Alexandria Engineering Journal (2020) 59, 3111–3118

HOSTED BY

ELSEVIER

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Fractional ARIMA with an improved cuckoo search optimization for the efficient Short-term power load forecasting

Fei Wu [a], Carlo Cattani [b], Wanqing Song [a,*], Enrico Zio [c,d]

[a] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, PR China
[b] Engineering School, DEIM, Tuscia University, Viterbo 01100, Italy
[c] Politecn Milan, Dipartimento Energia, Via La Masa 34, I-20156 Milan, Italy
[d] MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France

Received 1 January 2020; revised 20 April 2020; accepted 22 June 2020
Available online 7 July 2020

KEYWORDS

Fractional Auto-regressive Integrated Moving Average;
Long-range dependence;
Hurst exponent;
Improved Cuckoo Search

Abstract Short-term power load forecasting plays a key role in power supply systems. Many methods have been used in short-term power load forecasting during the past years. A new short-term power load forecasting method is proposed in this study. First, the study represents a Fractional Auto-regressive Integrated Moving Average (FARIMA) model based on long-range dependence (LRD). The LRD model is governed by the Hurst exponent, which shows whether the model exhibits the LRD or not. Then, the study employs Cuckoo Search (CS) algorithm based on two parameters dynamic adjustment for parameter optimization of the forecasting model. As test problem, we use the real power consumption data, and test it for different forecasting models. Our results indicate that the FARIMA model and the improved optimization algorithm show relatively high accuracy and effectiveness in forecasting short-term power load.

1. Introduction

Short-term power load forecasting is the premise and foundation for efficient power dispatching systems or Programme Planning, reasonable arrangement of the “start-up”, “stop-up” and the maintenance of the generating units. It also ensures stable and effective operation of power systems. Hence it makes this study useful in engineering applications. The short-term power load forecasting usually refers to the daily load forecasting or weekly load forecasting, where the forecasting step size is either 30 min or 60 min. The importance of short-term power load forecasting is that it can produce accurate forecasting of daily production and domestic electricity consumption, effectively economize the time and cost of power supply departments.

With the gradual recognition of the importance of power load forecasting, many methods and models have been applied to the short-term power load forecasting since the 1960 s. Such

* Corresponding author.
E-mail addresses: cattani@unitus.it (C. Cattani), swqls@126.com (W. Song), enrico.zio@polimi.it (E. Zio).

Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2020.06.049

methods and models can be roughly divided into two categories: traditional forecasting methods and artificial intelligence (AI) forecasting methods [1–5]. The traditional forecasting methods include the ARMA model, Grey Model (GM), the Kalman filtering methods, etc. The ARMA and the ARIMA represent the classical time series forecasting models [6–8]. However, the obvious shortcoming of the classical models is that such models do not directly consider the influence of other random variables. The GM algorithm also represents a traditional power load forecasting algorithm [9–11]. In fact, many forecasting practices show that the GM is not effective for all problem statements. In order to use the Kalman filtering method [12–14], the mathematical model and noise statistics need to be identified. The filtering designed with inaccurate mathematical model and noise statistical characteristics can lead to large state estimation errors or even filter divergence. In 90ies, the developed techniques of artificial intelligence (AI) technology generated many novel AI forecasting methods, which were also applied to power load forecasting. The artificial neural networks (ANNs) are the most widely used method in AI forecasting methods [15–18]. However, ANNs have some disadvantages, such as local minimization, slow convergence speed, and different structure selection. Support Vector Regression (SVR) algorithm is widely used because of Disadvantages of ANN methods [19–22]. However, once the appropriate kernel function is chosen, the SVR model will become rather difficult to use. Deep-Learning methods have achieved the state-of-art performance in such forecasting fields as the Convolutional Neural Networks (CNN), the Recurrent Neural Networks (RNN) and some other improved forms [23,24].

Fractional Auto-regressive Integrated Moving Average model is known as FARIMA (*p, d, q*)[25]. This model can be regarded as the ARIMA model with fractional order characteristic. This characteristic enables the FARIMA model to exhibit simultaneously both long-range dependence (LRD) and short-range dependence (SRD). The series of power load data is investigated with LRD, so that the FARIMA model can be used for effective forecasting. In order to use the FARIMA model, one need to study the LRD of the forecasting series. Hence, this paper comes out the definition of the LRD and the essential parameter: Hurst exponent (H) to evaluate the LRD. If $0 < H < 0.5$, the time series exhibits a short-range dependence; while $0.5 < H < 1$, the time series is LRD. Many methods have been proposed to calculate H, like the Variance time method (V-T methods), the log–log correlogram and the Rescaled range analysis (R/S) [26–28]. Hurst exponent is not only used to evaluate the dependence, but also to derive the difference d-parameter in FARIMA model. Once the d-parameter is calculated and the parameters p, q are evaluated according to the ARIMA model, we can get the FARIMA forecasting model. However, the d-parameter is not constant. When the optimal parameter d is calculated, the forecasting accuracy can be easily improved. In order to get the optimal d-parameter, we propose an improved Cuckoo Search (ICS) algorithm [29,30]. Instead of simply isotropic random walk, the algorithm can search global optimum by Levy flights. The conventional CS algorithm searches for the optimal process by Levy flight [31,32]. In this study, the two parameters dynamic adjustment strategy is used to optimize the process to achieve the purpose of improving the algorithm. The improved optimization algorithm can effectively optimize the parameter d. So that we use the FARIMA algorithm combined with the improved optimization algorithm for forecasting the real power load data. Our forecasting simulation illustrates the superiority of this optimization model, compared with the AI forecasting models, and uses four methods: RBF, RNN, FARIMA and ICS + FAIRMA. The main novelty of this paper is summarized as the following steps:

