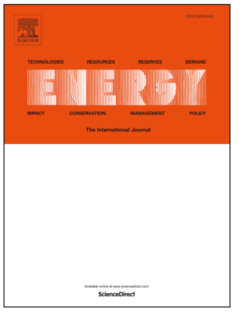
Journal Pre-proof



Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption

Weiheng Jiang, Xiaogang Wu, Yi Gong, Wanxin Yu, Xinhui Zhong

PII: S0360-5442(19)32474-0

DOI: https://doi.org/10.1016/j.energy.2019.116779

Reference: EGY 116779

To appear in: *Energy*

Received Date: 10 July 2019

Revised Date: 12 December 2019

Accepted Date: 14 December 2019

Holt–Winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption

Weiheng Jiang[a], Xiaogang Wu[a], Yi Gong[b], Wanxin Yu[a], Xinhui Zhong[a]

[a]*College of Communication Engineering, Chongqing University, Chongqing 400044, China.*
[b]*Southern University of Science and Technology, Shenzhen 518055, China.*

Abstract

Electricity consumption forecasting is essential for intelligent power systems. In fact, accurate forecasting of monthly consumption to predict medium- and long-term demand substantially contributes to the appropriate dispatch and management of electric power systems. Most existing studies on monthly electricity consumption forecasting require large datasets for accurate prediction, which is severely undermined when scarce data are available. However, in practical scenarios, data is not always sufficient, thereby hindering the accurate forecasting of monthly electricity consumption. The Holt–Winters exponential smoothing allows to accurately forecast periodic series with relatively few training samples. Based on this method, we propose a hybrid forecasting model to predict electricity consumption. The fruit fly optimization algorithm is used to select the best smoothing parameters for the Holt–Winters exponential smoothing. We used electricity consumption data from a city in China to comprehensively evaluate the forecasting performance of the proposed model compared to similar methods. The results indicate that the proposed model can substantially improve the prediction accuracy of monthly electricity consumption even when few training samples are available. Moreover, the computation time of the proposed model is the shortest among the evaluated hybrid benchmark algorithms.

Keywords: Monthly electricity consumption forecasting, Holt–Winters exponential smoothing, Fruit fly optimization algorithm

1. Introduction

Electricity consumption forecasting is critical for the successful management of current electric power systems, and it has attracted the attention from academy and industry [1]. In particular, highly accurate medium- and long-term electricity consumption forecasting is key for dispatching and planning in power systems. However, medium- and long-term electricity consumption has uncertain, complex, and nonlinear relationships depending on several external factors, such as political conditions, economy, human activity, and population behaviours [2], thus hindering its accurate forecasting. To improve the accuracy of medium- and long-term electricity consumption forecasting, various prediction methods and models have been proposed.

Medium- or long-term electricity consumption forecasting methods can be classified into two types, namely, stand-alone and hybrid models [1] [3]. The stand-alone models are further divided according to the underlying technique into three categories, namely, statistical [4] [5] [6] [7], computational intelligence [8] [9] [10] [11] [12], and mathematical programming models [13] [14]. Hybrid models combine different techniques: statistical–statistical [15], statistical–computational intelligence [16] [17] [18], computational intelligence–computational intelligence [19] [20], and statistical– mathematical programming [13] [14]. Furthermore, stand-alone models can be either linear or nonlinear [21], and they represent the more extensively investigated models.

Statistical stand-alone models mainly include exponential smoothing [4] [22], regression analysis [5] [23], and time-series methods [24] [25]. For instance, Tsekouras et al. [23] proposed nonlinear multivariable regression for medium-term energy consumption forecasting of power systems in an annual basis. In [25], Pakistan's electricity demand was estimated by applying the smooth transition autoregressive model.

Computational intelligence stand-alone methods include metaheuristic [8] [26], machine-learning [9] [10] [11] [12] [27], knowledge-based [28] [29], and uncertainty-based methods [30] [31]. For instance, a regression convolutional neural network has been used to extract features from data, and then a regressive support vector machine trained with these features predicted electricity consumption, achieving a low prediction error [9]. Monthly electricity demand has also been forecasted by using other artificial neural networks [32] [33]. In addition, instead of point forecasting, a novel long-term probability forecasting model based on fuzzy Bayesian theory and expert prediction has been proposed to predict the Chinese per-capita electricity consumption and its variation from 2010 to 2030 [29].

Unlike stand-alone models, hybrid models are more promising given their excellent representation ability of nonlinear and random factors. Ju et al. [19] employed chaotic gravitational search to determine three free parameters in a support vector regression (SVR) model and substantially improve prediction performance. To simultaneously reduce load forecasting error and enhance stability, Singh et al. [14] proposed a multiobjective algorithm based on the follow-the-leader algorithm and integrating an artificial neural network to notably improve forecasting accuracy. Likewise, in [20], a novel approach for monthly electricity consumption forecasting by integrating the wavelet transform and a neuro-fuzzy system is proposed, achieving accurate forecasting. He et al. [34] proposed a method of probability density forecasting based on least absolute shrinkage and a selection operator-quantile regression neural network and evaluated the prediction accuracy through empirical analyses on datasets from the Guangdong province in China and California in the United States. In [16], a bottom-up approach is combined with hierarchical linear models for long-term electricity consumption forecasting in an industrial region considering various energy efficiency scenarios. The model was used to generate long-term point and probability distribution forecasts from 2015 to 2050.

Although several studies have considered medium-term electricity consumption forecasting through a variety of models, they are only suitable if many historical data samples are available, retrieving low forecasting accuracy otherwise, especially in methods based on computational intelligence and machine learning. In China, however, the informatisation of electrical companies in some areas is recent, limiting the availability of aggregate historical electricity consumption data. Therefore, accurately forecasting electricity consumption with relatively few samples remains a current and challenging problem.

