



A Novel Decomposition and Combination Technique for Forecasting Monthly Electricity Consumption

Xi Zhang¹ and Rui Li^{2*}

¹School of Economics, Beijing Technology and Business University, Beijing, China, ²School of Economics and Management, University of Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:
Jian Chai,
Xidian University, China

Reviewed by:
Hao Xiao,
Hunan University, China
Dongmei Guo,
Central University of Finance and
Economics, China

*Correspondence:
Rui Li
lirui21@ucas.ac.cn

Specialty section:
This article was submitted to
Smart Grids,
a section of the journal
Frontiers in Energy Research

Received: 10 October 2021
Accepted: 01 November 2021
Published: 01 December 2021

Citation:
Zhang X and Li R (2021) A Novel
Decomposition and Combination
Technique for Forecasting Monthly
Electricity Consumption.
Front. Energy Res. 9:792358.
doi: 10.3389/fenrg.2021.792358

With the share of electricity in total final energy consumption increasing quickly, the world is becoming increasingly dependent on electricity, which makes it more and more important to improve the forecasting accuracy of electricity consumption to ensure the normal operation of economic activities. In this paper, a novel decomposition and combination technique to forecast monthly electricity consumption is proposed. First, we use STL decomposition to obtain the trend, season, and residual components of the time series. Second, we use SARIMA, SVR, ANN, and LSTM to forecast trend, season, and residual component, respectively. Third, we use time correlation principle to improve the forecasting accuracy of season component. Fourth, we integrated the residual component predicted by SARIMA, SVR, ANN, and LSTM into a new sequence to improve the forecasting accuracy of residual component. In order to verify the performance of the proposed forecast model, monthly electricity consumption data in China is introduced as an example for empirical analysis. The results show that after STL decomposition, time correlation modification, and residual modification, the forecasting accuracy of each model has been gradually improved. We believe that the proposed forecast model in this paper can also be used to solve other mid- and long-term forecasting problems with obvious seasonal characteristics.

Keywords: monthly electricity consumption, STL, time correlation modification, ANN, LSTM

INTRODUCTION

Background

Resource depletion and global climate change are serious problems that human society is facing and will face for a long time. To escape from this dilemma, the global energy mix needs two transformations: clean energy substitution on the energy supply side and electric energy substitution on the energy consumption side. This paper focuses on electricity consumption. According to statistics, global electrification of the final consumption continues to follow an increasing trend, and the share of electricity in total final energy consumption is close to 20% in 2020.

As the world becomes more and more dependent on electricity, planning for electricity production is crucial. In addition, electricity is difficult to store, so it is usually used immediately after it is generated. This further increases the need for power companies to plan their electricity supply in a proactive manner. Reliable forecast of future electricity consumption

level is the primary guiding principle of planning. In particular, high forecasting accuracy of medium- and long-term electricity consumption is the key to power system scheduling and planning. In contrast, inaccurate forecast of electricity consumption can backfire. Overestimation will waste scarce energy resources, huge capital investment, and long construction time. Underestimation will lead to more serious negative consequences, such as power shortage. Clearly, if effective early warning is given in advance based on high forecasting accuracy of electricity consumption, some measures can be adopted to avoid negative consequences. However, electricity consumption is uncertain, complex, and nonlinear, which depends on political conditions, economy (Lin and Liu, 2016), human activities, population behavior (Hussain et al., 2016), climate factors (Hernández, 2013), and other external factors affecting the forecasting accuracy of electricity consumption.

Literature Review and Motivation

At present, many techniques are used to forecast electricity consumption, which can be roughly divided into three categories: nonlinear intelligent model, statistical analysis model, and gray forecasting model. Nonlinear models mainly include the artificial neural network (Kandananond, 2011; Kaytez et al., 2015; Liu et al., 2017; Ghadimi et al., 2018; Bedi and Toshniwal, 2019; Hamzaçebi et al., 2019), support vector machine (Pai and Hong, 2005; Kavaklioglu, 2011; Cao and Wu, 2016), and Markov chain (Zhao et al., 2014). In addition to the nonlinear intelligent models mentioned above, statistical analysis models, such as regression analysis method (Mohamed and Bodger, 2005; Wang et al., 2018) and autoregressive integrated moving average (Yuan et al., 2016), have also been widely used in electricity consumption forecasting. The gray forecasting model proposed by Deng enjoys high popularity in many forecasting applications because it can describe the characteristics of uncertain systems even in the face of a small amount of data. Therefore, some literature forecast electricity consumption based on the gray model (Akay and Atak, 2007; Bahrami et al., 2014; Zhao and Guo, 2016; Xu et al., 2017; Ding et al., 2018; Wu et al., 2018).

These methods can generally provide good forecasts. However, the statistical analysis models have the limitation of linear (or near linear) assumption, the gray forecasting models are usually only suitable for time series that approximate exponential growth, and the nonlinear intelligent models often suffer from overfitting or the difficulty of parameter selection. To remedy these shortcomings, some decomposition and combination techniques have been proposed in recent years and achieve better performance: the SARIMA model with residual modification (Wang et al., 2012), wavelet transform combined with machine learning and time series models (Nguyen and Nabney, 2010), weighted hybrid model where trend and seasonal components are predicted by combined method, and SARIMA, respectively (Zhu, 2011), bagging ARIMA and exponential smoothing methods (de Oliveira and Cyrino Oliveira, 2018), convolutional neural networks

and fuzzy time series (Sadaei et al., 2019), and structural combination of seasonal exponential smoothing forecasts (Rendon-Sanchez and de Menezes, 2019).

For the above existing researches, there are still some issues that need to be further studied. First, the statistical analysis models assume linearity and have good forecasting accuracy for periodic and regular sequences. The nonlinear intelligent model can forecast nonlinear and irregular time series better, but it has the problem of overfitting. How could the advantages of the two methods be combined to improve the forecasting accuracy? Second, except for the fluctuations of monthly electricity consumption affected by extreme weather changes, and sudden major economic and health events, the monthly electricity consumption also shows strong periodicity and regularity, so the comprehensive utilization of these two characteristics is meaningful to increase forecasting accuracy.

Contributions

To bridge the gap discussed above in the Literature review and motivation section, this paper develops a novel decomposition and combination forecasting technique. The primary research contents of this paper include three parts. First is the research on the monthly electricity consumption forecast based on STL decomposition. Second is the research on a time correlation modification based on annual periodicity and adjacent similarity to improve the forecasting accuracy of the season component. Third, considering the residual component has nonlinear and irregular characteristics, the individual model may only extract a certain feature of the sequence. Therefore, we integrate the residual component predicted by four models into a new sequence to improve the forecasting accuracy of the residual component. The main contributions of this paper are as follows:

- 1) A novel decomposition and combination forecasting model utilizing STL decomposition, time correlation principle (embodied as annual periodicity and adjacent similarity), and hybrid forecasting principle is proposed.
- 2) The monthly electricity consumption data of China are applied to evaluate the performance of the proposed model.