(1) LRD model forecasting is used to solve the problem of nonlinear series preprocessing by judging the dependence of series.
(2) An improved cuckoo algorithm is proposed to optimize the important parameter of the model and improve the forecasting accuracy of power load.

The organization of this paper is as follows. The LRD description and the Hurst exponent computational details are given in Section 2. In Section 3 we show the deduction and the calculation processes of the FARIMA model. The key features of the ICS algorithm are given in Section 4. In Section 5 we describe the specific steps for optimizing d-parameter by using the ICS algorithm. Forecasting experiments based on real power load data are carried out in Section 6. Concluding remarks are given in Section 7.

2. Long-range dependence and hurst exponent estimation

Long-range dependence (LRD) is often associated with self-similar random processes or random fractals. When the auto-correlation refers to the scale behavior of the finite-dimensional distribution in a continuous-time or a discrete-time processes, The LRD denotes Long tail behavior of auto-correlation function for stationary time series [33–35].

The definition of LRD can be summarized as follows: for $X_1, X_2, \ldots, X_n$ are sampled observations of the given process$X(t)$, the mean and variance of $X(t)$read as $\mu = E(X_i)$ and$\sigma^2 = var(X_i)$, respectively. The auto-correlation between X_i and $X_j$can be described as:

$$\rho(k) = \frac{v(i,j)}{\sigma^2} \tag{1}$$

where$v(i,j) = \left[(X_i - \mu)(X_j - \mu)\right]$ is the auto-covariance between$X_i$andX_j.

If the following expression is true:

$$\sum_{K=-\infty}^{\infty} \rho(k) = \infty \tag{2}$$

The correlation is so slowly decaying that its partial sum diverges. This process can be considered as a process with long memory or exhibiting LRD. More formally, the LRD is defined as follows:

$$\rho(k)c\tilde{k}^{2H-2} \tag{3}$$

where $c$is the non-negative constant, $H$is the Hurst exponent which can be used to evaluate the LRD of time series.

Many methods have been proposed to calculate the Hurst exponent. For example, the variance time (V-T) method, the log-log correlogram and the Rescaled range analysis (R/S). In the following we use the R/S method for the Hurst exponent estimation. The specific steps are as follows:

For the observation series $\{X_t\}$, let its length T be divided into $k$adjacent subintervals of length n, where$T = kn$. The average value of each subinterval gives the new series $Y(i)$. The difference between the maximum and minimum values of the new series $Y(i)$and the standard deviation$S^2(n)$ can be calculated. The ratio between the two values read as follows:

$$(R/S)_n = (\max(Y_i) - \min(Y_i))/\sqrt{S^2(n)} \tag{4}$$

$$Y(i) = \sum_{j=1}^{n} X_j - \bar{X}_n \tag{5}$$

where $\bar{X}_n$and$S^2(n)$ represent the mean and the variance of the subinterval, respectively. The value R/S scales looks likecn^Has$n \to \infty$, where c is a constant, which is independent on n. The method is robust in the case of heavy tails. In addition, the heavier tails mean the closer to the expected true value.

3. The farima forecasting model

The model of the Fractional Auto-regressive Integrated Moving Average is usually expressed as the FARIMA (p, d, q). It was first proposed by Hosking [25]. Where parameters p, d and q represent the autoregressive order, the difference order and the moving average order, respectively.

The FARIMA model is capable to express long and short range dependences at the same time. So, it is often applied in network traffic forecasting. It can also be regarded as the special form of the ARIMA model.

The FARIMA model can be defined as follows:

If the given series$\{X_t\}$ is stationary and satisfies the difference equation:

$$\Phi(B)\Delta^d X_t = \theta(B)a_t \tag{6}$$

where$\{a_t\}$is the white noise sequence, $\Phi(B)$is the polynomial of order p, which denotes the Auto-regressive term, B is the backward moving operator, which satisfies the equation: $BX_t = X_{t-1}$, $\theta(B)$is the polynomial of order q. The expressions for $\Phi(B)$and $\theta(B)$have the form:

$$\Phi(B) = 1 - \Phi_1(B) - \Phi_2(B)^2 - \cdots - \Phi_p(B)^p \tag{7}$$

$$\theta(B) = 1 - \theta_1(B) - \theta_2(B)^2 - \cdots - \theta_p(B)^p \tag{8}$$

where$\Phi(B)$and$\theta(B)$do not have common zero's and do not equal zero for$|B| \leqslant 1$. Let$\Delta = (1 - B)$be the backward-shift operator, andΔ^ddenotes the fractional differencing operator, we have

$$\Delta^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \tag{9}$$

$$\binom{d}{k} = \Gamma(d+1)/[\Gamma(k+1)\Gamma(d-k+1)] \tag{10}$$

where Γdenotes the gamma function.