We propose a model integrating the fruit fly optimization algorithm (FOA) to perform Holt–Winters exponential smoothing (HW) for accurate medium-term electricity consumption forecasting, especially when insufficient historical data are available. HW is commonly used to forecast seasonal time-series data [35] [36]. As its performance mainly depends on the selection of appropriate smoothing parameters, we use the FOA for achieving global optimization [37] [2]. Therefore, we apply the FOA for HW parameter selection, establishing the proposed FOA-MHW model. To the best of our knowledge, this is the first enhanced HW model that can retrieve high accuracy despite the scarcity of samples from real datasets. Moreover, the algorithm execution time is the shortest among hybrid benchmark algorithms for prediction.

The rest of this paper is organised as follows. Section 2 introduces the HW method and FOA, and then the corresponding hybrid forecasting model is introduced. Section 3 details the dataset used in this study and reports the performance evaluation of the proposed model under varying lengths of training datasets compared to similar approaches. Finally, we draw conclusions in Section 4.

2. Proposed FOA-MHW model

In this section, we describe the HW model and FOA to then introduce the proposed FOA-MHW model.

2.1. Holt–Winters exponential smoothing

Exponential smoothing is an important time-series forecasting method, which basically pre-processes original data to eliminate randomness and improves the importance of recent data during prediction. The processed data are called smoothing values, which are used to establish the forecasting model and finally predict target values [38]. However, simple exponential smoothing cannot suitably remove the randomness of time series, and more capable approaches such as the HW method have been developed [35] [36]. A previous analysis has shown that, even with a small number of samples, HW can provide accurate forecasting results [38]. HW can be classified into four types, namely, multiplicative, additive, modified, and extended, according to the calculation of seasonal indices [4]. The multiplicative HW (MHW) is the most popular method, because its calculation of seasonal indices is more robust and accurate. Hence, we adopt MHW as the primary forecasting model in the proposed FOA-MHW.

MHW simultaneously forecasts trend, stationarity, and seasonal components of time series. The basic form of the HW forecasting model is given by

$$\hat{y}_{t+m} = (T_t + b_t m) S_{t+m-L}, \tag{1}$$

where $\hat{y}_{t+m}$ is the m-time-ahead forecasted data, t is the period of training data, and three variables from time series are included: stationarity T_t, trend b_t, and seasonality S_t. T_t corrects the time series by ruling out seasonal factors, b_t corrects trend values to eliminate seasonal interference, and S_t forecasts seasonal indices to exclude random interference [38]. The recursive expressions of these three variables are respectively defined as

$$T_t = \alpha \frac{y_t}{S_{t-L}} + (1-\alpha)(T_{t-1} + b_{t-1}), \tag{2}$$

$$b_t = \beta(T_t - T_{t-1}) + (1-\beta) b_{t-1}, \tag{3}$$

$$S_t = \gamma \frac{y_t}{T_t} + (1-\gamma) S_{t-L}, \tag{4}$$

where α, β, and γ are smoothing parameters of the HW model with values in $[0,1]$, and y_t represents the observation data. Based on the analysis of first-order exponential smoothing [38], larger values of α, β, and γ lead to smaller influence of long-term historical electricity values on prediction. Hence, as the prediction results depend on the smoothing parameters, their proper selection was an important part of this study.

The initialisation of equations (2)–(4) is given by

$$T_0 = \frac{1}{n} \sum_{t=1}^{n} y_t, \tag{5}$$

$$S_{0k} = \frac{\bar{y_k}}{T_0}, \tag{6}$$

$$b_0 = 0, \tag{7}$$

where

$$\bar{y_k} = \frac{1}{n/L} \sum_{t=1 \wedge t=jL+k}^{n} y_t, (k = 1, ..., L,\ j = 1, 2, ..., n/L - 1), \tag{8}$$

T_0 is the average value of the training data, and S_{0k} $(k = 1, ..., L)$ is the initial value of the seasonal index.

2.2. Fruit fly optimization algorithm

The FOA is an intelligent optimization algorithm developed by Pan [37] and often used to search the global optimal solution for parameter tuning based on the food finding behaviour of the fruit fly. The fruit fly provides representative modelling because it is superior to other species regarding sensing and perception, particularly osphresis and vision. The flowchart of the evolution of a fruit fly population is shown in Figure 1, and the details of the algorithm are summarised below [39].

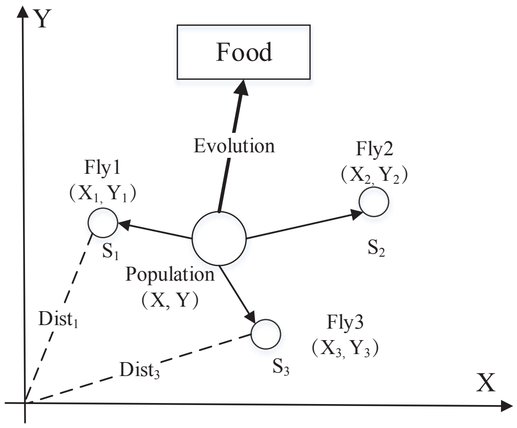


Figure 1: Evolution of fruit fly population [2].

(1) Parameter setting: Population size *sizepop*, maximum number of iterations *maxgen*, random flight range *FR*, and initial fruit fly swarm location (x_0,y_0).

(2) Initialisation: Give random direction and distance for food search using osphresis by an individual fruit fly.

$$X_i = x_0 + rand(FR) \tag{9}$$

$$Y_i = y_0 + rand(FR) \tag{10}$$

(3) Calculate smell concentration judgment value: Calculate the distance from a fruit fly to the origin ($Dist_i$) and the smell concentration judgment value (S_i).

$$Dist_i = \sqrt{X_i^2 + Y_i^2} \tag{11}$$

$$S_i = 1/Dist_i \tag{12}$$

(4) Calculate smell concentration: Feed S_i into fitness function to obtain smell concentration ($Smell_i$) of individual fruit fly location (i.e. equation (5)).

$$Smell_i = Fuction(S_i) \tag{13}$$

(5) Find the best individual: Determine the fruit fly with maximal smell concentration (finding the maximum) among the fruit fly swarm.

$$[bestSmell, bestIndex] = max(Smell_i) \tag{14}$$

bestSmell represents the highest smell concentration among the current fruit fly swarm and *bestIndex* indicates the fruit fly with the highest smell concentration in the fruit fly swarm.