The remainder of the paper is organized as follows. The Electricity consumption month-ahead forecasting model section introduces the proposed forecasting model. The Case study section presents the simulation results and discussion, in which the performance of the proposed forecasting model is evaluated. Finally, conclusions are drawn in the Conclusion section.

ELECTRICITY CONSUMPTION MONTH-AHEAD FORECASTING MODEL

This section first briefly introduces individual models, including the STL algorithm, SARIMA, SVR, ANN, and LSTM model. Then

the operation process of the proposed decomposition and combination method is described.

Seasonal-Trend Decomposition Using Loess Decomposition

For seasonal time series, academics generally use STL decomposition proposed by Cleveland et al. (1990) to obtain trend, season, and residual components. STL is a decompose model in the form of addition. In STL, loess is used to divide the time series into trend component, seasonal component, and residual component. Division is addition, that is, adding up the parts to get the original series. Specifically, the steps of STL decomposition are 1) detrending; 2) periodic subsequence smoothing; establish a sequence for each seasonal component and smooth it separately; 3) smoothing periodic substring low-pass filtering; recombine substring to smooth; 4) detrending the seasonal series; 5) detrending the original series using the seasonal components calculated in the previous steps; and 6) smoothing the de-seasonal sequence to obtain the trend component.

Seasonal Autoregressive Integrated Moving Average

SARIMA is one of the most widely used linear models for time series prediction. The general equation of this model is given by Eq. 1.

$$\phi_p(B)(1-B)^d\Phi_p(B^s)(1-B^s)^Dy_t = \theta_q(B)\Theta_q(B^s)a_t. \quad (1)$$

Here y_t is time series, a_t is white noise, and B is the lag operator. D represents the seasonal differentiation order, and d represents the regular differentiation order.

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (2)$$

$$\theta_p(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_p B^p. \quad (3)$$

Eqs. 2, 3 represent the autoregressive and moving average polynomial, respectively. They represent the dependence of future values of time series on past values as well as errors.

$$\Phi_p(B) = 1 - \mu_1 B^s - \mu_2 B^{2s} - \dots - \mu_p B^{ps}. \quad (4)$$

$$\Theta_q(B) = 1 - \nu_1 B^s - \nu_2 B^{2s} - \dots - \nu_q B^{qs}. \quad (5)$$

Similarly, Eqs. 4, 5 represent the seasonal autoregressive and seasonal moving average polynomials, respectively. Addition of these polynomials to the ARIMA equation helps in capturing the seasonal variation in time series. Differentiation is necessary for converting the nonstationary time series to a stationary one. S represents the order of seasonality.

Support Vector Machine

SVM was first proposed by Vapnik (1963) based on the statistical learning theory and principle of structural risk minimization, which possess good performance even for small samples. The

basic idea of support vector regression is to map original data to high-dimensional feature space and perform linear regression in the space. It can be formulated into:

$$f(x) = w^T \varphi(x) + b, \quad (6)$$

where $\varphi(x)$ is a nonlinear mapping function, $f(x)$ is the estimation value, and w^T and b are weights. It can be translated into an optimization problem:

$$\begin{aligned} & \text{Min } \frac{1}{2} w^T w + C \sum_{t=1}^T (\xi_t + \xi_t^p), \\ & \text{s.t } \begin{cases} w^T \varphi(x_t) + b - y_t \leq \varepsilon + \xi_t, (t = 1, 2, \dots, T) \\ y_t - (w^T \varphi(x_t) + b) \leq \varepsilon + \xi_t^p, (t = 1, 2, \dots, T), \\ \xi_t, \xi_t^p \geq 0 \end{cases} \end{aligned} \quad (7)$$

where C is the penalty parameter, and ξ_t and ξ_t^p are the nonnegative slack variables. Generally speaking, the parameters of SVR have a great influence on the accuracy of the regression estimation. Thereby, the grid search method is employed to automatically choose the optimal parameters of SVR in this paper.

Artificial Neural Network

ANN is an information processing method based on the biological neural network. Neural networks can theoretically simulate any complex nonlinear relationship through nonlinear units (neurons) and have been widely used in the field of forecast. The structure of artificial neural network consists of input layer, hidden layer, and output layer. The most widely used ANN model is the BP neural network model based on the BP algorithm. The neural network is determined by determining the weight between each layer. Therefore, the neural network is trained to set all the weights before being used for prediction. The initial weights are set randomly, and the output data can be obtained according to certain rules when the training process is going forward. The weights are modified based on the difference between the output data and the expected data during the fallback process. The forward and backward process is repeated until the difference between the output data and the required data is small enough.

Long Short-Term Memory

Traditional artificial neural networks (ANN) attempt to establish direct mapping between input historical data and output forecast data to achieve prediction methods. However, due to the absence of time correlation in data series, the neural network model cannot capture the relationship between data and time, which limits its application in time series prediction methods. Therefore, recursive neural network (RNN) is proposed to overcome this shortcoming. By adding cyclic connections on neurons, RNN can establish sequence-to-sequence mappings between input and output data. Therefore, the output of each time step is affected by the input of the previous time step. Therefore, RNN is used to realize the memory feature (Sutskever et al., 2014; LeCun et al., 2015).

The structure of RNN is shown in Figure 1. Each node represents a single time-step neuron. The connection weight

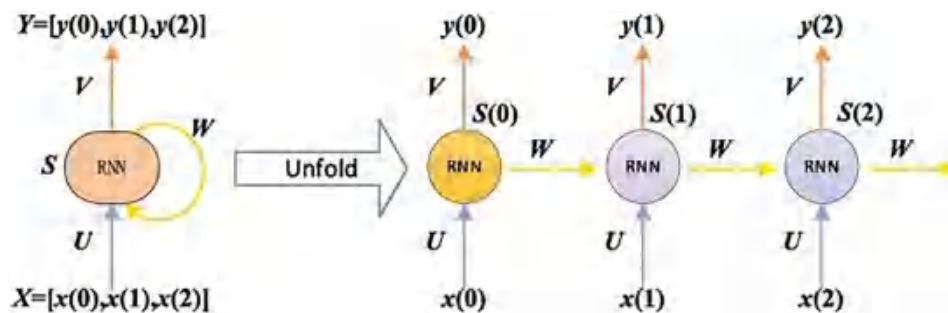


FIGURE 1 | The structure of recursive neural network (RNN) (Wang et al., 2020).