It is clear that for$d = 0$, the FARIMA (p, d, q) model represents the common ARMA (p, q). If $d \in (0, 0.5)$, then the LRD or persistence occurs in the FARIMA model. Especially, FARIMA (*0*, *d*, *0*) is the simplest and most fundamental form of FARIMA model. According to Section 2, parameter$d$is related to LRD, so it can be drawn from H: $d = H - 0.5$.

Taking into account the above equations and d-parameter definition, the fundamental steps of FARIMA can be summarized as follows:

Select real power load data for a period of time;
According to the LRD of power load series, the primary FARIMA prediction model is established.
Use the fractional difference filtering for power load data;
The ARMA model is used to identify the power load data after the differential filtering;
Following the principle of the ARMA model, the parameters p and q are evaluated.
Use a period of past power load data into the model to predict the future power load

4. Improved optimization algorithms for forecasting model

Although the FARIMA model is considered as an effective method for power load forecasting, its accuracy still remains poor. As a result, an optimization algorithm is used to optimize both parameters of the forecasting model respectively. In the following we will use the improved Cuckoo Search algorithm (ICS) to optimize the parameters of the FARIMA model.

The Cuckoo Search (CS) algorithm is the natural heuristic algorithm, proposed by Yang, X. and Deb, S. in 2009. The CS algorithm simulates certain species of cuckoo's parasitic brood process to solve global optimization problems. The CS algorithm adopts Levy flight search mechanism, which makes the optimization algorithm more effective. The one thousand trajectories simulation of the Levy flights are shown in Fig. 1.

The CS algorithm follows three basic rules adopted by the Cuckoo bird:

Rule 1: Only one egg at the time is laid down by cuckoo. Cuckoo dumps its egg in a randomly chosen nest.

Rule 2: only the best nest of hatching cuckoo eggs will be preserved for the next generation.

Rule 3: The number of available host nests is fixed. Eggs laid by a cuckoo bird is discovered by the host bird with a probability $p_a \in [0, 1]$.

On the basis of these three rules, the Levy flight has the form:

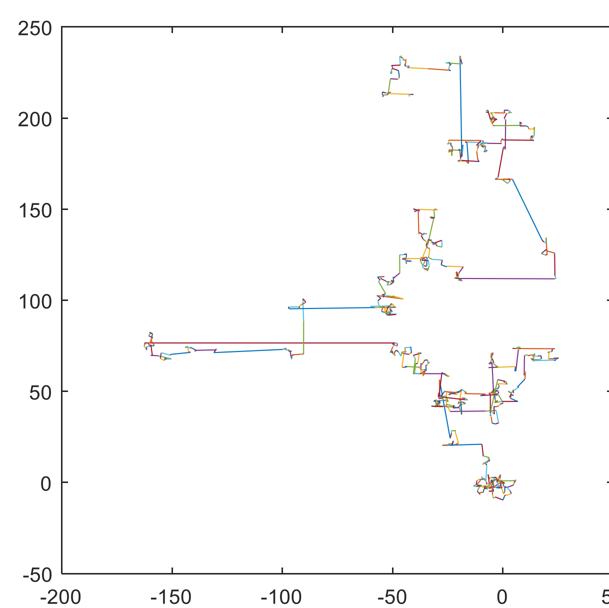


Fig. 1 The 1000 Trajectories simulation of the Levy flight.

$$X_i^{t+1} = X_i^t + \alpha \oplus levy(\beta) \tag{11}$$

where i is the X_i^t solution for t generation t, $\alpha > 0$ denotes a step size control factor, $\oplus$ denotes point to point multiplication, $levy(\beta)$ denotes the random search path:

$$Levy(\beta) \sim u = t^{-\lambda}, 1 < \lambda \leqslant 3 \tag{12}$$

In Eq. (12), $Levy(\beta)$ is calculated by the Mantegna Method [36,37], as:

$$Levy(\beta) \sim \frac{\varphi \times \mu}{|v|^{1/\beta}} \tag{13}$$

where μ and v are the random numbers, which obey the normal distribution/ For $\beta = 1.5$, φ can be calculated by the following:

$$\varphi = \left(\frac{\Gamma(1+\beta) \times \sin\left(\pi \times \frac{\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2} \times \beta \times 2^{\frac{\beta-1}{2}}\right)} \right)^{1/\beta} \tag{14}$$

In addition, during the iteration process of CS, after updating the nest according to Eq. (14), the discovery probability p_a is compared to the uniformly distributed random numbers $rand(0,1)$. If $rand > p_a$, the cuckoo eggs are discovered and abandoned, the next X_i^{t+1} is randomly changed. Finally, calculate the fitness function:

$$Y_i^t = fitness(X_i^t) \tag{15}$$

Then nests with better fitness values will be reserved, and it is written as X_i^{t+1}.