(6) Fruit fly swarm movement: Select the best fruit fly individual, *bestIndex*, and maintain the best direction for the fruit fly swarm to fly towards that best location using vision.

$$smellBest = bestSmell \tag{15}$$

$$x_0 = X_{bestIndex} \tag{16}$$

$$y_0 = Y_{bestIndex} \tag{17}$$

(7) Population evolution: Perform iterative optimization and repeat the steps 2 to 6. If the smell concentration is not better than that at the previous iteration or the maximum number of iterations, *maxgen*, is reached, terminate the algorithm.

2.3. Proposed forecasting model

As mentioned in section 2.1, the forecasting performance of HW is mainly determined by smoothing parameters α, β, and γ. Therefore, we use the FOA to search the optimal values of these parameters and improve the accuracy of monthly electricity consumption forecasting.

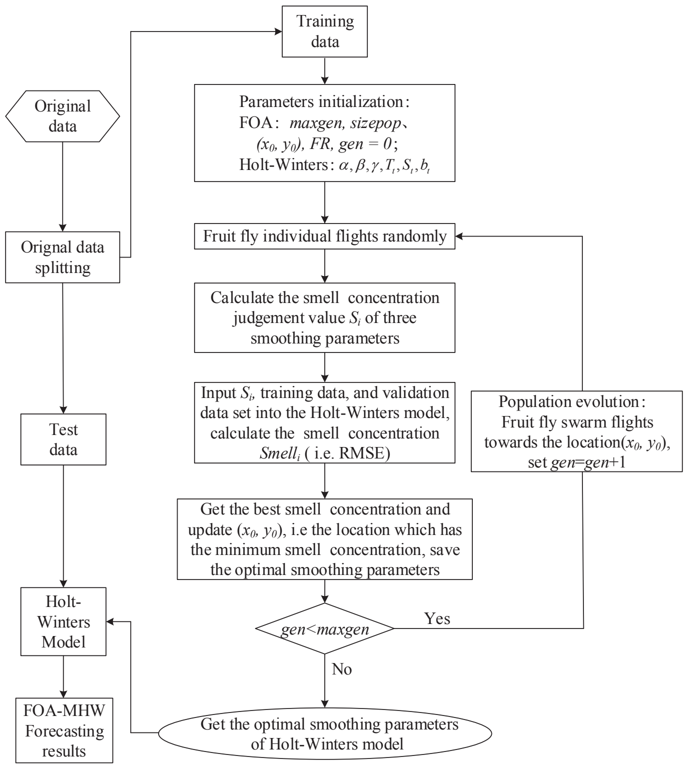


Figure 2: Flowchart of proposed FOA-MHW model.

The flowchart of the proposed FOA-based smoothing parameter selection for HW, called FOA-MHW, is shown in Figure 2, and its algorithm is detailed below.

(1) Dataset splitting: Divide sample dataset into training and test datasets. Samples from the last period of the training dataset constitute the validation dataset, which is used to conduct the optimal smoothing parameter selection for HW in step 5.

(2) Parameter initialisation: Initialise population size *sizepop*, maximum number of iterations *maxgen*, random flight ranges (FX^{α}, FY^{α}), (FX^{β}, FY^{β}), and (FX^{γ}, FY^{γ}), and initial fruit fly swarm locations (x_0^{α}, y_0^{α}), (x_0^{β}, y_0^{β}), and (x_0^{γ},y_0^{γ}), which correspond to the initial locations of parameters α, β, and γ, respectively.

Hence, there are three fruit flies searching the optimal smoothing parameters for the HW model.

(3) Fruit fly swarm flight: Every fruit fly randomly searches for food within a predetermined distance, with the locations of the fruit fly swarms being (X_i^{α}, Y_i^{α}), (X_i^{β}, Y_i^{β}), and (X_i^{γ}, Y_i^{γ}):

$$X_i^{\alpha} = x_0^{\alpha} + FX^{\alpha}, Y_i^{\alpha} = y_0^{\alpha} + FY^{\alpha}, \tag{18}$$

$$X_i^{\beta} = x_0^{\beta} + FX^{\beta}, Y_i^{\beta} = y_0^{\beta} + FY^{\beta}, \tag{19}$$

$$X_i^{\gamma} = x_0^{\gamma} + FX^{\gamma}, Y_i^{\gamma} = y_0^{\gamma} + FY^{\gamma}. \tag{20}$$

(4) Calculate smell concentration judgment value: This value, S_i, for each fruit fly is given by

$$S_i^{\alpha} = \alpha = \frac{1}{D_i^{\alpha}}, \tag{21}$$

$$S_i^{\beta} = \beta = \frac{1}{D_i^{\beta}}, \tag{22}$$

$$S_i^{\gamma} = \gamma = \frac{1}{D_i^{\gamma}}, \tag{23}$$

where

$$D_i^{\alpha} = 1/\sqrt{X_i^{\alpha 2} + Y_i^{\alpha 2}}, \tag{24}$$

$$D_i^{\beta} = 1/\sqrt{X_i^{\beta 2} + Y_i^{\beta 2}}, \tag{25}$$

$$D_i^{\gamma} = 1/\sqrt{X_i^{\gamma 2} + Y_i^{\gamma 2}}, \tag{26}$$

with D_i^{α}, D_i^{β}, and D_i^{γ} being the distance from the corresponding fruit fly to the origin. We feed the smell concentration judgment value (i.e. candidate values of α, β, and γ) to train the HW model and obtain accurate forecasting results. Then, we calculate the root-mean-square error (RMSE) between the forecasting results and validation dataset.

(5) Calculate smell concentration: The RMSE is used as smell concentration value ($Smell_i$), and it is calculated as follows:

$$RMSE = \sqrt{\sum_m (y_m - \hat{y_m})^2}, \tag{27}$$

where y_m represents the data in the validation dataset and $\hat{y_m}$ represents the forecasting results.