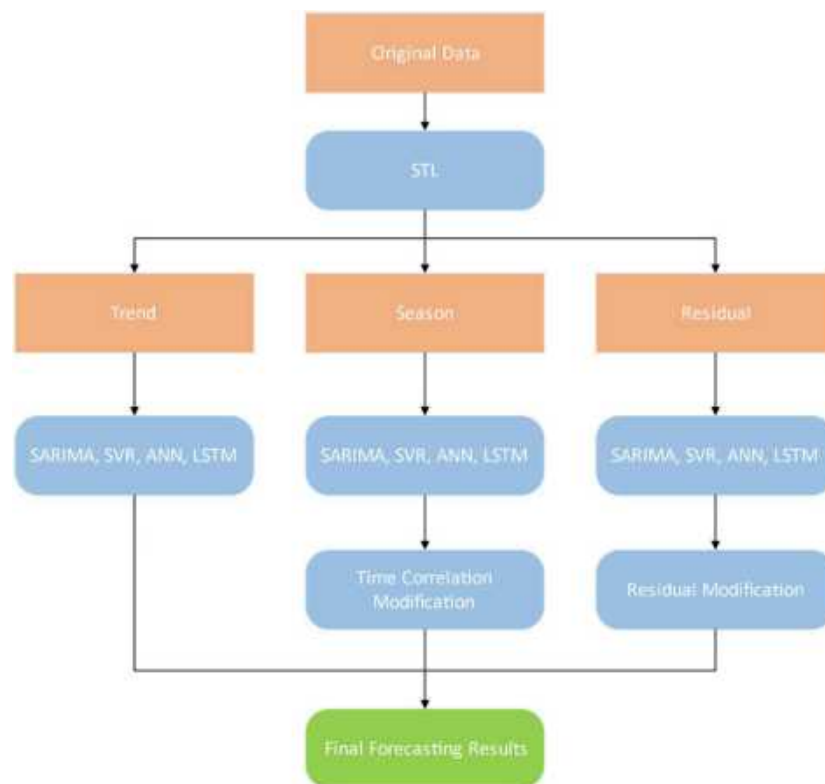


FIGURE 2 | The proposed forecast framework.

of input neuron is W_1 , the self-connection weight of each neuron is W_2 , and the connection weight of output neuron is W_3 . The input data sequence enters the network in turn according to the time step, and the weight coefficient is recycled.

The Proposed Forecast Framework

The proposed forecast framework utilizing STL decomposition, time correlation modification (embodied as annual periodicity and adjacent similarity) and residual modification is illustrated in Figure 2.

In Figure 2, The proposed forecast framework consists of four steps:

In the first step, we use the Seasonal–Trend decomposition using Loess (STL decomposition) to obtain the trend, season, and residual components of the time series.

In the second step, we use SARIMA, SVR, ANN, and LSTM to forecast trend, season, and residual component, respectively.

In the third step, we use the time correlation principle to improve the forecasting accuracy of the season component. The season component presents time correlation characteristics, which embodies as annual periodicity and adjacent similarity. Here the annual periodicity means that data from the same month in the next year are

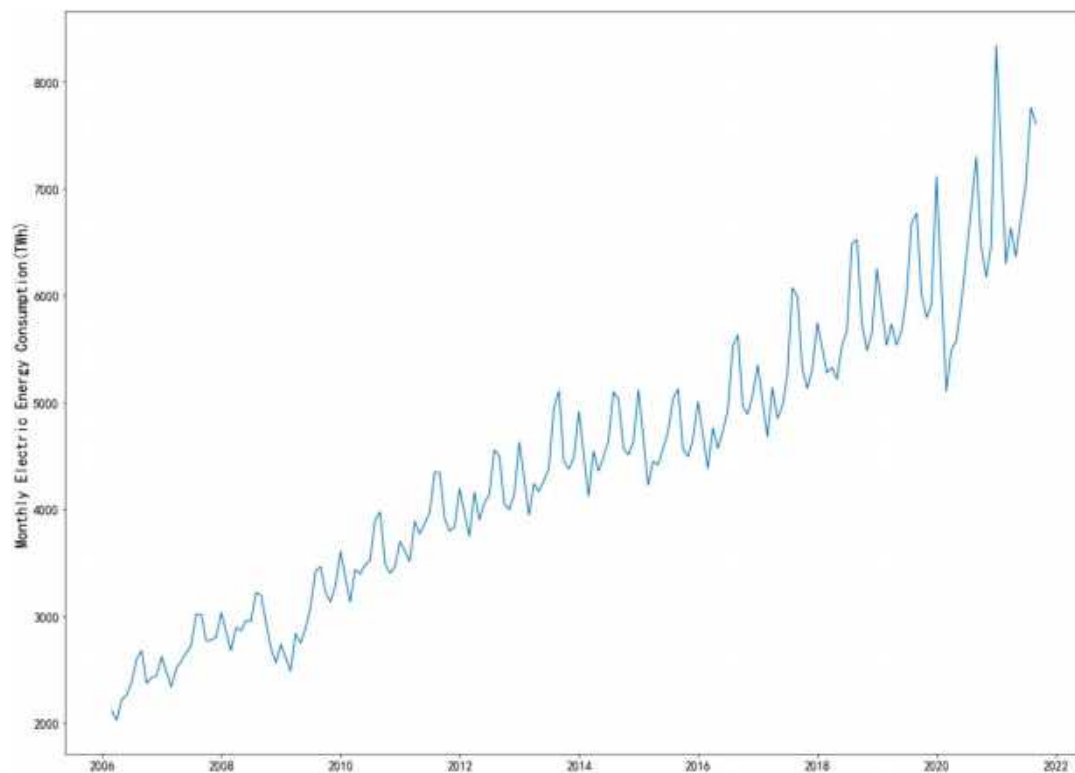


FIGURE 3 | The monthly electricity consumption of China from 1 and 2 2006 to August 2021.

similar. The adjacent similarity means that data are close to each other in adjacent months. In the second step, only adjacent similarity is used. We divide the season component into 11^2 subsequences, each of which represents a certain month. Then the exponential smoothing method is used to forecast each subsequence. The forecasting results are weighted with the season component predicted by each model (SARIMA, SVR, ANN, LSTM) to improve the forecasting accuracy of the season component. The weight is calculated based on the last forecasting error of the model.

