, and Table 14 shows the additive method under different weight conditions. Algo-

rithm 2 shows the pseudocode for reactive power prediction using Holt-Winters' exponential model.

Figure 15 represents the reactive power flow/demand over 3 different months. Three different patches of reactive power demand have been evaluated. Figure 16 shows the distributed seasonal trend of the reactive power samples. The trend pattern of the data reveals that the prediction is a linear combination of exponential data and sine waveform. Here, a residual of 1 unit has been observed. The seasonal pattern has an accuracy level of 0.98 units, which is 7.09% efficient as compared to ARIMA and 11.38% efficient as compared to the SARIMA model. As desired, the additive and multiplicative trends of the proposed model merge at the end of 3 months of data. However, there exists an RMSE level of 3.28 between the two sets of predictions.

Figure 17 represents Holt-Winters' additive and multiplicative trend analyses for reactive power prediction. The prediction has been done for 5 min intervals of time for the next 5 days. The error between additive and multiplicative forecasting is merged together and found to be zero. This indicates that the prediction is accurate for all seasonal variations.

Table 15 represents the prediction for an 8-day evening peak using HW model. As observed during 8-9 pm, maximum evening peak has been observed of a maximum magnitude of 917.415 MW on day 3, with an average demand of 843.39 MW on day 1. All the peak values were also validated against their actual values based on the history data set. Table 16 represents a comparative analysis of error matrices between the benchmarking model and the proposed model. The median analysis of different quantities such as nMBE, nMAE, nRMSE, and RMSE has been conducted. Normalized MBE, which evaluates the underestimation and overestimation of data inside a data set, is found to be negative for ARIMA, whereas the least positive value of 0.42 has been found for HW-GFPA under validation. Similarly, for the testing data set, it is 0.43. The normalized RMSE, which determines the variance present in between actual measured values and forecasted value, is 0.803 (least) for proposed HW-GFPA for validation and of 0.799 for testing.