There are four parameters in the CS algorithm: nests number N, discovery probability p_a, step size α and λ in Levy flight. Due to the number of nests and Levy flight parameter are basically determined after initialization, the strategy of dynamic adjustment of two parameters is equal to adjusting the α and p_a. First of all, the original step size is also fixed, which reduces the prediction accuracy when the number of iterations is adjusted. In the following, we propose a method of exponentially decreasing α with the number of iterations to adjust the α. The specific equation is as follows:

$$\alpha'(t) = \alpha(t) \cdot \max\_it \cdot \exp\left(-\frac{i}{\max\_it}\right) \tag{19}$$

where i is the current number of iterations, $\max\_it$ is the maximum iterations. Then there is the improvement of discovery probability p_a. We change its probability to two-stage changing probability. In the first half of the max_it, the negative sinusoidal adaptive decline is adopted, while in the second half of the max_it, the negative cosine adaptive decline is adopted to search for the best nest location. The equation is as follows:

$$p_\alpha = \begin{cases} p_{\alpha\max} - \sin\left(\frac{\pi}{2} \cdot \frac{i-1}{\max\_it-1}\right)(p_{\alpha\max} - p_{\alpha\min}), i \leqslant \frac{\max\_it}{2} \\ p_{\alpha\max} - \cos\left(\frac{\pi}{2} \cdot \frac{i-1}{\max\_it-1}\right)(p_{\alpha\max} - p_{\alpha\min}), i \geqslant \frac{\max\_it}{2} \end{cases} \tag{20}$$

where $p_{\alpha\max}$ and $p_{\alpha\min}$ the maximum and minimum of the discovery probability. Therefore, the optimization ability of CS algorithm is improved by adjusting the above two parameters.

5. d-Parameter optimization

In this section, the improved optimization algorithm ICS will be used to optimize the model parameter, because it plays an important role in FARIMA prediction. The process of the d-parameter optimization is based on the following steps.

Step 1: Initialize the objective function $f(d)$;

Step 2: Set the CS algorithm parameters, i.e., number of nests N, the discovery probability $p_a = 0.25$, the maximum and minimum of the discovery probability $p_{\alpha\max} = 0.5$ and $p_{\alpha\min} = 0$, step size α and λ in Levy flight, initial position $nest_0 = [d_1^0, d_2^0, ..., d_N^0]$ and set the first Root Mean Square Errors (RMSE), which are calculated by substituting the parameters into the objective function as the fitness values $fbest_0 = [y_1^0, y_2^0, ..., y_N^0]$, set maximum iterations $\max\_it$.

Step 3: Choose the best nest position of the previous generation d_h^i $(1 \leqslant i \leqslant \max\_it, h \in [1, N])$, calculate step size. According to Eq. (19), then search for the nest location based on the Levy flight. Record new nest as s, take s into objective function and then calculate new fitness value $fnew$. Compare $fnew$ with the optimal fitness value of the previous generation y_h^i, if $fnew < y_h^i$, the nest location based on Levy flight become the new best nest position.

Step 4: Update p_a according to Eq. (14), set the random number $rand(0,1)$. If $rand > p_a$, choose the nest position randomly, and replace it with the worst position.

Step 5: If the number of iterations is satisfied, stop searching; otherwise, step back to Step 3.

Step 6: Output the nest position with the least fitness value as optimum solution d_{best}.

The flowchart of the d-parameter optimization by ICS optimization algorithm is shown in Fig. 2.

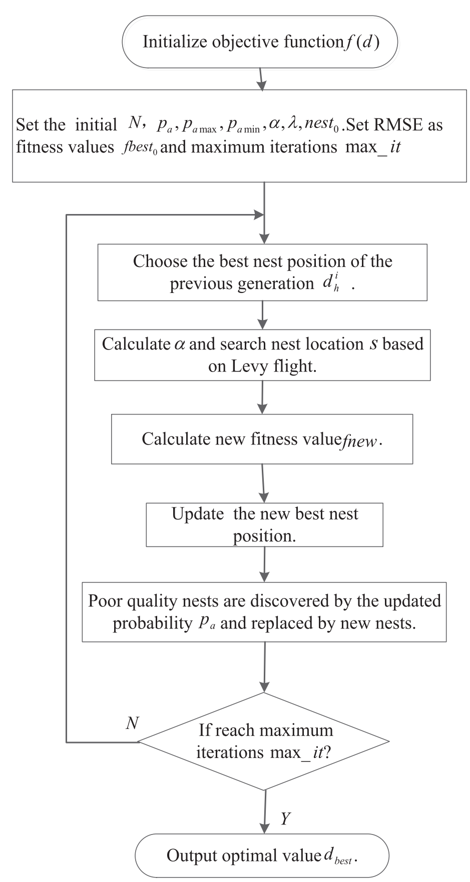


Fig. 2 The flowchart of the d-parameter optimization by ICS optimization algorithm.