In the fourth step, because the residual component has nonlinear and irregular characteristics, the individual model may only extract a certain feature of the sequence, so the forecasting accuracy is low. In fact, it is rare that a single forecasting model is always best in all cases. Each model has its own unique strengths and weaknesses. When multiple forecasting models are available, consider a combined approach, which is a good way to take full advantage of the strengths of each model. Therefore, we integrate the residual component predicted by SARIMA, SVR, ANN, and LSTM into a new sequence, and replace the residual component predicted by the above four methods with the new

sequence to improve the forecasting accuracy of the residual component.

CASE STUDY

Data Collection

We evaluate the performance of the proposed forecasting method using the monthly electricity consumption data of China. However, these figures cannot be used directly as Chinese New Year always lasts for a few days in January or February. Almost all companies and factories have stopped operating. As a result, electricity consumption in January and February is sometimes abnormal. To avoid this problem, we treat the January and February averages as observations of a new month 1 and 2 each year, i.e., each year has 11 monthly values with a period length of 11. This study collects electricity consumption data from the beginning of 1 and 2 2006 to the end of August 2021 to keep relevant to the current situation of electricity development. These original data are shown in Figure 3.

Experimental Design

We select the data from 1 and 2 2006 to December 2018 as the training dataset (i.e., the first 143 data points) and the remaining data as the test dataset (i.e., the last 29 data points). The training data set is further divided into the optimization training data set and the verification data

²In the later empirical research, we treat the January and February averages as observations of a new month 1 and 2 each year.

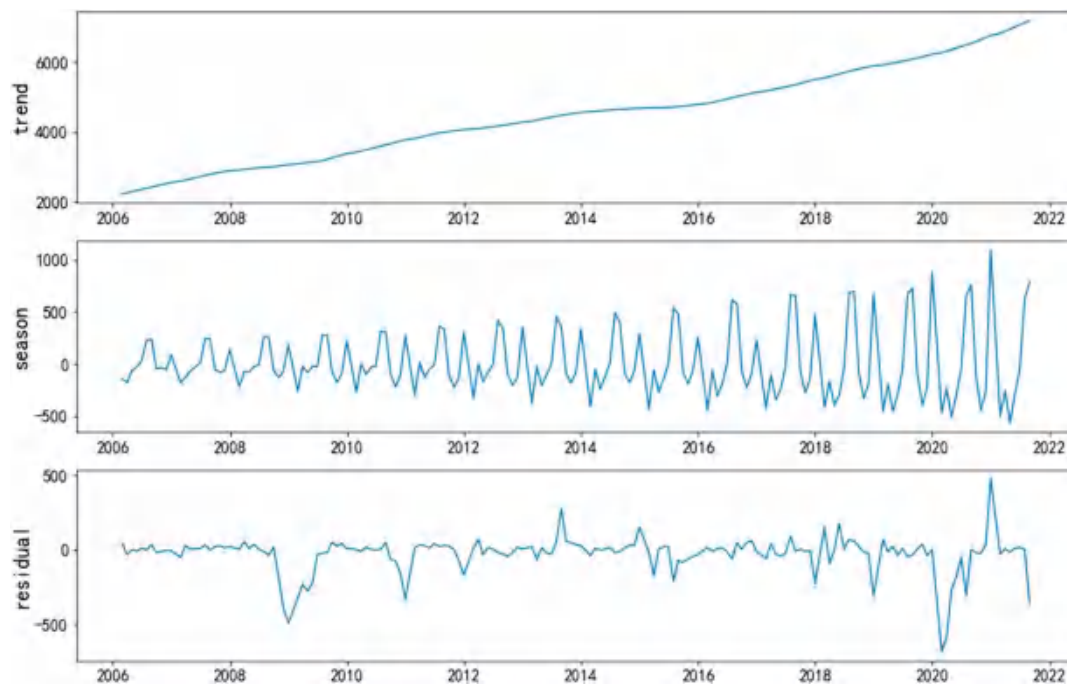


FIGURE 4 | The trend, seasonal, and residual for monthly electricity consumption data decomposed by Seasonal-Trend decomposition using Loess (STL).

TABLE 1 | Performance evaluations of different models with or without Seasonal-Trend decomposition using Loess (STL).

Horizons		One-step ahead			Two-step ahead			Three-step ahead		
Indices		Mean absolute error (MAE) (TWh)	Mean absolute percentage error (MAPE) (%)	Correlation coefficient(COR)	MAE (TWh)	MAPE (%)	COR	MAE (TWh)	MAPE (%)	COR
Without STL	SARIMA	230	3.58	0.88	259	4.06	0.86	280	4.40	0.85
	SVR	346	5.59	0.79	346	5.59	0.79	383	6.22	0.76
	Artificial neural network (ANN)	266	4.22	0.87	266	4.22	0.87	317	5.07	0.86
	LSTM	336	5.36	0.78	336	5.36	0.78	406	6.35	0.71
With STL	STL-SARIMA	126	1.98	0.96	121	1.93	0.96	131	2.10	0.96
	STL-SVR	145	2.23	0.95	145	2.23	0.95	197	3.04	0.94
	STL-ANN	150	2.34	0.95	150	2.34	0.95	189	2.99	0.94
	STL-LSTM	169	2.78	0.96	173	2.83	0.96	211	3.41	0.93

set. The optimization training data set contains the first 121 data points, and the verification training data set contains the last 22 data points. Optimization training and validation data sets were used to determine the hyperparameters of SVR, ANN, and LSTM models, while test data sets are used to evaluate forecasting performance.

Three error indices, mean absolute error (MAE), mean absolute percentage error (MAPE), and correlation coefficient (COR), are applied to evaluate the model performance according to forecast results. The official functions of the three error indices are:

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n},$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i},$$

$$COR = \frac{Cov(y_i, \hat{y}_i)}{\sigma_y \sigma_{\hat{y}}},$$

where y is the actual value, $\hat{y}$ is the forecasted value, and i is the index value of the data.



FIGURE 5 | Trend, season, random error ratio.

TABLE 2 | Performance evaluations of different models with time correlation modification and residual modification.