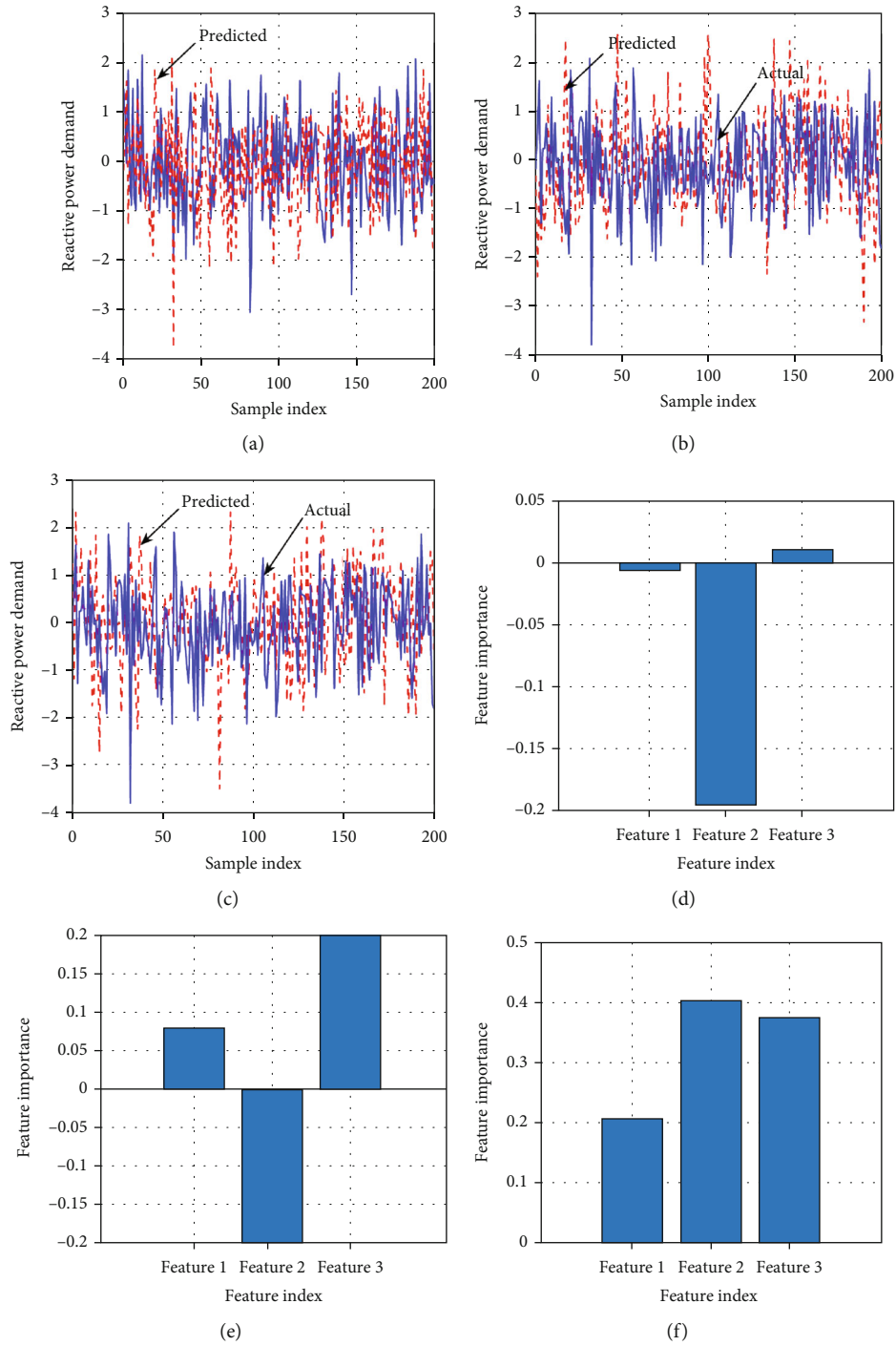


FIGURE 9: Reactive power forecasting: (a) sample-1 with a learning rate of 0.032, (b) sample-2 with a learning rate of 0.067, (c) sample-3 with a learning rate of 0.070, (d) feature index for sample-1, (e) feature index for sample-2, and (f) feature index for sample-3.

TABLE 10: Comparative analysis of different curve fitting techniques.

Type	R-square		Adj-R		DFE		RMSE	
	Ref.	Test	Ref.	Test	Ref.	Test	Ref.	Test
Sum of sine	0.50	0.4077	0.40	0.2073	50	68	20	29.93
Smoothing spine	0.99	0.725	0.30	0.312	25	36	50	42.13
Polynomial	0.20	0.218	0.30	0.33	70	83	50	37.41
Interpolate	1	1	NaN	NaN	1	0	0	NaN

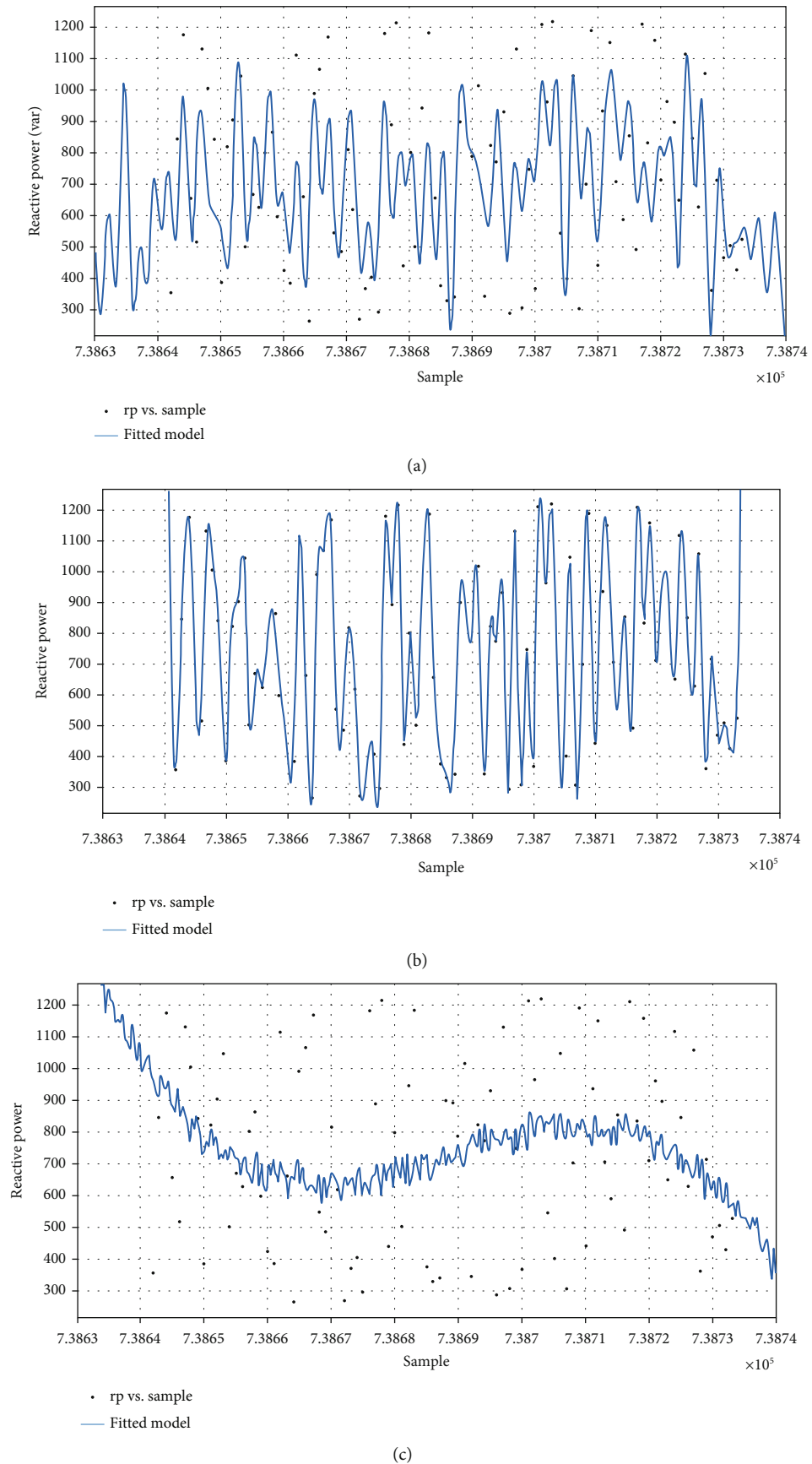
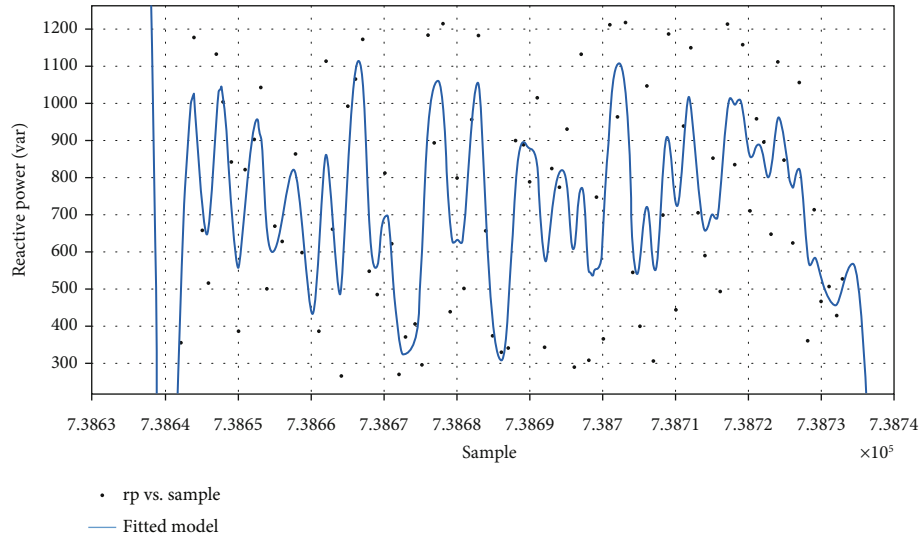