6. Experiments and numerical analysis

In this section, we use the forecasting model and the optimization algorithm to forecast the power load data. Our simulation refers to the actual data provided by the EIRGIRD GROUP (Ireland). The simulation utilizes the data collected in February 2018. In order to show the superiority of the ICS method and to compare with the existing mainstream methods, this experiment also takes ANN method: RBF model, Deep Learning method: RNN model and the original method: FARIMA model into forecasting.

Since this paper focuses on short-term power forecasting, the forecasting steps are set to 24 h to forecast next 24 h. Fig. 3 shows the forecasting results and actual data of the power load on weekdays and weekends using the above four forecasting methods.

On the basis of the forecasted data, this experiment first calculated the Mean absolute error (MAE) and the mean absolute percentage errors (MAPE) for different models. Two conventional error criteria can show the specific effect of each method, They are computed as follows:

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\mathrm{PL}_i^f - \mathrm{PL}_i^r\right|}{\left|\mathrm{PL}_i^r\right|}\times 100\% \tag{21}$$

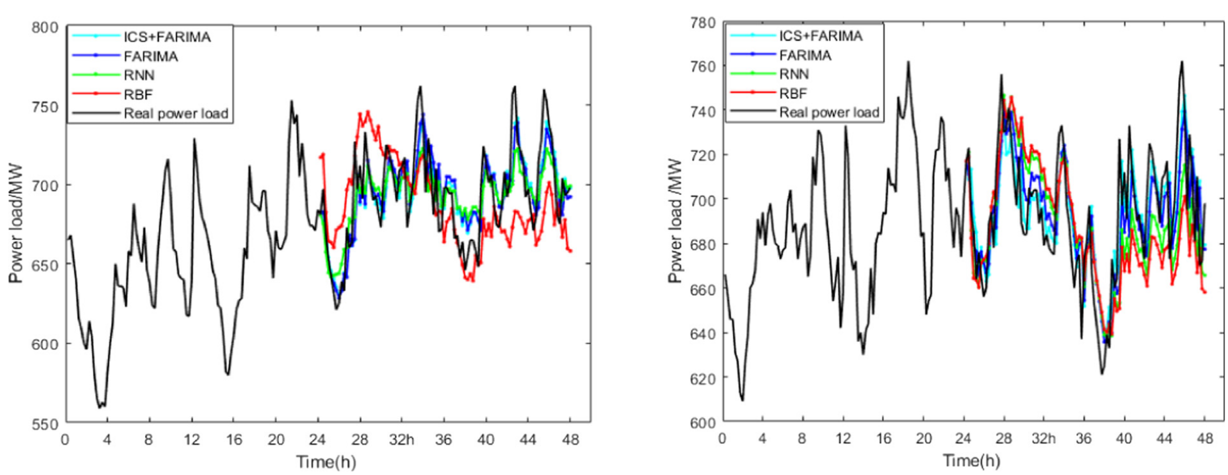
$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{PL}_i^f - \mathrm{PL}_i^r\right| \tag{22}$$

where N is the number of power load data; PL_i^f is the predicted power load data of validation sample i; and PL_i^r is the real power load data of validation sample i. In this experiment, we have four methods to forecast power load data on weekdays (Tuesday and Wednesday) and weekends (Saturday and Sunday) respectively. The detailed results are shown in Table 1.

Table shows that the short-time power load forecasting in a week as given by four models (RBF, RNN, FARIMA and FARIMA + ICS).

In order to describe the accuracy of forecasting more comprehensively, we have considered in the next experiment 4, 11, 18 and 25 during Sundays in February into forecasting. Moreover we have used three new error criteria to describe forecasting effect of the above four models. The three new error criteria are: root mean-squared error (RMSE), normalised MAPE (NMAPE) and normalised RMSE (NRMSE) defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{PL}_i^f - \mathrm{PL}_i^r\right)^2} \tag{23}$$



(a) weekdays (b) weekends

Fig. 3 Forecasted power consumption during weekdays and weekends.

Table 1 Comparison of the four forecasting methods during weekdays and weekends.

Methods	Error criteria	Weekdays		Weekends		Average
		Tues.	Wed.	Sat.	Sun.	
RBF	MAPE	0.03229	0.03132	0.03285	0.03231	0.03219
	MAE	22.3595	20.2877	22.9815	21.9443	21.8933
RNN	MAPE	0.02533	0.02658	0.02817	0.02791	0.02699
	MAE	17.7124	17.3113	19.6599	19.0431	18.4317
FARIMA	MAPE	0.02272	0.02241	0.02388	0.02068	0.02242
	MAE	15.8421	14.7176	16.6083	14.1739	15.3355
ICS + FARIMA	MAPE	**0.02173**	**0.02153**	**0.02282**	**0.01924**	**0.02132**
	MAE	**15.1634**	**14.2012**	**15.8475**	**13.1878**	**14.5998**

Table 2 Comparison of the four forecasting methods during each Sunday in February.