Horizons		One-step ahead			Two-step ahead			Three-step ahead		
Indices		MAE (TWh)	MAPE (%)	COR	MAE (TWh)	MAPE (%)	COR	MAE (TWh)	MAPE (%)	COR
TCM	STL-SARIMA-TCM	118	1.86	0.97	115	1.83	0.96	127	2.04	0.96
	STL-SVR-TCM	131	2.05	0.97	131	2.05	0.97	182	2.84	0.96
	STL-ANN-TCM	134	2.11	0.97	136	2.14	0.96	168	2.68	0.96
	STL-LSTM-TCM	158	2.61	0.97	165	2.71	0.97	167	2.70	0.95
RM	STL-SARIMA-TCM-RM	113	1.82	0.97	113	1.80	0.97	123	1.98	0.96
	STL-SVR-TCM-RM	135	2.14	0.97	136	2.15	0.97	187	2.92	0.96
	STL-ANN-TCM-RM	129	2.06	0.97	132	2.10	0.96	156	2.48	0.96
	STL-LSTM-TCM-RM	129	2.10	0.97	131	2.11	0.97	141	2.27	0.96

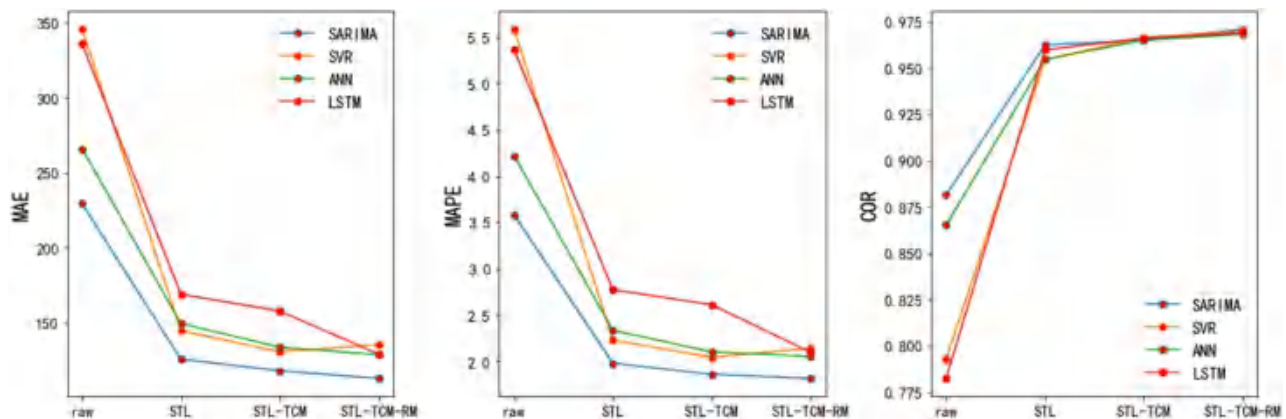


FIGURE 6 | Comparison between different models for 1-month ahead forecasting.

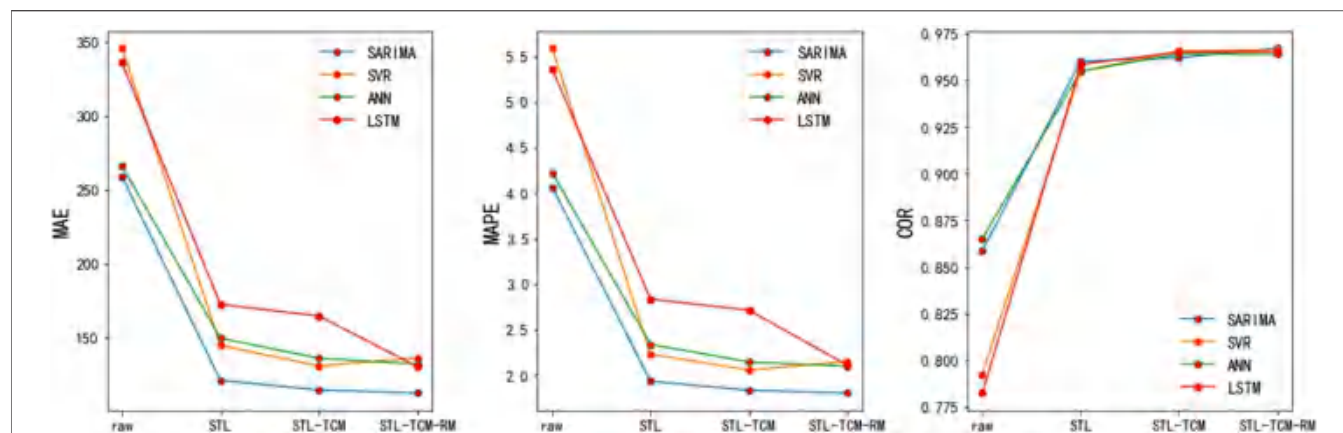


FIGURE 7 | Comparison between different models for 2-month ahead forecasting.

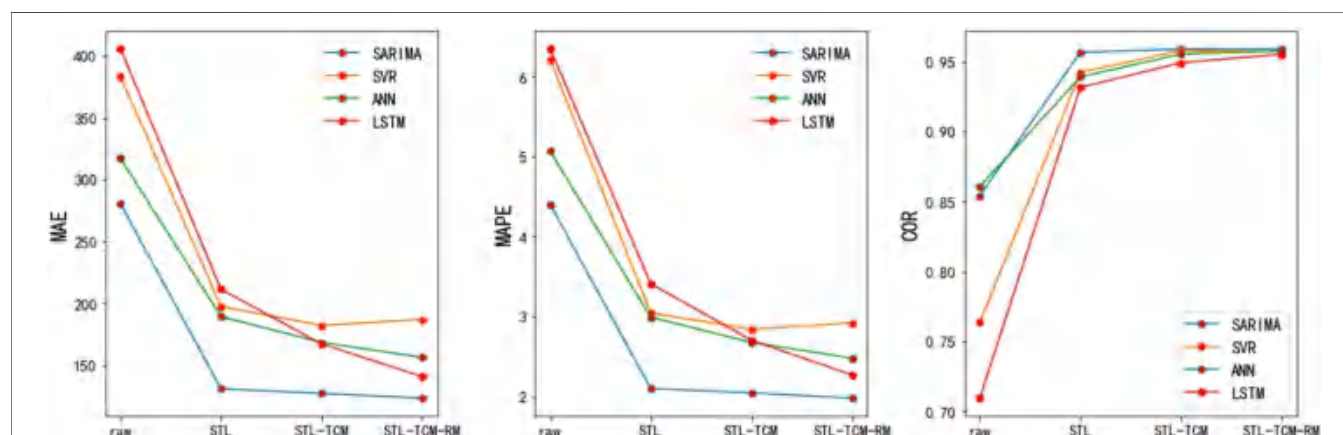


FIGURE 8 | Comparison between different models for 3-month ahead forecasting.