FIGURE 10: Continued.



(d)

FIGURE 10: Curve fitting model: (a) sum of sine curve, (b) smoothing spine, (c) polynomial, and (d) interpolate model.

```

Require: Define objective function,  $f(Q) = q(q_1, q_2, q_3, \dots, q_n)$ 
Ensure: the Initial optimal solution for the reactive power cost function, Initialize  $q_n$ 
for  $DR_{in} = 1 : IterationNo.count$  do
  for  $=1: N$  do
    Plot initial distribution for  $(Q_k^n, Q_k^{n+1})$ ; where  $Q$ =reactive power
    Random choose  $p, d, q$ 
    perform pollination eq. (27)
    find new optimized solution
    check the boundary eq. (28)
    update pollen
  end for
end for
    
```

ALGORITHM 1: Reactive power local sample evaluation.

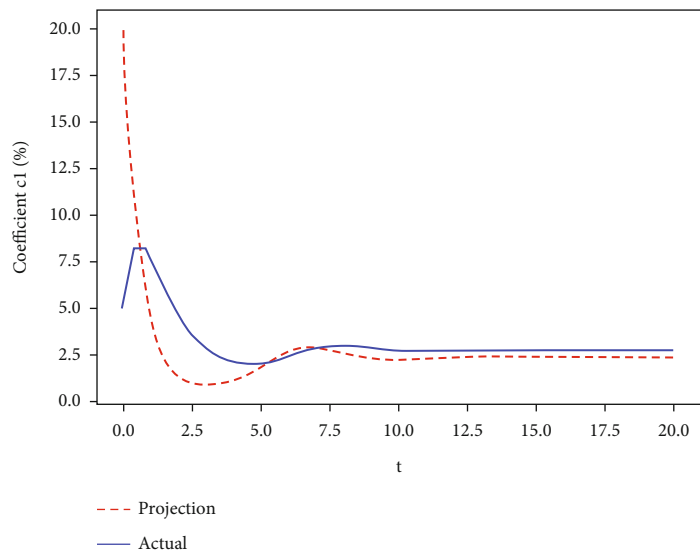
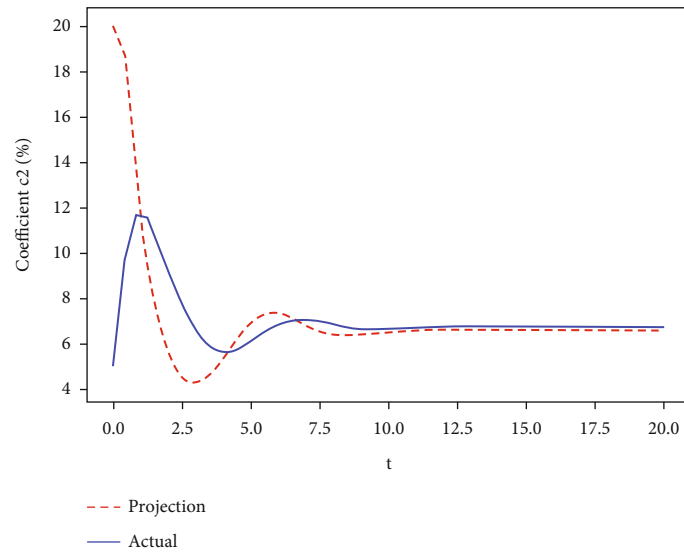


FIGURE 11: Performance of $c1$ against time in search space.