(6) Find the best individual: Determine the three fruit flies with the highest smell concentration (minimum of $Smell_i$) among their fruit fly swarms, as equation (28).

$$[bestSmell, bestIndex] = min(Smell_i) \tag{28}$$

(7) Fruit fly swarm movement: Select the best three fruit fly individuals with *bestIndex* and maintain the smell concentration judgment value. Considering the best direction, the fruit fly swarm uses vision to fly towards the best location as follows:

$$x_0^{\alpha} = X_{bestIndex}^{\alpha}, y_0^{\alpha} = Y_{bestIndex}^{\alpha}, \tag{29}$$

$$x_0^{\beta} = X_{bestIndex}^{\beta}, y_0^{\beta} = Y_{bestIndex}^{\beta}, \tag{30}$$

$$x_0^{\gamma} = X_{bestIndex}^{\gamma}, y_0^{\gamma} = Y_{bestIndex}^{\gamma}. \tag{31}$$

(8) Population evolution: Repeat steps 3 to 7. If the smell concentration is not better than that at the previous iteration or the maximum number of iterations, *maxgen*, is reached, retrieve the best smell concentration judgment values that would correspond to optimal smoothing parameters α_{opt}, β_{opt}, and γ_{opt}, and go to step 9.

(9) Forecasting: Feed the optimal smoothing parameters and training dataset into the HW model to obtain forecasting results.

3. Numerical evaluation and results

We selected two examples to verify the effectiveness of the proposed FOA-MHW model, namely, monthly electricity consumption forecasting of a southern city in China and the more specific case of the telecommunications and television industry in that city. To determine the performance of the proposed algorithm, five related algorithms were considered as benchmarks: seasonal index (SI) model [40], MHW model with default parameters (MHW-default) [38], FOASVR [39], GASVR [41], and GA-MHW models. The GA-MHW is a variant that we developed to evaluate the algorithm execution time. Basically, a genetic algorithm is used to search the optimal smoothing parameters for the HW model for improving forecasting accuracy. In addition, to verify the performance gain of our FOA-MHW model on a small training dataset, we evaluated varying numbers of training samples. The performance metric for evaluation is the mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\hat{y}_t - y_t}{y_t}\right| \tag{32}$$

where N is the number of forecasting datapoints, which is 6 in our simulation, y_t is the real electricity consumption, and $\hat{y}_t$ is the forecasting result.

The experiment was implemented on MATLAB R2016b using an in-house software executed on a computer with Intel Core i5-6500 3.2 GHz CPU, 16 GB RAM, and running the Microsoft Windows 7 Pro operating system.

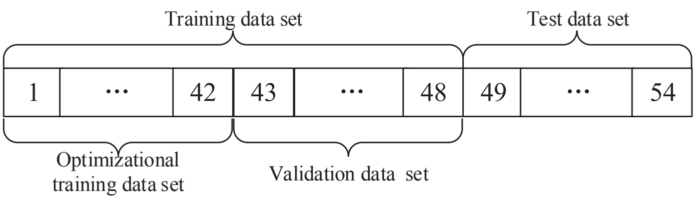


Figure 3: Details of dataset used in this study.

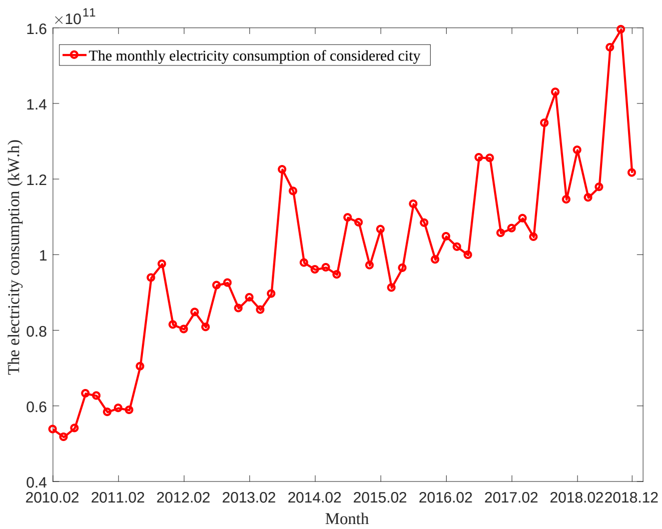


Figure 4: Monthly electricity consumption of city considered in this study.

3.1. Dataset

The sample dataset was provided by a Chinese energy supply company and includes the monthly electricity consumption of Chongqing city in China from January 2010 to December 2018 and the consumption from the telecommunications and television industry of Chongqing city during the same period. As the electricity meters of users collect and report consumption every 2 months to the power management system, we only have 6 datapoints per year. Therefore, there are 54 datapoints in the sample dataset. We selected the data from January 2010 to December 2017 as training dataset (i.e. first 48 datapoints), and the remaining data as test dataset (i.e. last 6 datapoints). We further divided the training dataset into optimization training and validation datasets, with the former containing the first 42 datapoints and the latter containing the last 6 datapoints. The optimization training and validation datasets were used to conduct the FOA for determining the smoothing parameters in the HW model, while the test dataset was used to evaluate the forecasting performance. Details of the sample dataset are shown in Figure 3.

3.2. Forecasting results from city

The actual monthly electricity consumption in the study period is shown in Figure 4, and the forecasting process is detailed as follows.

(1) Dataset processing: According to the method presented in 3.1, we obtained the training and test datasets.

(2) Forecasting: We implemented the process of FOA-MHW model described in 2.3 and initialised *maxgen* and *sizepop* to 40 and 50, respectively. FX^{η} and FY^{η} followed an independent uniform distribution in $[5, 10]$, while $\eta \in \alpha, \beta, \gamma$ and the initial locations of the fruit fly swarms followed independent uniform distributions in $[0, 1]$ (i.e. $x_0^{\eta}, y_0^{\eta} \in [0, 1], \eta \in \alpha, \beta, \gamma$).

(3) Comparative analysis: We implemented the SI, HW-default, and GASVR models in the same dataset, where the default smoothing parameters of the HW model were $\alpha = 0.2, \beta = 0.1$, and $\gamma = 0.6$ [38].