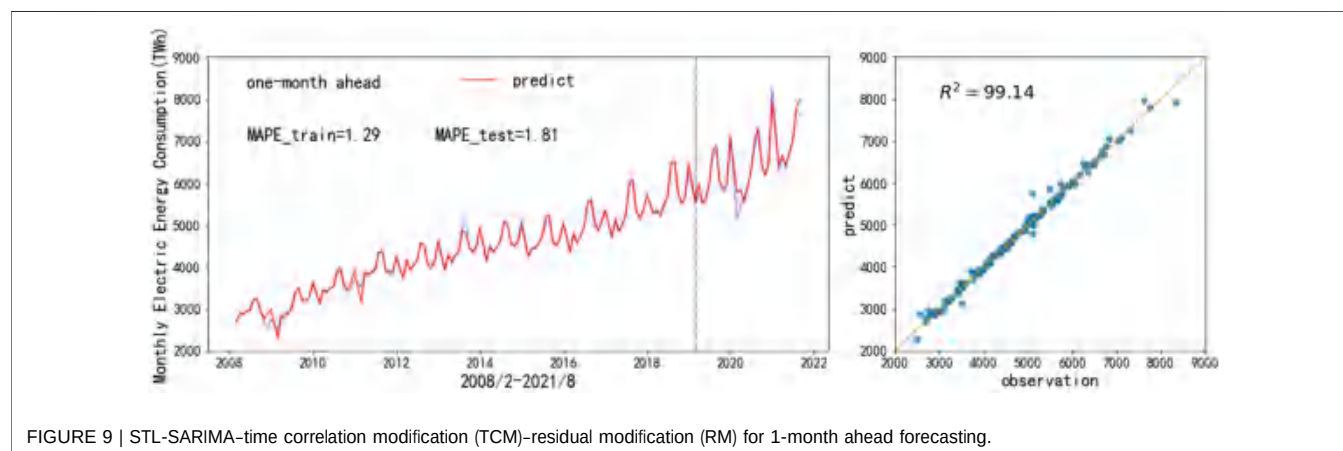


FIGURE 9 | STL-SARIMA-time correlation modification (TCM)-residual modification (RM) for 1-month ahead forecasting.

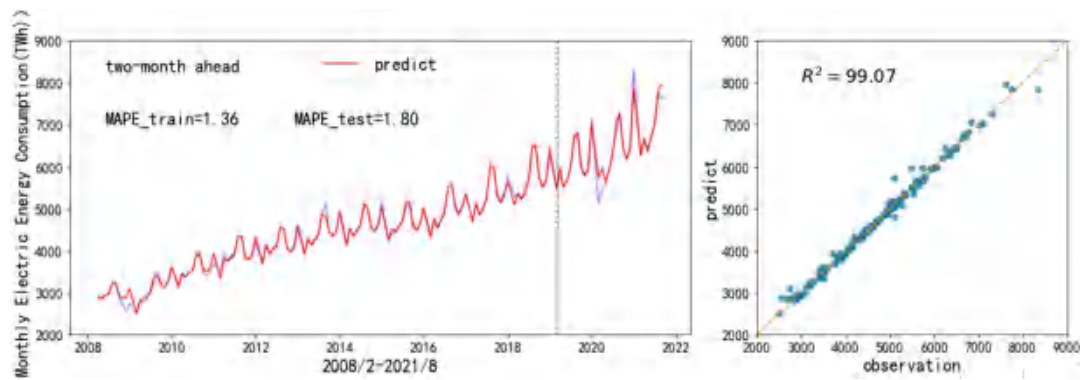


FIGURE 10 | STL-SARIMA-TCM-RM for 2-month ahead forecasting.

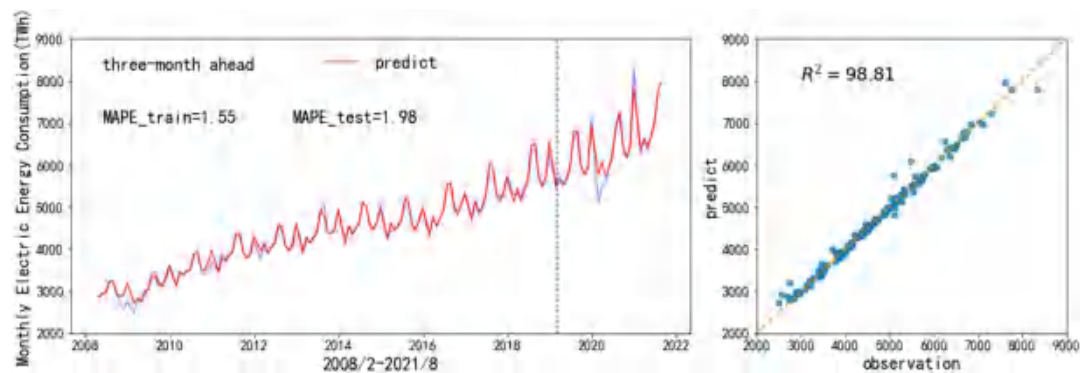


FIGURE 11 | STL-SARIMA-TCM-RM for 3-month ahead forecasting.

TABLE 3 | Performance evaluation of STL-SARIMA-TCM-RM in 2019, 2020, and 2021.

Horizons Indices	One-step ahead			Two-step ahead			Three-step ahead		
	MAE (TWh)	MAPE (%)	COR	MAE (TWh)	MAPE (%)	COR	MAE (TWh)	MAPE (%)	COR
2019	70	1.16	0.99	52	0.86	0.99	48	0.76	0.99
2020	166	2.77	0.97	193	3.19	0.97	220	3.66	0.96
2021	97	1.36	0.98	82	1.11	0.99	89	1.24	0.98

Results

Seasonal-Trend decomposition using Loess Decomposition

Figure 4 shows the STL decomposition results of the monthly electricity consumption. The trend component of the electricity consumption of China is increasing year by year, and the growth trend has accelerated since 2016. This is mainly because in 2016, eight departments in China jointly issued The Guidelines on Promoting the Substitution of Electric Energy, with a view to increasing the proportion of electric energy in the final energy consumption to 27%. Electric energy substitution is an important way to achieve carbon peak and carbon neutrality by replacing

coal, oil, gas, and wood with electricity in energy consumption. The season component vibrates more and more. Due to financial crisis, extreme weather events, and epidemic, there are several relatively large negative and positive shocks on the residual component. If the original sequence is directly used, these huge shocks will seriously threaten the forecasting accuracy of the model.

Table 1 shows the performance evaluation results of four models without STL decomposition and with STL decomposition. The model comparisons demonstrate that STL decomposition is effective in boosting the forecasting accuracy of monthly electricity consumption. Compared with any single

model (SARIMA, SVR, ANN, LSTM), the models with STL decomposition leads to reductions in all of the evaluation indices (MAE, MAPE and COR).

According to Table 1, the divide-and-conquer strategy improves the forecasting accuracy. Next, we analyze the source of errors, that is, the percentage of trend, season, and residual component forecasting errors to the total errors. As shown in Figure 5, for any model, most of the errors come from residual component forecast. SARIMA, in particular, was the least effective. This is because the residual component has nonlinear and irregular characteristics, and SARIMA is not good at forecasting these kinds of sequences. In addition, a single model may only extract a certain feature of the sequence, so the forecasting accuracy is low. For trend component, it can be seen that SARIMA has the highest forecasting accuracy, while for machine learning algorithms, such as SVR, ANN, and LSTM, the accuracy is not high. Therefore, it can be concluded that the traditional statistical method is better for simple sequence like trend component. The forecasting errors of season component also account for a large part.

Time Correlation Modification

As shown in Figure 5, the errors caused by season component account for 18%–28%. In this section, we use the periodicity of the seasonal series to improve the forecasting accuracy of the season component. As we can see, the season component presents time correlation characteristics, which embodies as annual periodicity and adjacent similarity. We divide the season component into 11 subsequences, each of which represents a certain month. Then exponential smoothing method is used to forecast each subsequence. The forecasting results are weighted with the season component predicted by each model (SARIMA, SVR, ANN, and LSTM) to improve the forecasting accuracy of the season component. The weight is calculated based on the last forecast error of the model. Rows 3–6 in Table 2 show that the forecasting accuracy has been improved after time correlation modification (TCM).

Residual Modification

Figure 5 shows that most of the errors come from a residual component. This is because the residual component has nonlinear and irregular characteristics; a single model may only extract a certain feature of the sequence, so the forecasting accuracy of a single model is low. Therefore, we need to improve the forecasting accuracy of the residual component. We integrate the residual component predicted by SARIMA, SVR, ANN, and LSTM into a new sequence, and replace the residual component predicted by the above four methods with the new sequence to optimize each model. Rows 7–10 in Table 2 show that the forecasting accuracy has been improved after residual modification (RM).

Comparison Between Different Models

Figures 6–8 show that after STL decomposition, time correlation modification, and residual modification, the forecasting accuracy

of each model has been gradually improved. Among them, the forecasting accuracy improved the most after STL decomposition. This is mainly because there are many random disturbances in the original sequence, and the model will be affected by these disturbances if it is not decomposed.

DISCUSSION

According to Figures 6–8, STL–SARIMA–TCM–RM is the most accurate forecasting model.³. As we can see, on the one hand, compared with machine learning, SARIMA is better at forecasting trend, season, and other sequences with clear patterns. That is why it is so accurate. On the other hand, SARIMA is not good at forecasting an irregular random term. Therefore, residual modification can improve the forecasting accuracy of SARIMA most significantly. Figures 9–11 show the STL–SARIMA–TCM–RM forecasting performance in the data set, as well as a scatter plot of forecasting results and actual values.

Considering that the test set includes COVID-19, we divide the test set into 2019, 2020, and 2021. Table 3 shows that the 2019 forecast results are significantly better than that for 2020 and 2021.

CONCLUSION

This paper provides a novel decomposition and combination method to forecast electricity consumption. This approach first uses STL to decompose the sequence into trend, season, and residual components. Then the three decomposed subsequences are forecasted, and the season component forecasting results are modified according to the annual periodicity, and the forecasting results of the residual component of each model are integrated. The results show that STL–SARIMA–TCM–RM is the most accurate forecasting model.

In addition to electricity forecasting, we believe that the forecasting method proposed in this paper can also be used to solve other mid- and long-term forecasting problems with obvious seasonal characteristics, including tourist flow forecasting, energy consumption forecasting, traffic flow forecasting, and so on. Furthermore, this paper only focuses on univariate time series analysis and does not consider other factors affecting electricity consumption. If these factors can be introduced into the proposed learning method, the predictive performance may be better.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

³The STL–SARIMA–TCM–RM model is also superior to the existing models (Zhao et al., 2014; Cao and Wu, 2016; Jiang et al., 2020; Liu et al., 2020). Their models' MAPE all above 2, and the STL–SARIMA–TCM–RM model is less than 2.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

- Akay, D., and Atak, M. (2007). Grey Prediction with Rolling Mechanism for Electricity Demand Forecasting of Turkey. *Energy* 32 (9), 1670–1675. doi:10.1016/j.energy.2006.11.014
- Bahrami, S., Hooshmand, R.-A., and Parastegari, M. (2014). Short Term Electric Load Forecasting by Wavelet Transform and Grey Model Improved by PSO (Particle Swarm Optimization) Algorithm. *Energy* 72, 434–442. doi:10.1016/j.energy.2014.05.065
- Bedi, J., and Toshniwal, D. (2019). Deep Learning Framework to Forecast Electricity Demand. *Appl. Energy* 238, 1312–1326. doi:10.1016/j.apenergy.2019.01.113
- Cao, G., and Wu, L. (2016). Support Vector Regression with Fruit Fly Optimization Algorithm for Seasonal Electricity Consumption Forecasting. *Energy* 115, 734–745. doi:10.1016/j.energy.2016.09.065
- Cleveland, R. B., Cleveland, W., McRae, J. E., and Terpenning, I. J. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess (With Discussion). *J. Off. Stat* 6, 3–73.
- de Oliveira, E. M., and Cyrino Oliveira, F. L. (2018). Forecasting Mid-long Term Electric Energy Consumption through Bagging ARIMA and Exponential Smoothing Methods. *Energy* 144, 776–788. doi:10.1016/j.energy.2017.12.049
- Ding, S., Hipel, K. W., and DangDang, Y.-g. (2018). Forecasting China's Electricity Consumption Using a New Grey Prediction Model. *Energy* 149, 314–328. doi:10.1016/j.energy.2018.01.169
- Ghadimi, N., Akbarimajid, A., Shayeghi, H., and Abedinia, O. (2018). Two Stage Forecast Engine with Feature Selection Technique and Improved Meta-Heuristic Algorithm for Electricity Load Forecasting. *Energy* 161, 130–142. doi:10.1016/j.energy.2018.07.088
- Hamzaçebi, C., Hüseyin Avni, Es., Es, H. A., and Çakmak, R. (2019). Forecasting of Turkey's Monthly Electricity Demand by Seasonal Artificial Neural Network. *Neural Comput. Applic* 31 (7), 2217–2231. doi:10.1007/s00521-017-3183-5
- Hernández, D. (2013). Energy Insecurity: A Framework for Understanding Energy, the Built Environment, and Health Among Vulnerable Populations in the Context of Climate Change. *Am. J. Public Health* 103 (4), e32–e34. doi:10.2105/ajph.2012.301179
- Hussain, A., Rahman, M., and Memon, J. A. (2016). Forecasting Electricity Consumption in Pakistan: The Way Forward. *Energy Policy* 90, 73–80. doi:10.1016/j.enpol.2015.11.028
- Jiang, W., Wu, X., Gong, Y., Yu, W., and Zhong, X. (2020). Holt-Winters Smoothing Enhanced by Fruit Fly Optimization Algorithm to Forecast Monthly Electricity Consumption. *Energy* 193, 116779. doi:10.1016/j.energy.2019.116779
- Kandananond, K. (2011). Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. *Energies* 4 (8), 1246–1257. doi:10.3390/en4081246
- Kavaklioglu, K. (2011). Modeling and Prediction of Turkey's Electricity Consumption Using Support Vector Regression. *Appl. Energy* 88 (1), 368–375. doi:10.1016/j.apenergy.2010.07.021
- Kaytez, F., Taplamacioglu, M. C., Cam, E., and Hardalac, F. (2015). Forecasting Electricity Consumption: A Comparison of Regression Analysis, Neural Networks and Least Squares Support Vector Machines. *Int. J. Electr. Power Energy Syst.* 67, 431–438. doi:10.1016/j.jepes.2014.12.036
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature* 521 (7553), 436–444. doi:10.1038/nature14539
- Lin, B., and Liu, C. (2016). Why Is Electricity Consumption Inconsistent with Economic Growth in China. *Energy Policy* 88, 310–316. doi:10.1016/j.enpol.2015.10.031
- Liu, Y., Wang, W., and Ghadimi, N. (2017). Electricity Load Forecasting by an Improved Forecast Engine for Building Level Consumers. *Energy* 139, 18–30. doi:10.1016/j.energy.2017.07.150
- Liu, Y., Zhao, J., Liu, J., Chen, Y., and Ouyang, H. (2020). Regional Midterm Electricity Demand Forecasting Based on Economic, Weather, Holiday, and Events Factors. *IEEE Trans. Elec. Electron. Eng.* 15, 225–234. doi:10.1002/tee.23049
- Mohamed, Z., and Bodger, P. (2005). Forecasting Electricity Consumption in New Zealand Using Economic and Demographic Variables. *Energy* 30 (10), 1833–1843. doi:10.1016/j.energy.2004.08.012

FUNDING

The research was supported by National Natural Science Foundation of China (72003009).

- Nguyen, H. T., and Nabney, I. T. (2010). Short-Term Electricity Demand and Gas Price Forecasts Using Wavelet Transforms and Adaptive Models. *Energy* 35 (9), 3674–3685. doi:10.1016/j.energy.2010.05.013
- Pai, P.-F., and Hong, W. C. H. (2005). Support Vector Machines with Simulated Annealing Algorithms in Electricity Load Forecasting. *Eng. Convers. Manag.* 46 (17), 2669–2688. doi:10.1016/j.enconman.2005.02.004
- Rendon-Sanchez, J. F., and de Menezes, L. M. (2019). Structural Combination of Seasonal Exponential Smoothing Forecasts Applied to Load Forecasting. *Eur. J. Oper. Res.* 275 (3), 916–924. doi:10.1016/j.ejor.2018.12.013
- Sadaei, H. J., de Lima e SilvaSilva, P. C., Guimarães, F. G., and Lee, M. H. (2019). Short-Term Load Forecasting by Using a Combined Method of Convolutional Neural Networks and Fuzzy Time Series. *Energy* 175, 365–377. doi:10.1016/j.energy.2019.03.081
- Sutskever, I., Vinyals, O., and QuocLe, v. (2014). Sequence to Sequence Learning with Neural Networks. *Adv. Neural Inf. Process. Syst.* 4 (January), 3104–3112.
- Zhu, L., Huang, X., Shi, H., Cai, X., and Song, Y. (2011). Transport Pathways and Potential Sources of PM10 in Beijing. *Atmos. Environ.* 45 (3), 594–604. doi:10.1016/j.atmosenv.2010.10.040
- Vapnik, V. (1963). Pattern Recognition Using Generalized Portrait Method. *Autom. Remote Control* 24, 774–780.
- Wang, F., Xuan, Z., Zhen, Z., Li, K., Wang, T., and Shi, M. (2020). A Day-Ahead PV Power Forecasting Method Based on LSTM-RNN Model and Time Correlation Modification under Partial Daily Pattern Prediction Framework. *Eng. Convers. Manag.* 212, 112766. doi:10.1016/j.enconman.2020.112766
- Wang, Y., Wang, J., Zhao, G., and Dong, Y. (2012). Application of Residual Modification Approach in Seasonal ARIMA for Electricity Demand Forecasting: A Case Study of China. *Energy Policy* 48, 284–294. doi:10.1016/j.enpol.2012.05.026
- Wang, Z.-X., Li, Q., Pei, L.-L., and Pei, L. L. (2018). A Seasonal GM(1,1) Model for Forecasting the Electricity Consumption of the Primary Economic Sectors. *Energy* 154, 522–534. doi:10.1016/j.energy.2018.04.155
- Wu, L., Gao, X., Xiao, Y., Yang, Y., and Chen, X. (2018). Using a Novel Multi-Variable Grey Model to Forecast the Electricity Consumption of Shandong Province in China. *Energy* 157, 327–335. doi:10.1016/j.energy.2018.05.147
- Xu, N., Dang, Y., and Gong, Y. (2017). Novel Grey Prediction Model with Nonlinear Optimized Time Response Method for Forecasting of Electricity Consumption in China. *Energy* 118, 473–480. doi:10.1016/j.energy.2016.10.003
- Yuan, C., Liu, S., and Fang, Z. (2016). Comparison of China's Primary Energy Consumption Forecasting by Using ARIMA (The Autoregressive Integrated Moving Average) Model and GM(1,1) Model. *Energy* 100, 384–390. doi:10.1016/j.energy.2016.02.001
- Zhao, H., and Guo, S. (2016). An Optimized Grey Model for Annual Power Load Forecasting. *Energy* 107, 272–286. doi:10.1016/j.energy.2016.04.009
- Zhao, W., Wang, J., and Lu, H. (2014). Combining Forecasts of Electricity Consumption in China with Time-Varying Weights Updated by a High-Order Markov Chain Model. *Omega* 45, 80–91. doi:10.1016/j.omega.2014.01.002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.