FIGURE 12: Performance of c_2 against time in search space.TABLE 11: Comparative analysis of different optimization algorithms for cost function: $c_1 = 0.17$.

Sample	Algo.	DRtn	Mean	E(DRtn)	Min
300	PSO	1.11	0.37	0.82	0.01
	GA	1.17	0.42	0.73	0.22
	FPA	1.15	0.44	0.71	0.01
450	PSO	3.81	0.49	4.64	2.15
	GA	5.47	2.98	2.98	5.47
	FPA	2.98	2.98	0.49	2.98
600	PSO	5.47	2.15	0.49	4.64
	GA	5.47	0.49	1.32	2.98
	FPA	1.32	5.47	5.47	5.47
750	PSO	0.49	0.49	2.15	4.64
	GA	1.32	3.81	4.64	4.64
	FPA	2.98	5.47	-0.34	2.15

TABLE 12: Comparative analysis of different optimization algorithms for cost function: $c_2 = 0.21$.

Sample	Algo.	DRtn	Mean	E(DRtn)	Min
300	PSO	-1.08	-1.69	-1.32	-1.99
	GA	-1.03	-1.65	-1.39	-1.82
	FPA	-1.04	-1.63	-1.41	-1.99
450	PSO	2.54	1.85	-2.28	-0.22
	GA	-0.22	2.54	-0.22	1.85
	FPA	-2.28	2.54	0.47	-0.22
600	PSO	2.54	-2.28	0.47	0.47
	GA	-0.90	-0.22	-2.28	-0.90
	FPA	-1.59	-2.28	2.54	1.85
750	PSO	-1.59	0.47	-0.90	-0.90
	GA	-2.28	0.47	0.47	1.16
	FPA	1.85	1.16	0.47	0.47

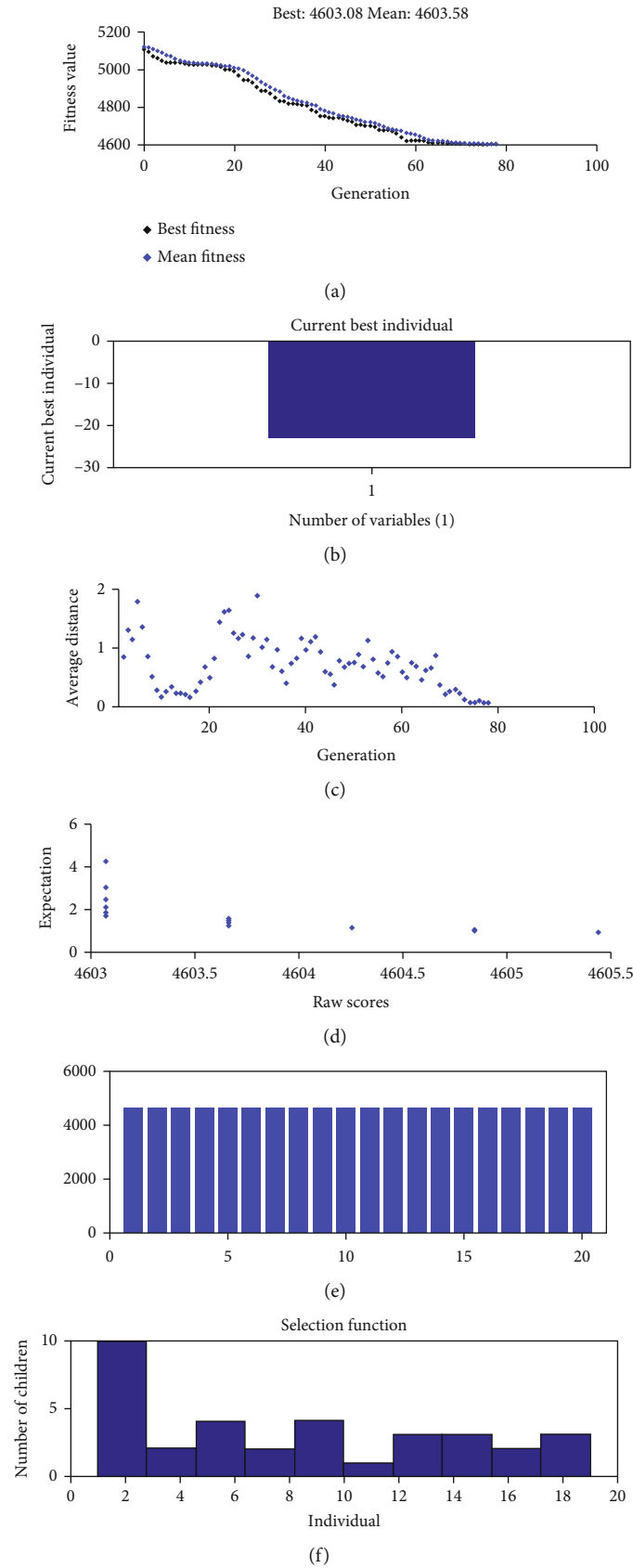


FIGURE 13: Flower pollination optimization with initial pollen level 2.33: (a) fitness vs. pollen generation, (b) current best individual, (c) average distance between individual, (d) fitness scaling, (e) fitness of individual, and (f) selection function for children/pollen.