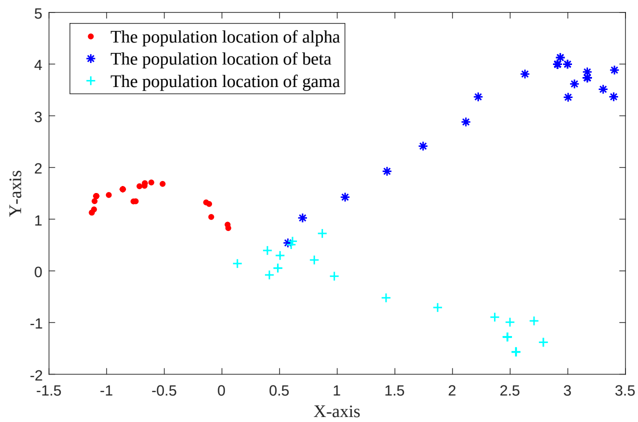


Figure 5: Search route of fruit fly swarms for α, β, and γ in city example.

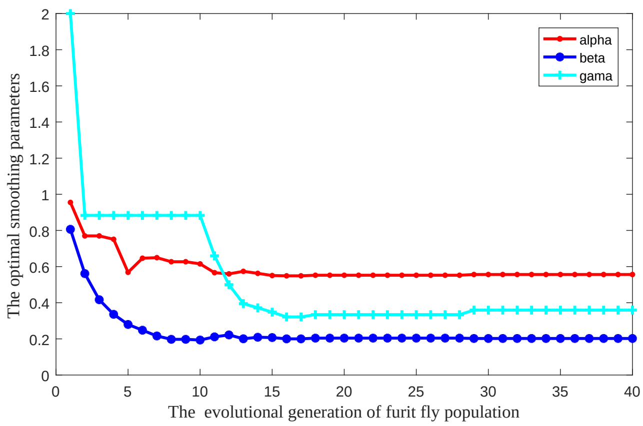


Figure 6: Update of optimal smoothing parameters in city example.

By performing the forecasting process described above, we obtained the search route of the three fruit fly swarms shown in Figure 5 and the update of the optimal smoothing parameters shown in Figure 6. The three swarms can find the food quickly, thus suggesting fast convergence of the smoothing parameter values. The optimal smoothing parameters selected by the FOA were $\alpha = 0.5562, \beta = 0.2022$, and $\gamma = 0.3590$. Such optimal values would be very difficult to obtain through manual tuning.

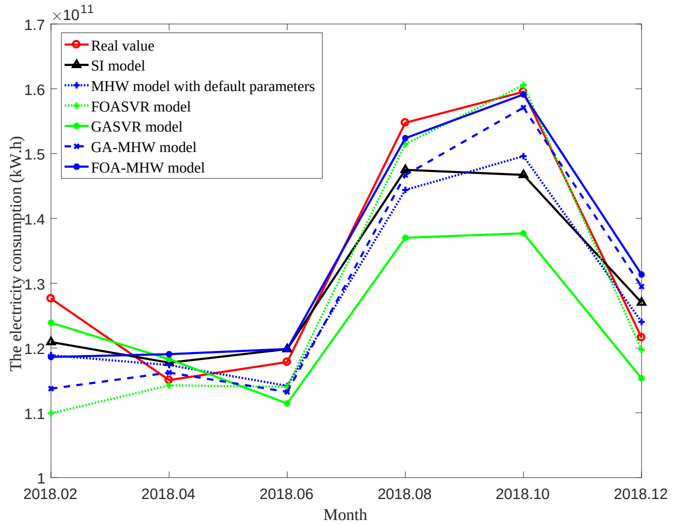


Figure 7: Forecasting results of evaluated models in city example.

The forecasting results of the evaluated models are shown in Figure 7, whereas the relative error per month during 2018 and total MAPE are listed in Table 1. The proposed FOA-MHW model has the smallest MAPE of 3.65%, followed by the SI, MHW-default, and GA-MHW models, which achieve better performance than the FOASVR and GASVR models. Hence, although the monthly electricity consumption is highly variable, the optimal parameter selection of the FOA improves prediction for the FOA-MHW model to suitably overcome randomness. As the order of the SI and MHW-default models is lower than that for the FOA-MHW model, their performance is inferior. In addition, the machine learning-based models, FOASVR and GASVR, cannot be suitably trained on a small dataset, making their prediction performance unsatisfactory.

The algorithm execution times of the evaluated models are listed in Table 2. The execution time of the stand-alone models (i.e. SI and MHW-default) is much shorter than that of hybrid models. In fact, intelligent algorithms demand a long time to find optimal parameters for subsequent forecasting. Still, among the hybrid models, the proposed FOA-MHW has the shortest execution time. Moreover, the execution time of the models based on statistical learning (i.e. FOA-MHW and GA-MHW) is shorter than that of the models based on machine learning (i.e. FOASVR and GASVR).

3.3. *Forecasting results from telecommunications and television industry*

To further evaluate the performance of the proposed FOA-MHW model and gain insights on its operation, we conducted another simulation using electricity consumption data from a specific industry (i.e. telecommunications and television), with the consumption being shown in Figure 8. The consumption in this industry has a better periodicity than that in the city (see Figure 4). The forecasting process and parameter settings were the same as those reported in section 3.2.

The search routes of the three fruit fly swarms are shown in Figure 9, and the update of the optimal smoothing parameters is shown in Figure 10. The three fruit fly swarms can find the optimal smoothing parameters very quickly, reaching the following values: $\alpha = 0.7992, \beta = 0.3556$, and $\gamma = 0.9893$.

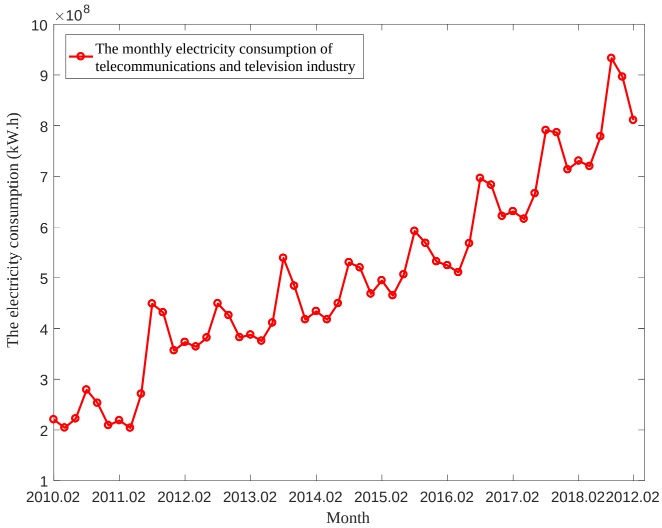


Figure 8: Monthly electricity consumption of telecommunications and television industry from city considered in this study.