Methods	Error criteria	1st	2nd	3rd	4th	
		Feb.4	Feb.11	Feb.18	Feb.25	Average
RBF	NMAPE	3.0329	2.8022	3.3786	3.2308	3.1111
	RMSE	27.1687	23.5084	28.9891	27.2223	26.7221
	NRMSE	3.9301	3.6152	4.0937	4.0568	3.9239
RNN	NMAPE	2.4276	2.3138	2.5972	2.7908	2.5323
	RMSE	20.9917	20.0663	23.0569	24.0105	22.0313
	NRMSE	2.9812	3.0108	3.2623	3.5023	3.1891
FARIMA	NMAPE	2.2624	2.2397	2.3886	2.0687	2.2398
	RMSE	20.7051	19.1651	20.8811	18.7117	19.8657
	NRMSE	2.9559	2.8652	2.9827	2.7201	2.8809
ICS + FARIMA	NMAPE	**2.1865**	**2.1168**	**2.2834**	**1.9243**	**2.1277**
	RMSE	**19.7475**	**18.3004**	**19.8345**	**18.0592**	**18.9854**
	NRMSE	**2.8101**	**2.7207**	**2.8426**	**2.6202**	**2.7484**

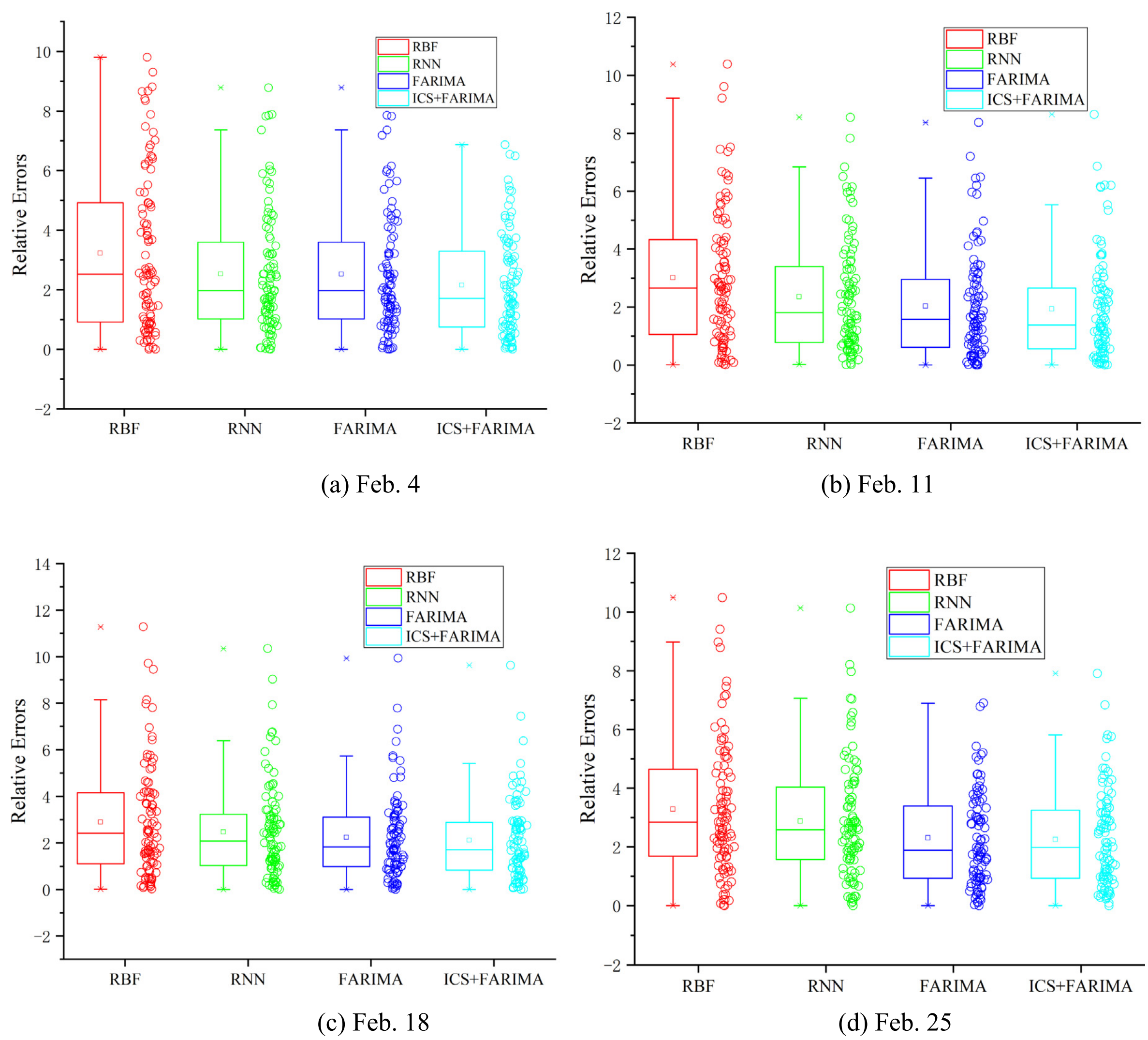


Fig. 4 Box-scatter plot representation of relative errors during the four Sundays.

$$\text{NMAPE}(\%) = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\text{PL}_i^f - \text{PL}_i^r\right|}{\text{PL}_N} \times 100 \tag{24}$$

$$\text{NRMSE}(\%) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\text{PL}_i^f - \text{PL}_i^r}{\text{PL}_N}\right)^2} \times 100 \tag{25}$$

where $PL_N$is the nameplate capacity of the power load system. The forecasting results of the four Sundays are brought into the above formulas for calculation. Then the final results are shown in the Table 2. From the data in Table 2, the proposed ICS + FARIMA combining method performance the best result in short-term power load forecasting compared with the other methods. This experiment also gives a graphical insight about the forecasting accuracy of the above four forecasting methods. The Box-scatter plot representation of relative errors (which denote each independent of NMAPE) during the four Sundays are shown in Fig. 4.

7. Conclusion

A new short-term power load forecasting model with a LRD and an algorithm for model optimization are proposed in this paper. The FARIMA model can be regarded as the special form of the ARIMA model with the LRD. The difference order d-parameter of the model can be deduced from the Hurst exponent H, which is used to evaluate the LRD. The d-parameter plays an important role in the forecasting formulas: excellent global optimization capability of the improved optimization algorithm allows us to optimize the parameter. The experiments, based on the real power load data illustrates the effectiveness of the optimized forecasting model and the proposed ICS + FARIMA optimized forecasting model has been shown to have the best performance than the other three forecasting models.

Declaration of Interest Statement

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Acknowledgments

This project was funded by the Key Project of Science and Technology Commission of Shanghai Municipality (Grant No. 18511101600) and the Natural Science Foundation of Shanghai (Grant No. 14ZR1418500).

References

[1] W.J. Lee, J. Hong, A hybrid dynamic and fuzzy time series model for mid-term power load forecasting, Int. J. Electr. Power Energy Syst. 64 (Jan 2015) 1057–1062.
[2] R. Hu, S.P. Wen, Z.G. Zeng, T.W. Huang, A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm, Neurocomputing 221 (Jan 2017) 24–31.
[3] C. Guan, P.B. Luh, L.D. Michel, Y.T. Wang, P.B. Friedland, Very Short-Term Load Forecasting: Wavelet Neural Networks With Data Pre-Filtering, IEEE Trans. Power Syst. 28 (1) (Feb 2013) 30–41.
[4] W.Q. Song, C. Cattani, C.H. Chi, Multifractional Brownian Motion and Quantum-Behaved Particle Swarm Optimization for Short Term Power Load Forecasting: An Integrated Approach, Energy 194 (1) (2020) 116847.
[5] N. Liu, Q.F. Tang, J.H. Zhang, W. Fan, J. Liu, A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids, Appl. Energy 129 (Sep 2014) 336–345.
[6] A. Kavousi-Fard, F. Kavousi-Fard, A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA, J. Exp. Theor. Artif. Intell. 25 (4) (2013) 559–574.
[7] J.K. Liang, C. Carlo, W.Q. Song, Power Load Prediction Based on Fractal Theory, Adv. Mathematical Phys. (2015) 1–6.
[8] M.H. Amini, A. Kargarian, O. Karabasoglu, ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation, Electr. Power Syst. Res. 140 (Nov 2016) 378–390.
[9] H.R. Zhao, S. Guo, An optimized grey model for annual power load forecasting, Energy 107 (Jul 2016) 272–286.
[10] S. Bahrami, R.A. Hooshmand, M. Parastegari, Short term electric load forecasting by wavelettransform and grey model improved by PSO (particle swarm optimization) algorithm, Energy 72 (Aug 2014) 434–442.
[11] J. Shi, Z.H. Ding, W.J. Lee, Y.P. Yang, Y.Q. Liu, M.M. Zhang, Hybrid Forecasting Model forVery-Short Term Wind Power Forecasting Based on Grey Relational Analysis and Wind Speed Distribution Features, IEEE Trans. Smart Grid 5 (1) (Jan 2014) 521–526.
[12] E.A. Blood, B.H. Krogh, M.D. Ilic, Electric Power System Static State Estimation through Kalman Filtering and Load Forecasting, IEEE Power & Energy Society General Meeting 1–11 (2008) 5601–5606.
[13] C.N. Ko, C.M. Lee, Short-term load forecasting using SVR (support vector regression)-basedradial basis function neural network with dual extended Kalman filter, Energy 49 (Jan 2013) 413–422.
[14] C. Guan, P.B. Luh, L.D. Michel, Z.Y. Chi, Hybrid Kalman Filters for Very Short-Term LoadForecasting and Prediction Interval Estimation, IEEE Trans. Power Syst. 28 (4) (Nov 2013) 3806–3817.
[15] L. Hernandez, C. Baladron, J.M. Aguiar, B. Carro, A.J. Sanchez-Esguevillas, J. Lloret, Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks, Energies 6 (3) (Mar 2013) 1385–1408.
[16] R.A. Hooshmand, H. Amooshahi, M. Parastegari, A hybrid intelligent algorithm based shortterm load forecasting approach, Int. J. Electr. Power Energy Syst. 45 (1) (Feb 2013) 313–324.
[17] D.K. Chaturvedi, A.P. Sinha, O.P. Malik, Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network, Int. J. Electr. Power Energy Syst. 67 (May 2015) 230–237.
[18] F. Yu, X.Z. Xu, A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network, Appl. Energy 134 (Dec 2014) 102–113.
[19] Z. Y. Hu, Y. K. Bao, and T. Xiong, "Electricity Load Forecasting Using Support Vector Regression with Memetic Algorithms", Scientific World Journal, 2013, Art. no. 292575.
[20] A. Kavousi-Fard, H. Samet, F. Marzbani, A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting, Expert Syst. Appl. 41 (13) (Oct 2014) 6047–6056.
[21] A.S. Ahmad et al, A review on applications of ANN and SVM for building electrical energy consumption forecasting, Renew. Sustain. Energy Rev. 33 (May 2014) 102–109.
[22] A. Selakov, D. Cvijetinovic, L. Milovic, S. Mellon, D. Bekut, Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank, Appl. Soft Comput. 16 (Mar 2014) 80–88.
[23] O. Abedinia, D. Raisz, N. Amjady, Effective prediction model for Hungarian small-scale solar power output, IET Renew. Power Gener. 11 (13) (Nov 2017) 1648–1658.