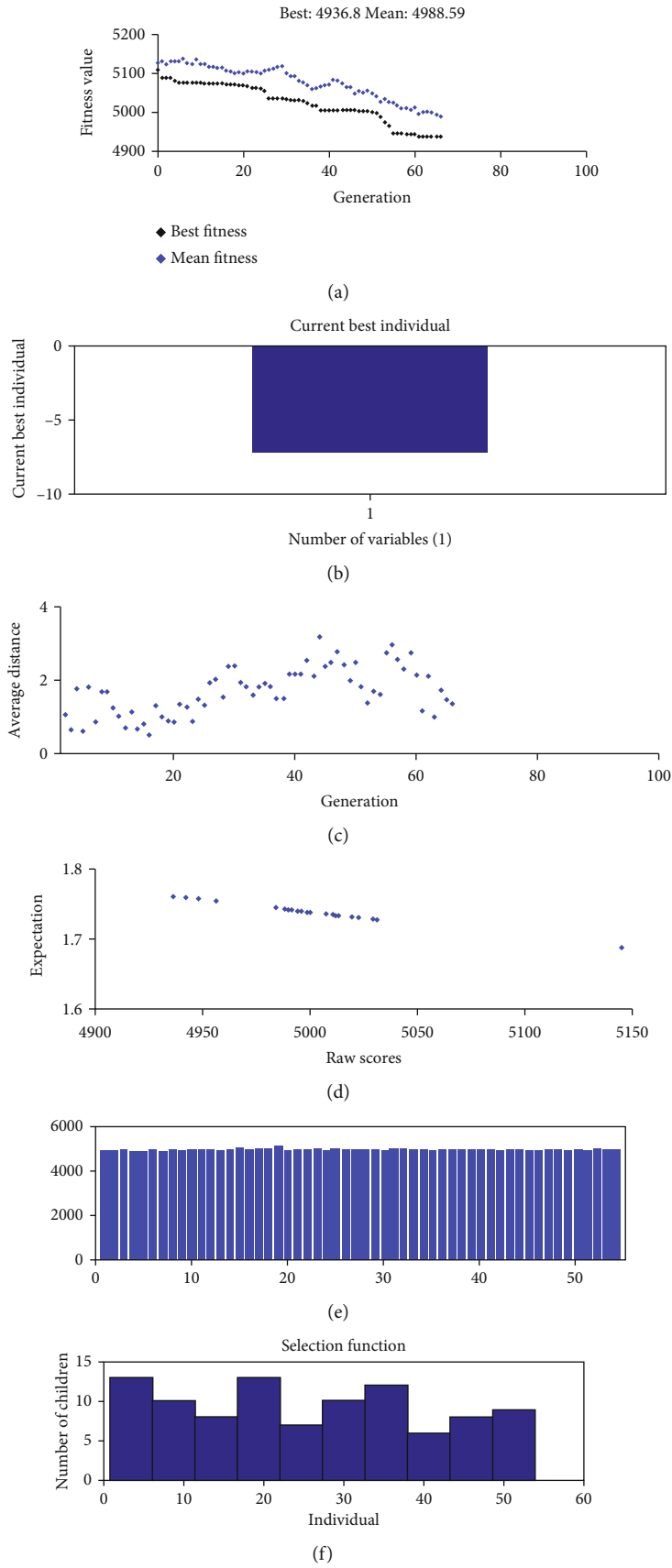


FIGURE 14: Flower pollination optimization with initial pollen level 3.07: (a) fitness vs. pollen generation, (b) current best individual, (c) average distance between individual, (d) fitness scaling, (e) fitness of individual, and (f) selection function for children/pollen.

TABLE 13: Seasonal multiplication method under different weight conditions.

Month	(0, 0.1)	Error	(0, 0.4)	Error	(0, 0.6)	Error	(0, 1)	Error	No. of sample
May-22	203, 66	8, 0.13	78, 71	8, 0.13	112, 57	11, 0.38	114, 78	84, 0.85	028
June-22	248, 04	11, 0.72	83, 59	8, 0.22	104, 63	11, 0.38	117, 34	79, 0.90	1047
July-22	132, 81	11, 0.64	117, 61	9, 0.73	97,61	13, 0.77	133, 64	81, 0.92	1047
Aug-22	141, 73	9, 0.98	118, 57	7, 0.88	103,72	11, 0.84	170, 55	79, 0.62	1093

TABLE 14: Seasonal additive method under different weight conditions.