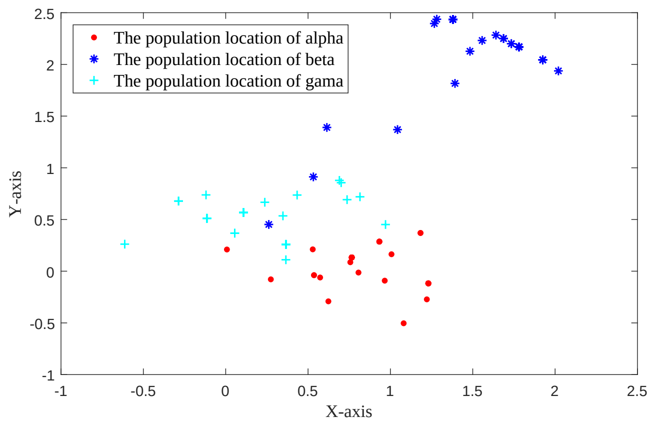


Figure 9: Search route of fruit fly swarms for α, β, and γ in telecommunications and television industry example.

From Figure 11 and Table 3, the FOA-MHW model obtains the best performance over the other benchmark models, as its MAPE is the lowest, at 1.89%. Hence, the proposed FOA-MHW model outperforms the other models in different scenarios.

The algorithm execution times of the evaluated models are listed in Table 4. The results are similar to those in section 3.2, with the execution time of the stand-alone models (i.e. SI and MHW-default) being far shorter than that of the hybrid models.

Table 1: Relative error and MAPE of evaluated models (%) during 2018 for city example

Month	SI model	MHW-default	FOASVR	GASVR	GA-MHW	FOA-MHW
Feb.	5.40	7.00	23.3	3.07	10.88	7.20
Apr.	2.16	1.84	1.28	2.63	1.01	3.29
July.	1.52	3.28	5.78	5.59	3.91	1.54
Aug.	4.82	6.82	3.20	11.58	5.21	1.66
Oct.	8.15	6.35	0.96	13.80	1.52	0.39
Dec.	4.26	1.81	2.71	5.35	6.45	7.84
MAPE	4.38	4.52	6.20	7.00	4.84	3.65

Table 2: Algorithm execution times of evaluated models in city example

Model	SI	MHW-default	FOASVR	GASVR	GA-MHW	FOA-MHW
Execution time (ms)	5.2	0.4	7325	9595	2589	1683

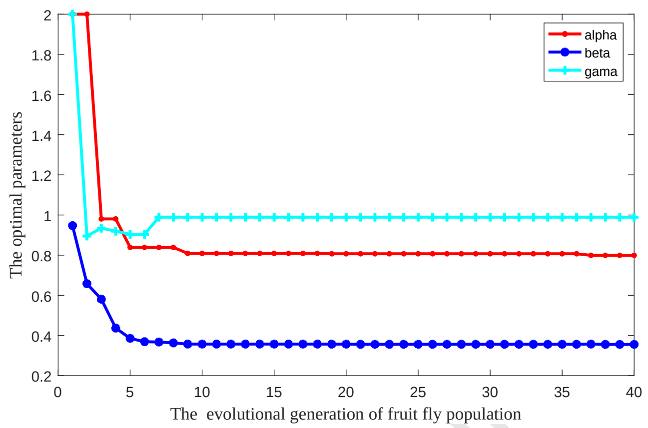


Figure 10: Update of optimal smoothing parameters in telecommunications and television industry example.

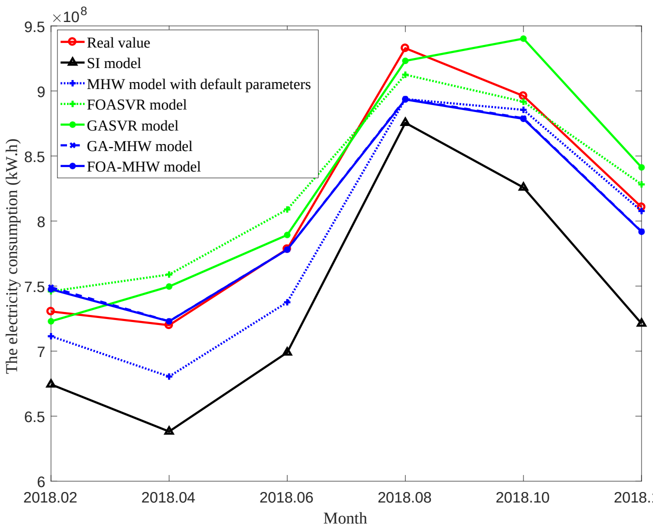


Figure 11: Forecasting results of evaluated models in telecommunications and television industry example.

Again, the proposed FOA-MHW model provides the shortest execution time among the hybrid models.

Finally, to evaluate the performance of the proposed FOA-MHW model on a small dataset, the monthly electricity consumption with varying training samples was forecasted. At least 3 years (i.e. 18 datapoints) and no more than 8 years (i.e. 48 datapoints) of training data were considered for this evaluation. The MAPE of the evaluated models is shown in Figure 12. With fewer training samples available, the MAPE of the FOASVR, GASVR, and MHW-default models increases. The MAPE of these models increases linearly as the availability of training samples decreases. In contrast, the MAPE of the proposed FOA-MHW and GA-MHW models slightly increases with fewer training samples, with the proposed FOA-MHW model having a small MAPE of 3.58% even with only 3 years of training data.