[24] H. Shi, M.H. Xu, R. Li, Deep Learning for Household Load Forecasting-A Novel Pooling Deep RNN, IEEE Trans. Smart Grid 9 (5) (Sep 2018) 5271–5280.

[25] Hosking, J.R.M, Fractional differencing, Biometrika, 1981, p. 165.

[26] H.Y. Wang, W.Q. Song, E. Zio, F. Wu, Y.J. Zhang, Remaining Useful Life Prediction for Lithium-ion Batteries Using Fractional Brownian Motion and Fruit-fly Optimization Algorithm, Measurement (2020), in press.

[27] J. Barunik, L. Kristoufek, On Hurst exponent estimation under heavy-tailed distributions, Physica a-Statistical Mechanics Appl. 389 (18) (Sep 2010) 3844–3855.

[28] L.T. Yan, The fractional derivative for fractional Brownian local time with Hurst index large than 1/2, Math. Z. 283 (1–2) (Jun 2016) 437–468.

[29] Y.D. Gao, V. Francesco, L. Ming, et al, Multi-Scale Permutation Entropy Based on Improved LMD and HMM for Rolling Bearing Diagnosis, Entropy 19 (4) (2017) 176.

[30] W.Q. Song, X.X. Cheng, C. Cattani, E. Zio, Multifractional Brownian Motion and Quantum-Behaved Partial Swarm Optimization for Bearing Degradation Forecasting, Complexity (2020), https://doi.org/10.1155/2020/854 3131.

[31] G.G. Wang, A.H. Gandomi, X.J. Zhao, et al, Hybridizing harmony search algorithm with cuckoo search for global numerical optimization, Soft. Comput. 20 (1) (2016) 273–285.

[32] X.T. Li, M.H. Yin, Modified cuckoo search algorithm with self adaptive parameter method, Inf. Sci. 298 (2015) 80–97.

[33] M. Li, "Long-range dependence and self-similarity of teletraffic with different protocols at the large time scale of day in the duration of 12 years: Autocorrelation modeling", Physica Scripta, Accepted, 2020. https://doi.org/10. 1088/1402-4896/ab82c4

[34] M. Li, Multi-fractional generalized Cauchy process and its application to teletraffic, Physica A (2020), https://doi.org/10.1016/j.physa.2019.123982.

[35] Song W., Liu H., Zio E., Generalized Cauchy process: difference iterative forecasting model, Chaos, Solitons & Fractals, 2020, in press.

[36] M. R.N., "Fast, accurate algorithm for numerical simulation of Levy stable stochastic processes," vol. 49, ed: Physical Review E, 1994, p. 4677.

[37] H. Liu, W.Q. Song, M. Li, A. Kudreyko, E. Zio, "Fractional Lévy stable motion: Finite difference iterative forecasting model", Chaos, Solitons & Fractals 133 (4) (2020) 109632.