Month	(0, 0.1)	Error	(0, 0.4)	Error	(0, 0.6)	Error	(0, 1)	Error	No. of sample
May-22	138, 43	98, 0.33	137, 49	77, 0.37	142, 59	63, 0.38	117, 29	6, 0.67	1028
June-22	126, 38	91, 0.48	128, 41	78, 0.44	147, 63	67, 0.16	121, 69	6, 0.38	1047
July-22	132, 40	93, 0.55	117, 39	92, 0.51	123, 52	61, 0.87	128, 68	7, 0.61	1047
Aug-22	135, 47	87, 0.51	133, 56	87, 0.58	134, 50	54, 0.66	133, 81	11, 0.74	1093

```

Require: Define map evaluation,  $f(Month) = m(m_1, m_2, m_3, \dots, m_4)$ 
Ensure: Monthly reactive power sample data
Value: Predict distinct boundary error
if Data no trend no season then
    Apply exponential smoothing
else
    if data no season then
        apply Holt-Winters' model
        Predict reactive power for each month
        add(error valuation, measure)
    end if
end if
    
```

ALGORITHM 2: Reactive power prediction: Holt-Winters' exponential model.

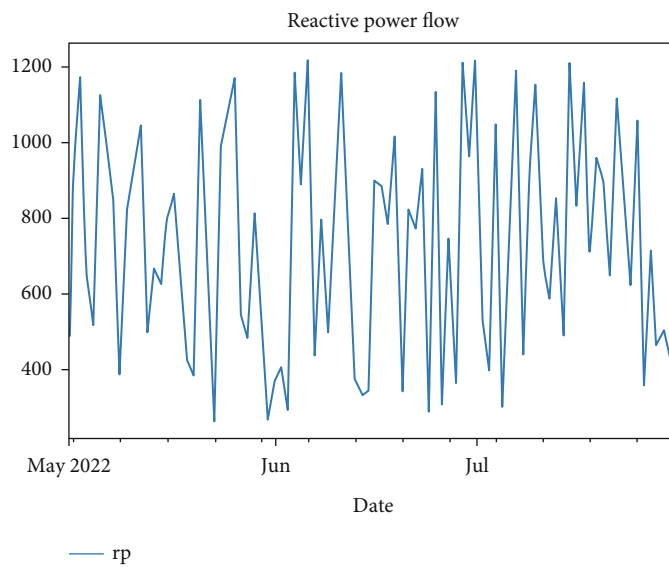


FIGURE 15: Reactive power sample over three different months.

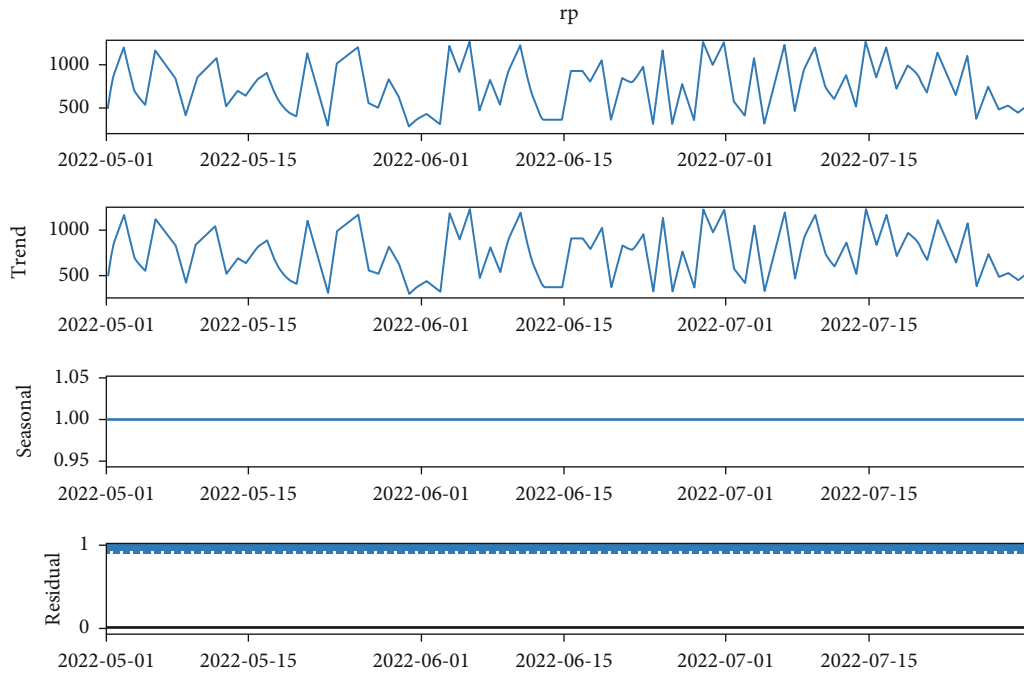


FIGURE 16: Distributed seasonal trend of reactive power sample over three different months.

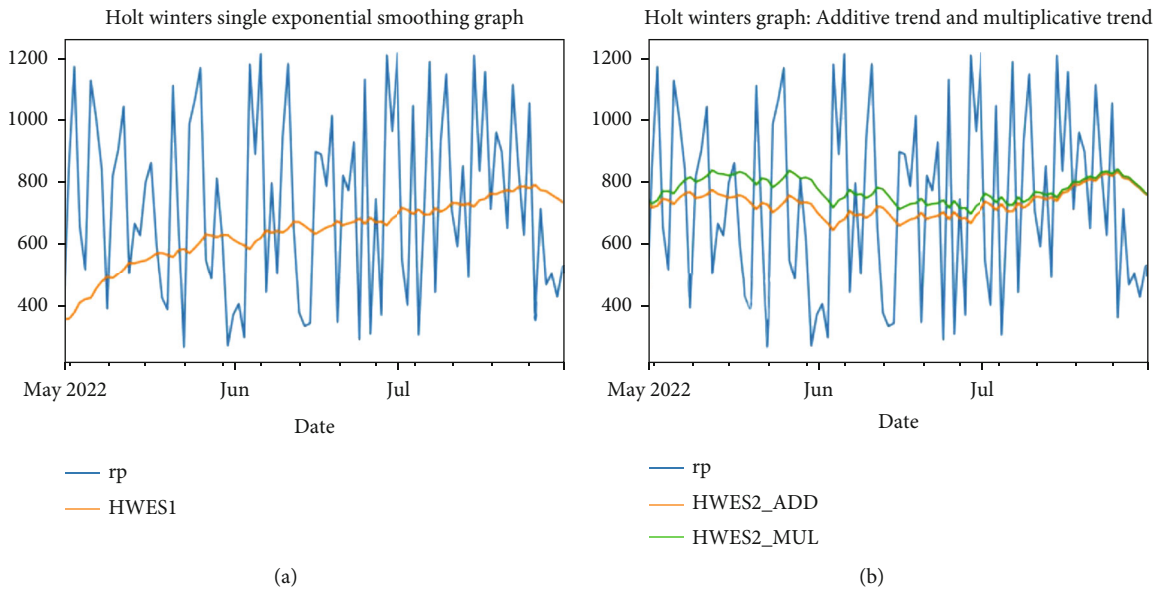


FIGURE 17: (a) Holt-Winters' single exponential smoothing graph. (b) Holt-Winters' additive and multiplicative trends.

TABLE 15: Prediction using HW model for 8-day evening peak analysis.

Sample time (PM)	D1	D2	D3	D4	D5	D6	D7	D8
5-6	937.285	746.834	886.056	917.408	870.676	864.006	881.823	800.852
6-7	874.765	838.080	771.428	899.362	814.018	777.647	833.484	723.414
7-8	792.601	777.439	701.350	767.383	920.583	918.354	742.873	765.647
8-9	850.266	726.181	917.415	785.319	689.934	908.772	882.846	825.250
9-10	762.078	920.868	916.909	745.847	895.593	728.439	770.720	837.624
Average	843.399	801.880	838.632	823.064	838.161	839.443	822.349	790.558

TABLE 16: Comparative analysis of error metrics between the benchmarking model and proposed model.

Performance	Metric	ARIMA	SARIMA	XGBoost	HW-GFPA
Validation	nMBE	0.68	0.61	0.49	0.42
	nMAE	0.830	0.828	0.824	0.822
	nRMSE	0.803	0.804	0.804	0.803
	RMSE	0.700	0.628	0.504	0.432
Testing	nMBE	-0.26	0.48	0.67	0.43
	nMAE	0.802	0.824	0.830	0.840
	nRMSE	0.813	0.805	0.803	0.799
	RMSE	-0.267	0.494	0.690	1.060

6. Conclusion

In this research paper, an attempt has been made to find the reactive power demand based on the usage of real power. Out of 1st-year load data, only three months of critical load reactive power has been taken into consideration from the lab experimental setup.

The problem formulation, as discussed in Section 2, reveals that reactive power forecasting can be modeled based on the load demand. After detailed modeling, it is understood that the global flower pollination algorithm can be best suited for the point-by-point optimization of the system. Again, in this research paper, Holt-Winters' model has been used for forecasting reactive power.

In order to validate the robustness of the proposed model, ARIMA and SARIMA models have been taken into consideration as the benchmarking models. The GFPA has been validated against PSO and GA regarding its effectiveness in predicting the time series data for two different cost functions. Again, under the result section, a detailed sample forecasting for three different situations such as normal load, critical peak demand during the day time, and critical peak demand during the night time has been presented as forecasted reactive power for 8 days in the month of August 2022, which is also validated in terms of MAPE and average error of actual data during that time interval for the same month.

The result analysis shows that the proposed Holt-Winters' model enabled with GFPA can predict effectively reactive power based on the active power against ARIMA and SARIMA models. In this research article, unit price commitment based on reactive power has not been taken into consideration; however, future work can be made based on this research matter.

Data Availability

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

Conflicts of Interest

There is no conflict of interest on behalf of all authors.

References

- [1] N. H. Khan, Y. Wang, D. Tian, M. A. Z. Raja, R. Jamal, and Y. Muhammad, "Design of fractional particle swarm optimization gravitational search algorithm for optimal reactive power dispatch problems," *IEEE Access*, vol. 8, pp. 146785–146806, 2020.
- [2] S. Granville, "Optimal reactive dispatch through interior point methods," *IEEE Transactions on Power Apparatus and Systems*, vol. 9, no. 1, pp. 136–146, 1994.
- [3] S. Alabd, M. H. Sulaiman, and M. I. Rashid, "Optimal power flow solutions for power system operations using moth-flame optimization algorithm," in *Proceedings of the 11th National Technical Seminar on Unmanned System Technology 2019*, Z. Md Zain, Ed., vol. 666 of Lecture Notes in Electrical Engineering, pp. 207–219, Springer, Singapore, 2021.
- [4] H. El-Desouki, A. El-Zonkoly, and G. Aly, "Optimal active and reactive power dispatch and excitation system controller design using Genetic Algorithm," *MEJ. Mansoura Engineering Journal*, vol. 28, no. 2, pp. 1–16, 2021.
- [5] M. S. Saddique, A. R. Bhatti, S. S. Haroon et al., "Solution to optimal reactive power dispatch in transmission system using metaheuristic techniques—status and technological review," *Electric Power Systems Research*, vol. 178, article 106031, 2020.
- [6] R. Jamal, B. Men, and N. H. Khan, "A novel nature inspired meta-heuristic optimization approach of GWO optimizer for optimal reactive power dispatch problems," *IEEE Access*, vol. 8, pp. 202596–202610, 2020.
- [7] M. H. Sulaiman, Z. Mustafa, M. M. Saari, and H. Daniyal, "Barnacles mating optimizer: a new bio-inspired algorithm for solving engineering optimization problems," *Engineering Applications of Artificial Intelligence*, vol. 87, article 103330, 2020.
- [8] S. Bhongade, A. Tomar, and S. R. Goigowal, "Minimization of optimal reactive power dispatch problem using BAT algorithm," in *2020 IEEE First International Conference on Smart Technologies for Power, Energy and Control (STPEC)*, pp. 1–5, Nagpur, India, 2020.
- [9] S. C. Swain, A. Gogoi, S. Sharma, R. Dash, and A. Acharya, "Some aspects of fuzzy logic controller for designing MPPT based SPV system," in *2018 International Conference on Recent Innovations in Electrical, Electronics Communication Engineering (ICRIEECE)*, pp. 3357–3361, Bhubaneswar, India, 2018.
- [10] M. Ettappan, V. Vimala, S. Ramesh, and V. T. Kesavan, "Optimal reactive power dispatch for real power loss minimization and voltage stability enhancement using artificial bee colony

- algorithm,” *Microprocessors and Microsystems*, vol. 76, article 103085, 2020.
- [11] T. Das, R. Roy, K. K. Mandal et al., “Optimal Reactive Power Dispatch Incorporating Solar Power Using Jaya Algorithm,” in *Computational Advancement in Communication Circuits and Systems*, vol. 575 of Lecture Notes in Electrical Engineering, Springer, Singapore.
- [12] T. T. Htay, K. K. Mon, and O. Z. Lin, “Optimal reactive power dispatch using particle swarm optimization algorithm for Yangon distribution network,” in , Article ID 012152IOP Conference Series: Earth and Environmental Science (IOP Publishing), vol. 463, Bangkok, Thailand, 2020.
- [13] M. Zand, H. R. Chamorro, M. A. Nasab, and S. H. Hosseinian, “Optimal reactive power dispatch using θ -social mimic optimization (θ -SMO),” *Journal of Intelligent Fuzzy Systems*, pp. 1–15, 2020.
- [14] A. Ballaji, R. Dash, V. Subburaj, J. R. Kalvakurthi, D. Swain, and S. C. Swain, “Design & development of MPPT using PSO with predefined search space based on fuzzy Fokker Planck solution,” *IEEE Access*, vol. 10, pp. 80764–80783, 2022.
- [15] K. Lenin, “Bamboo plant intellect deeds optimization algorithm for solving optimal reactive power problem,” in *Proceedings of the 7th International Conference on Advances in Energy Research*, Springer Proceedings in Energy, M. Bose and A. Modi, Eds., pp. 665–672, Springer, Singapore, 2021.
- [16] K. Lenin, “Levy Flight-Based White Wolf Algorithm for Solving Optimal Reactive Power Problem,” in *Advances in Communication and Computational Technology*, pp. 647–654, Springer, In, 2021.
- [17] N. Rana, M. S. Abd Latiff, and H. Chiroma, “Whale optimization algorithm: a systematic review of contemporary applications, modifications and developments,” *Neural Computing and Applications*, vol. 32, no. 20, pp. 16245–16277, 2020.
- [18] S. Mouassa, T. Bouktir, and A. Salhi, “Ant lion optimizer for solving optimal reactive power dispatch problem in power systems,” *Engineering Science and Technology, an International Journal*, vol. 20, no. 3, pp. 885–895, 2017.
- [19] A. M. Shaheen, R. A. El-Sehiemy, and S. M. Farrag, “A novel framework for power loss minimization by modified wind driven optimization algorithm,” in *2018 International Conference on Innovative Trends in Computer Engineering (ITCE)*, pp. 344–349, Aswan, Egypt, 2018.
- [20] W. S. Sakr, R. A. El-Sehiemy, and A. M. Azmy, “Adaptive differential evolution algorithm for efficient reactive power management,” *Applied Soft Computing*, vol. 53, pp. 336–351, 2017.