The MAPE of the SI model is the smallest with training data from less than 5 years, because this model can achieve high performance when data are regular and periodic, and the data from 2014 to 2018 are very periodic, as shown in Figure 8. In contrast, the data from 2010 to 2013 are not periodic enough, as shown in Figure 8, undermining the performance of the SI model. In addition, the MAPE of the FOA-MHW model is smaller than that of both the HW-default and GASVR models in any case. This is because the GASVR model based on machine learning requires massive training data to retrieve accurate results, but the FOA-MHW model has less restrictive data requirements. More specifically, as training data becomes scarce, the performance of models based on machine learning (i.e. FOASVR and GASVR) degrades, but that of the proposed FOA-MHW model remains mostly unchanged. Thus, the FOA-MHW model maintains high performance even when scarce training data are available, outperforming other methods based on machine learning.

4. Conclusion

This study was aimed to perform accurate monthly electricity consumption forecasting even under limited availability of

Table 3: Relative error and MAPE of evaluated models (%) during 2018 for telecommunications and television industry example

Month	SI model	MHW-default	FOASVR	GASVR	GA-MHW	FOA-MHW
Feb.	7.68	2.60	2.66	1.04	2.55	2.33
Apr.	11.35	5.48	7.08	4.13	0.42	0.41
July.	10.22	5.26	4.90	1.38	0.06	0.06
Aug.	6.15	4.19	2.57	1.04	4.17	4.23
Oct.	7.86	1.19	0.46	4.92	1.91	1.96
Dec.	11.04	0.37	2.25	3.76	2.31	2.34
MAPE	9.05	3.18	3.32	2.71	1.90	1.89

Table 4: Algorithm execution time of evaluated models in telecommunications and television industry example

Model	SI	MHW-default	FOASVR	GASVR	GA-MHW	FOA-MHW
Execution time (ms)	5.0	0.4	3629	5080	2930	1818

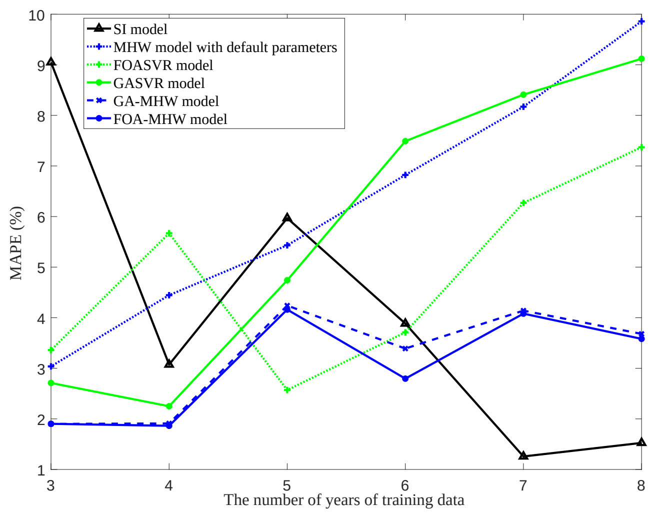


Figure 12: MAPE of evaluated models according to available training samples in telecommunications and television industry example.

historical data. To this end, we devised a hybrid forecasting model, called FOA-MHW, in which the FOA is adopted for parameter selection of the HW forecasting model. Based on actual electricity consumption data, the performance of the proposed model was verified under different scenarios. The proposed FOA-MHW model achieves the highest performance over other benchmark models, namely, SI, MHW-default, FOASVR, GASVR, and GA-MHW. Moreover, the execution time of the proposed FOA-MHW is shorter than that of other hybrid models. Specifically, the execution time of FOA-MHW is reduced to 23–50% that of FOASVR, 18–35% that of GASVR, and approximately 65% that of GA-MHW. Additional experiments confirmed that the proposed FOA-MHW model provides excellent forecasting performance even when a small number of training samples is available, with its MAPE being 3.58% with only 3 years of training data. This represents approximately twice the accuracy of the GASVR and FOASVR models, which are also based on machine learning. The proposed FOA-MHW model has been deployed in a Chinese energy supply company to forecast the monthly electricity consumption, and its high performance is being constantly verified in this practical scenario.

References

References

[1] K. B. Debnath, M. Mourshed, Forecasting methods in energy planning models, Renewable and Sustainable Energy Reviews 88 (2018) 297–325.

[2] H. Z. Li, S. Guo, C. J. Li, J. Q. Sun, A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm, Knowledge-Based Systems 37 (2) (2013) 378–387.

[3] Z. Shao, F. Chao, S. L. Yang, K. L. Zhou, A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting, Renewable and Sustainable Energy Reviews 75 (2017) 123–136.

[4] L. F. Tratar, E. Strmcnik, The comparison of holt-winters method and multiple regression method: A case study, Energy 109 (2016) 266–276.

[5] G. Aydin, Modeling of energy consumption based on economic and demographic factors: The case of turkey with projections, Renewable & Sustainable Energy Reviews 35 (35) (2014) 382–389.

[6] M. Imani, H. Ghassemian, Residential load forecasting using wavelet and collaborative representation transforms, Applied Energy 253 (2019) 113505.
[7] W. Jiang, H. Tang, L. Wu, H. Huang, H. Qi, Parallel processing of probabilistic models-based power supply unit mid-term load forecasting with apache spark, IEEE Access 7 (2019) 7588–7598.
[8] M. S. KIran, E. Özceylan, M. Gündüz, T. Paksoy, Swarm intelligence approaches to estimate electricity energy demand in turkey, Knowledge-Based Systems 36 (2012) 93–103.
[9] Y. Zhang, Q. Li, A regressive convolution neural network and support vector regression model for electricity consumption forecasting, arXiv preprint arXiv:1810.08878v1 (2018).
[10] P. Pełka, G. Dudek, Pattern-based forecasting monthly electricity demand using multilayer perceptron, in: Artificial Intelligence and Soft Computing, Springer International Publishing, Cham, 2019, pp. 663–672.
[11] M. Afrasiabi, M. Mohammadi, M. Rastegar, A. Kargarian, Multi-agent microgrid energy management based on deep learning forecaster, Energy 186 (2019) 115873.
[12] T. Windler, J. Busse, J. Rieck, One month-ahead electricity price forecasting in the context of production planning, Journal of Cleaner Production 238 (2019) 117910.
[13] M. Forouzanfar, A. Doustmohammadi, M. B. Menhaj, S. Hasanzadeh, Modeling and estimation of the natural gas consumption for residential and commercial sectors in iran, Applied Energy 87 (1) (2010) 268–274.
[14] P. Singh, P. Dwivedi, A novel hybrid model based on neural network and multi-objective optimization for effective load forecast, Energy 182 (2019) 606 – 622.
[15] Z. Tan, J. Zhang, J. Wang, J. Xu, Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models, Applied Energy 87 (11) (2010) 3606–3610.
[16] F. L. da Silva, F. L. C. Oliveira, R. C. Souza, A bottom-up bayesian extension for long term electricity consumption forecasting, Energy 167 (2019) 198–210.
[17] H. J. Sadaei, P. C. de Lima e Silva, F. G. Guimarães, M. H. Lee, Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series, Energy 175 (2019) 365 – 377.
[18] Z. Chang, Y. Zhang, W. Chen, Electricity price prediction based on hybrid model of adam optimized lstm neural network and wavelet transform, Energy 187 (2019) 115804.
[19] F. Y. Ju, W. C. Hong, Application of seasonal svr with chaotic gravitational search algorithm in electricity forecasting, Applied Mathematical Modelling 37 (23) (2013) 9643–9651.
[20] J. F. Chen, Q. Do, T. Nguyen, T. Doan, Forecasting monthly electricity demands by wavelet neuro-fuzzy system optimized by heuristic algorithms, Information 9 (3) (2018) 51.
[21] L. Hernandez, C. Baladron, J. M. Aguiar, B. Carro, J. Massana, A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings, IEEE Communications Surveys & Tutorials 16 (3) (2014) 1460–1495.
[22] J. F. Rendon-Sanchez, L. M. de Menezes, Structural combination of seasonal exponential smoothing forecasts applied to load forecasting, European Journal of Operational Research 275 (3) (2019) 916–924.
[23] G. J. Tsekouras, E. N. Dialynas, N. D. Hatziargyriou, S. Kavatza, A nonlinear multivariable regression model for midterm energy forecasting of power systems, Electric Power Systems Research 77 (12) (2007) 1560–1568.
[24] E. M. de Oliveira, F. L. Cyrino Oliveira, Forecasting mid-long term electric energy consumption through bagging arima and exponential smoothing methods, Energy 144 (2018) 776–788.
[25] S. Nawaz, N. Iqbal, S. Anwar, Modelling electricity demand using the star (smooth transition auto-regressive) model in pakistan, Energy 78 (2014) 535–542.
[26] M. H. Amjadi, H. Nezamabadi-Pour, M. M. Farsangi, Estimation of electricity demand of iran using two heuristic algorithms, Energy Conversion & Management 51 (3) (2010) 493–497.
[27] J. W. Taylor, R. Buizza, Neural network load forecasting with weather ensemble predictions, IEEE Trans on Power Systems 17 (7) (2002) 59–59.
[28] M. S. Kandil, S. M. El-Debeiky, N. E. Hasanien, Long-term load forecasting for fast developing utility using a knowledge-based expert system, IEEE Power Engineering Review 22 (4) (2002) 78–78.
[29] L. Tang, X. Wang, X. Wang, C. Shao, S. Liu, S. Tian, Long-term electricity consumption forecasting based on expert prediction and fuzzy bayesian theory, Energy 167 (2019) 1144 – 1154.
[30] J. Wang, X. Ma, W. Jie, D. Yao, Optimization models based on gm (1, 1) and seasonal fluctuation for electricity demand forecasting, International Journal of Electrical Power & Energy Systems 43 (1) (2012) 109–117.
[31] D. Ali, M. Yohanna, M. I. Puwu, B. M. Garkida, Long-term load forecast modelling using a fuzzy logic approach, Pacific Science Review A Natural Science & Engineering (2016) S2405882316300217.
[32] C. Hamzaçebi, H. A. Es, R. Çakmak, Forecasting of turkey's monthly electricity demand by seasonal artificial neural network, Neural Computing and Applications 31 (7) (2019) 2217–2231.
[33] T. Ahmad, H. Chen, Deep learning for multi-scale smart energy forecasting, Energy 175 (2019) 98 – 112.
[34] Y. He, Y. Qin, S. Wang, X. Wang, C. Wang, Electricity consumption probability density forecasting method based on lasso-quantile regression neural network, Applied Energy 233-234 (2019) 565 – 575.
[35] C. C. Holt, Forecasting seasonals and trends by exponentially weighted moving averages, International Journal of Forecasting 20 (1) (2004) 5–10.
[36] P. R. Winters, Forecasting sales by exponentially weighted moving averages, Management Science 6 (3) (1960) 324–342.
[37] W. T. Pan, A new fruit fly optimization algorithm: Taking the financial distress model as an example, Knowledge-Based Systems 26 (2) (2012) 69–74.
[38] S. Makridakis, S. C. Wheelwright, V. E. McGee, Forecasting: methods and applications., New York: Wiley, 1998.
[39] G. Cao, L. Wu, Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting, Energy 115 (2016) 734–745.
[40] J. Wang, X. Ma, J. Wu, Y. Dong, Optimization models based on gm (1,1) and seasonal fluctuation for electricity demand forecasting, International Journal of Electrical Power & Energy Systems 43 (1) (2012) 109 – 117.
[41] T. A. Oyehan, I. O. Alade, A. Bagudu, K. O. Sulaiman, S. O. Olatunji, T. A. Saleh, Predicting of the refractive index of haemoglobin using the hybrid ga-svr approach, Computers in Biology and Medicine 98 (2018) 85 – 92.

Highlights

- It is the first attempt to enhance the Holt--Winters exponential smoothing method.
- A high forecasting accuracy can be obtained by FOA-MHW method for the monthly electricity consumption.
- The proposed FOA-MHW method can obtain incredible performance even in the situation of insufficient training data.
- The proposed FOA-MHW model has the shortest execution time among with all the other hybrid models.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: