Short-term electric power demand forecasting using a hybrid model of SARIMA and SVR

Qihe Lou\textsuperscript{1,2}, Qi Lyu\textsuperscript{3,}* , Zhixiong Na\textsuperscript{4}, Dayan Ma\textsuperscript{4}, Xiaoguang Ma\textsuperscript{4}

\textsuperscript{1}North China Electric Power University, Beijing102206, China
\textsuperscript{2}State Grid Corporation of China, Beijing100031, China
\textsuperscript{3}Tsinghua University, Beijing 100084, China
\textsuperscript{4}State Grid e-commerce Co., Ltd, Beijing 10053, China

*lvq19@mails.tsinghua.edu.cn

Abstract. Short-term electric power demand forecasting is the most basic and important application of smart grid. With the rapid development of renewable energy and clean energy in recent years, power demand forecasting gains special attention again. It has a great impact on the planning of power generation units and the purchase and sale of power market. Furthermore, it is also conducive to the realization of demand response and resource allocation efficiency and reliability, thus contributing to the Photovoltaic power generation system as well. Although generous methods have been proposed, it remains to be an open challenge owing to its limited precision. In this paper, a hybrid method of Seasonal Auto-Regressive Integrated Moving Average (SARIMA) and Support Vector Regression (SVR) is proposed for hourly forecasting, which overcomes the difficulty of generalization of a single model. The core thought of this model is using SARIMA to fit the linear part of the series and correcting the deviation by SVR. We tested the accuracy of this method on a public data set, and the result shows that it performs much better than the single SARIMA model on the fitting effect. Also, we compared our model with the recent three works, and it demonstrates a decrease of 18.102%, 10.534% and 4.757% in forecasting error respectively. To our surprise, when we used the previous four hours of data to predict the current data by the single SVR, we got the best performance of all models. In the end, we analyzed the possible causes for this result.

1. Introduction

The electric power industry has become the most important factor affecting the development of society and economics in the world, and has attracted more and more countries’ attention. The electric power system is based on three levels: generation, transport, and distribution and marketing, among which generation is an extremely important part. Inaccurate power demand forecast will increase the operating cost of utility companies, especially in the market environment, because if the forecast is underestimate, it will lead to adverse events such as black-outs and load shedding, and affect corporate profits, while when the forecast is overestimate, it will lead to superfluous power generation, and for power cannot be stored in larger quantities [1], the excess power that could have been converted into benefit is wasted. The imbalance between supply and demand caused by inaccurate forecast in traditional power generation system has led to the integration of advanced communication technologies into traditional power grid, which are known as smart grids [2]. Meanwhile, with the recent
emergence of new players in the field of power system (Electric Vehicles (EV), Smart Customers, Renewable Energy). Power demand forecasting gets special attention again [3]. For example, photovoltaic (PV) system, as a new promising renewable energy system, provides increasingly larger shares of power generation and becomes the fastest growing generation technology today. PV power generation lays a great dependency on the weather factor, thus has higher requirements for the coordination of load and storage of the grid. So accurate power demand forecasting is of great importance and can provide powerful data support for the consumption balance and energy storage management of the photovoltaic power station. It is obvious that even a small improvement in forecasting accuracy means a huge cost saving and a great contribution to the environment.

Hippertet al. [4] indicates that electric power demand forecasting can be classified into four groups based on the period of time to be predicted: very short-term load forecast (VSTLF) [5], short-term load forecasting (STLF) [6-8], medium-term load forecast (MTLF) [9], and long-term load forecast (LTLF) [10]. According to [3], the most important forecasting horizons are weekly, daily, and hourly. Producing an accurate forecast of the next 24 hours is essential for power companies, since it can have a direct impact on the optimal hourly planning of the generation units, as well as on purchase/sale in exchange systems, etc. This is called “load profile”. Although electricity demand forecasting has been widely studied for many years, it is still a critical task. One of the most important reasons for difficulty of accurate forecasting is that even though the power demand seems like a univariate time series [11], it is subject to various influential factors which may have discriminative capability in influencing the power demand. Thus, to forecast precisely with a single simple model is a hard thing.

Electrical load is a typical time series, since it consists of successive hourly measurements, it may possess seasonal patterns such as daily, weekly, monthly, etc. In order to model such time series, SARIMA models can be used. Furthermore, SVR could realize nonlinearity and deals with arbitrarily structured data by using a kernel function. In this paper, we propose a hybrid model of SARIMA and SVR to forecast hourly demand. SARIMA is used to fit the linear part of the sequence, and SVR is used to make up the residual caused by a multiple of complex factors. In order to verify the effectiveness of our model mixing, we compared the results of the hybrid model with the single model; and for verifying the superiority of our model, we tested the actual effect of the model and the recent three works [12-14] on the same real open data set. The results showed that the hybrid model is much more effective than the single model, and it demonstrates a decrease of 18.102%, 10.534% and 4.757% in forecasting error with comparison to the recent three works.

The rest of the paper is organized as follows: in Section II, we introduce some related researches on electric power demand forecasting using (S)ARIMA, SVR and hybrid models. Section III provides some basic theoretical aspects of SARIMA and SVR. The experimental processes and results are described in Section IV. Section V discusses the results and future work. Finally, section VI embraces the conclusions.

2. Related work

Electricity power demand over a period of time is a seasonal non-stationary timeseries. Many attempts have been made by statisticians to create forecasting models for this kind of time-series. The methods of electric power forecasting model can be roughly classified into linear, nonlinear and hybrid methods. Linear methods include Multiple Linear Regression (MLR) [15], Spectral Decomposition (SD) [16,17], Auto-Regressive Integrated Moving Average (ARIMA) [18], Expert System (ES) [19] and so on, in which the ARIMA model has strong underlying mathematical and statistical properties that can support to generate forecasting intervals in the easiest way. The model is also very versatile and can capture a lot of patterns in forecasting. That's why ARIMA models are the basic and most general form of time series forecasting techniques. G. Juberias et al. [20] proposed a new ARIMA model for hourly load forecasting. The model is based on a time series analysis methodology and includes the action of meteorology as an explanatory variable. It is successfully applied in the Central Region Control Center in Red Electrica de Espana. Hongming He et al. [21] proposed a high frequency forecast model based
on ARIMA to forecast hourly and quarter-hourly electricity demand for next few days ahead and got a good result in estimating the relationships between user’s demand and various variables. ARIMA’s major assumption is stationary (No trends/seasonality, constant level, variance, and constellations), but if the series gets seasonality, SARIMA [22] could be used.

Nonlinear methods contain Fuzzy Logic [23], Decision Tree (DT), Support Vector Regression (SVR) [26], Artificial Neural Network (ANN) [27-29], in which SVR possesses advantages of good for fitting and generalize, global optimal solution and fast calculation. It has always been a research hotspot. Sajjad Fattaheian-Dehkordi et al. [30] applied SVR to practical hourly data of the Greater Tehran Electricity Distribution Company and the parameters are selected by using a grid optimization process and an investigation on different kernel functions. Finally, acceptable results are obtained compared with ANN and the real data. It is also worth mentioning that in [31], a hourly power demand forecasting was finished at Salagatan 18 in Uppsala using SARIMA and SVR respectively, which demonstrated the validity of SARIMA and SVR for power demand forecasting.

Hybrid models are usually robust for complex problems and often improve performance. Yi Yang et al. [32] proposed a hybrid model based on the seasonal ARIMA model and BP neural network. The seasonal ARIMA model is adopted to forecast the electric load demand day ahead and then the follow-up residual series is forecasted by BP neural network. Finally, summing up the forecasted residual series and the forecasted load demand series got by seasonal ARIMA model. The result shows it is quite useful to improve the accuracy of STLF. S. Karthika et al. [33] proposed a hybrid model ARIMA-SVM is used to predict the hourly demand. ARIMA is used to predict the demand after correcting the outliers using Percentage Error (PE) method and its deviation is corrected using SVM. It is observed that the MAPE error got reduced and its convergence speed increased. Very recently, Muralitharan Krishnan et al. [34] combined SARIMA and LSTM for forecasting the time series data. The results of the simulation reveal that better prediction is achieved by the proposed hybrid model. Based on the idea of the above hybrid model, this paper proposes a hybrid method of SARIMA and SVR. First, SARIMA is used to remove the seasonality and fit the linear part of the sequence, and then SVR is used to make up the deviation. It turned out to be quite good.

3. Theoretical framework

3.1. Overview

In order to improve the performance of the STLF, we propose a hybrid method. In our method, the total power demand is regarded as the sum of the linear part and the non-linear part of the series. The linear part depends on the influence of some fixed factors, which is captured by SARIMA, while the non-linear part is determined by many uncertain interference factors and compensated by SVR. It can be expressed by the following Eq (1),

\[ Y_t = L_t + N_t(1) \]

Where at discrete time t, the total power demand is \( Y_t \), \( L_t \) is the linear component and \( N_t \) is the nonlinear component.

3.2. SARIMA

For more than half a century, the Box-Jenkins ARIMA linear models have dominated many areas of time series forecasting. One of the attractive features of the Box-Jenkins approach for forecasting is that ARIMA processes are a very rich class of possible models and it is usually possible to find a process which provides an adequate description to the data. However, power demand data has strong seasonality and nonstationary, so SARIMA model is preferred to better capture its characteristics, it can be written as Eq (2) [35],

\[ \phi_p(B)\phi_p(B^s)(1 - B)^d(1 - B^5)^D Y_t = \theta_q(B)\theta_q(B^5)a_t(2) \]

with \( \phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \).
\[
\begin{align*}
\phi_p(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \cdots - \phi_p B^{ps}, \\
\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q, \\
\theta_q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_q B^{qs}
\end{align*}
\]

Where \( B \) is the backward shift operator and \( a_t \) is a white noises series with mean 0 and variance \( \sigma^2_{Y_t} \) represents total electric power demand at time \( t \).

The above model is often denoted as \( ARIMA(p, d, q) \times (P, D, Q)_{s} \) in which the index \( s \) refers to the seasonal period, \((1 - B)^d\) and \((1 - B^s)^D\) are the regular and seasonal difference operators, \( \phi_p(B) \) and \( \Phi_p(B^s) \) are the regular and seasonal AR polynomials, and \( \theta_q(B) \) and \( \Theta_q(B^s) \) are called the regular and seasonal MA polynomials respectively.

### 3.3. SVR

SVR constitutes a new and promising approach to data regression. It focuses on finding a hyperplane that is as flat as possible and approximates the data to regress as well.

Given a dataset of \( N \) samples, \( i.e \{(x_i, y_i), i=1,2,\ldots,N\} \), in which \( x_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\} \in \mathbb{R}^n \), \( y_i \in \mathbb{R} \) is the actual value corresponding to \( x_i \). A nonlinear mapping \( \varphi(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^l(\ l > n) \) is defined to map the input data into a high dimensional feature space, which may have infinite dimensions. Then, in the high dimensional feature space, there theoretically exists a linear function to formulate the nonlinear relationship between input data and output data. Here is the SVR function, which is written as Eq (3),

\[
f(x) = W^T \varphi(x) + b \tag{3}
\]

Where \( f(x) \) denotes the forecasting values, \( W \) is the \( l \)-dimensional weight factor, the coefficients \( b \) is the adjustable factor. Different from traditional regression models, SVR can tolerate a maximum deviation \( \varepsilon \) between \( f(x_i) \) and \( y_i \). So it aims at minimizing the empirical risk as Eq (4) shows,

\[
R_{emp}(f) = \frac{1}{N} \sum_{i=1}^{N} \Theta_\varepsilon(y_i, W^T \varphi(x_i) + b) \tag{4}
\]

Where \( \Theta_\varepsilon(y_i, W^T \varphi(x_i) + b) \) is the \( \varepsilon \)-insensitive loss function and defined as Eq (5),

\[
\Theta_\varepsilon(y_i, W^T \varphi(x_i) + b) = \begin{cases} 
|W^T \varphi(x_i) + b - y_i| - \varepsilon, & \text{if } |W^T \varphi(x_i) + b - y_i| \geq \varepsilon \\
0, & \text{otherwise}
\end{cases} \tag{5}
\]

However, it might not be feasible if there exists no hyperplane that approximates the data responses with \( \varepsilon \) precision. To deal with this, slight error is allowed by the introduction of slack variable \( \xi^+ \) and \( \xi^- \) that are non zero if the point lies above or below the hyperplane respectively[26].

The optimization problem can be restated as Eq (6),

\[
\min_{w, b, \xi^+, \xi^-} R_e(W, \xi^+, \xi^-) = \frac{1}{2} ||W||^2 + C \sum_{i=1}^{N} (\xi^+_i + \xi^-_i) \tag{6}
\]

Subject to

\[
\begin{align*}
& y_i - W^T \varphi(x_i) - b \leq \varepsilon + \xi^+_i \\
& -y_i + W^T \varphi(x_i) + b \leq \varepsilon + \xi^-_i \\
& \xi^+_i \geq 0 \\
& \xi^-_i \geq 0 \\
& i = 1, 2, \ldots, N
\end{align*}
\]

Where \( C \) is the parameter that trades off training errors and the maximum distance between training data and the hyperplane.

After the quadratic optimization problem with inequality constraints is solved, the parameter vector \( W \) in Equation 3 could be obtained by the following Eq (6),

\[
W = \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-) \varphi(x_i) \tag{7}
\]

Where \( \alpha_i^+, \alpha_i^- \) are the Lagrangian multipliers and obtained by solving the quadratic program. Finally, the SVR function is updated as follows,

\[
f(x) = \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-) K(x_i, x) + b \tag{8}
\]

Where \( K(x_i, x_j) \) is named the kernel function, and its value equals the inner product of the two vectors, \( x_i \) and \( x_j \), in the feature space \( \varphi(x_i) \) and \( \varphi(x_j) \), respectively; that is, \( K(x_i, x_j) = \ldots \)
\( \varphi(X_i) \circ \varphi(X_j) \). Any function that meets Mercer’s condition[26] can be used as the kernel function.

The commonly used kernel functions are polynomial kernel, Gaussian kernel and sigmoid function, of which Gaussian kernel is the most widely used. It could be depicted as 
\[
K(X_i, X_j) = \exp \left( -0.5 \frac{|X_i - X_j|^2}{\sigma^2} \right),
\]
where \( \sigma \) is a tunable parameter. The Gaussian RBF kernel is not only easier to implement, but also capable of nonlinearly mapping the training data into an infinite dimensional space, thus, it is suitable to deal with nonlinear relationship problems.

4. Experiments and results

4.1. Dataset

For comparison with previous works, we used part of the publicly available power usage dataset provided by University of Massachusetts[36] which is also utilized by [14], it records the power consumption of one apartment with ID29 in 2016, and our proposed model would be trained on the hourly data of the first 28 days in September and tested on the following 2 days. Unlike the data used in [14], we only use historical data and exclude weather and calendar information. The data is visually displayed in the figure 1 below.

![Figure 1: Electrical power consumption for Apartment 29 hourly in July, 2016.](image)

4.2. Evaluation metric

In this paper, two metrics are used to evaluate the forecasting accuracy of the model. One is the mean absolute percentage error (MAPE) (Eq (9)), and the other is the mean squared error (MSE) (Eq (10)). Smaller values of the error metrics indicate higher forecasting accuracy.

\[
MSE = \frac{1}{n} \sum_{t=1}^{n} (A_t - F_t)^2
\]

\[
MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \times 100\%
\]
Where \( n \) is the total number of forecast values, \( A_t \) and \( F_t \) denote the actual and forecast value at time \( t \), respectively.

### 4.3. SARIMA

Through visual inspection in figure 1, we can easily see that the electrical power consumption has periodicity. In order to enhance this observation, we can draw the single side amplitude spectrum of the sequence, as shown in the figure 2.

![Single side amplitude spectrum (normalized)](image)

Figure 2: the single side amplitude spectrum of the electrical power consumption series.

Our sampling frequency is set as 720, and the figure above shows that there is a spike at 3, then followed by 30. In order to get a not too big period, 30 is chosen. So the period \( s \) of the sequence is \( 720/30 = 24 \), which is in line with our life rule. In order to select the parameters \( p, d, q, P, D, Q \) of SARIMA more accurately, we use the function `auto_arima` of the `pmdarima` package in Python to select parameters automatically and supplement with way of manual selection. The core idea of it is to minimize the AIC by means of cross validation. Through multiple experiments, the best AIC was obtained for \( p=4 \), \( d=1 \), \( q=0 \), \( P=2 \), \( D=0 \), \( Q=2 \), the resulting fitted model was therefore an \( ARIMA(4,1,0) \times (2,0,2)_{24} \).

The fitting is shown in the figure 3.
Then we analyze the residual sequence as follows in figure 4.

As shown in the above figure, the residual of the training set of the sequence after \textit{ARIMA}(4,1,0) \times (2,0,2)_{24} fitting is very similar to a white noise. Furthermore, Ljung box test is
carried out on it, and the p value is approximately 0.668, far greater than 0.05, so the residual sequence can be considered as a white noise, which shows that our seasonal model assumption is very reasonable.

4.4. SVR
First, RBF is chosen for SVR for its superiority of expressing nonlinearity, then there are three parameters in SVR model, $\sigma, C$ and $L$. In which $\sigma$ is the width parameter, $C$ is the penalty coefficient, and $L$ refers to predicting the current data with the previous $L$ data. We use the way of cross validation to set $\sigma = [0.1, 0.5, 1, 2, 3, 4, 5, 10]$, $C = [0.1, 0.5, 1, 2, 3, 4, 5, 10]$, and $L$ equals 1 to 10, respectively. Finally, the minimum MAPE is obtained from $\sigma=2$, $C=1$ and $L=4$. The fitting of test set is shown in the figure 5.

![Figure 5: forecasting result of the error series by SVR in test set.](image)

So our final forecasting result is the sum of the result after SARIMA and the residual prediction. Finally, in order to compare the effect of the hybrid model and the single models, we use the single SVR to predict the original sequence directly, and use the last 48 hours of the training set as the validation set to select parameters. The best result corresponds to the following parameters, $\sigma=0.5$, $C=10$, $L=4$. We call this method Our-SVR distinguishing from the SVR in [12]. The comparison of forecasting results of hybrid model and single model is depicted in figure 6.
As shown in the figure above, our proposed hybrid model compensates the error well, making it 29.048% lower than SARIMA in MAPE, but surprisingly, the single SVR performs even better than hybrid method.

Table 1. Performances of all models.

| Models         | MSE     | MAPE    |
|----------------|---------|---------|
| SVR[12]        | 9.693E-05 | 10.391% |
| BDTR[13]       | 8.853E-05 | 9.512%  |
| PowerLSTM[14]  | 6.923E-05 | 8.935%  |
| SARIMA         | 12.359E-05 | 11.994% |
| SARIMA-SVR     | 8.801E-05 | 8.510%  |
| Our-SVR        | 7.193E-05 | 6.843%  |

Compare our hybrid model with recent three models[12-14], in which [13] is a way of Boosted Decision Tree Regression (BDTR) adopted by Bansal et al, [12] is adopted by Yu et al using a SVR method to predict power consumption with weather information and energy usage history information, and [14] is a Long Short-Term Memory (LSTM) based method using the data after feature selection.

To be consistent with [14], when calculating metric, our data will be divided by 60 on the original basis.

As the results demonstrate in table 1, our SARIMA-SVR model has 28.789% improvement in MSE and 29.048% improvement in MAPE compared with the single SARIMA model, but it is inferior to Our-SVR. When having a comparison with [12][13][14], our SARIMA-SVR brings an improvement in MAPE by 18.102%、10.534%、4.757%, respectively, and its MSE is between BDTR and PowerLSTM. Last, Our-SVR get best scores in both MSE and MAPE.

5. Discussion and future work
We propose a hybrid model based on SARIMA and SVR to forecast hourly electric power demand, which has the advantages of simplicity, capable of capturing linear information and compensating the
residual caused by nonlinear factors at the same time, which make it more effective than a single SARIMA model and become more robust. Moreover, compared with other methods with complex structure and multiple kinds of data, it can still achieve good results. However, in this experiment, we found that using the single SVR with the data of the previous four hours to predict the current value, has achieved better results than all models both on MSE and MAPE. We speculate that there is a strong correlation between the data, and our SVR can just exactly capture the regular information of the sequence. In this paper, we choose the dataset adopted by [14] for the sake of comparison, while in the future, dataset with larger data volume and more complex pattern would be utilized to test the capability of the model.

6. Conclusion
Electric power demand forecasting is a critical task for smart grid to achieve efficiency and reliability in demand response and resource allocation. Accurate forecasting can greatly reduce the cost and loss caused by imbalance between supply and demand.

The contribution of this paper is two-fold. Firstly, we proposed and implemented a hybrid method based on SARIMA and SVR. Its advantage is that it can compensate the influence of nonlinear factors while fitting the linearity. It becomes more robust and performs better than a single SARIMA model. Secondly, we provided a comparison of three recent power demand forecasting methods on a public dataset. The results show that our method achieve best scores in MAPE than all of the above methods, but higher MSE than PowerLstm which has complex network and uses data from multiple sources.

Acknowledgements
The authors gratefully acknowledge the financial support from National Key R&D Program of China (2018YFB1500800) and Self-built project of State Grid E-Commerce Co., Ltd. In 2010(1700/2020-72001B)

References
[1] Pelka P., Dudek G. (2019) Pattern-Based Forecasting Monthly Electricity Demand Using Multilayer Perceptron. In: Rutkowski L., Scherer R., Korytkowski M., Pedrycz W., Tadeusiewicz R., Zurada J. (eds) Artificial Intelligence and Soft Computing. ICAISC 2019. Lecture Notes in Computer Science, vol 11508. Springer, Cham
[2] Hassan, S.; Khosravi, A.; Jaafar, J.; Khanesar, M.A. (2016) A systematic design of interval type-2 fuzzy logic system using extreme learning machine for electricity load demand forecasting. Int. J. Electr. Power Energy Syst. 82: 1–10.
[3] Hernández-Callejo, Luis & Baladron, Carlos & Aguilar, Javier & Carro, Belén & Sanchez-Esguevillas, Antonio & Lloret, Jaime & Massana, Joaquim. (2014). A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings. Communications Surveys & Tutorials, IEEE. 16. 1460-1495. 10.1109/SURV.2014.032014.00094.
[4] H. S. Hippert, C. E. Pedreira, and C. R. Souza. (2001) “Neural Networks for Short-Term Load Forecasting: A review and Evaluation,” IEEE Trans. Power Syst., vol. 16, no. 1, pp. 44-51.
[5] K. Liu et al. (1996) “Comparison of very short-term load forecasting techniques,” in IEEE Transactions on Power Systems, vol. 11, no. 2, pp. 877-882, doi: 10.1109/59.496169.
[6] A. G. Bakirtzis, V. Petridis, S. J. Kiartzis, M. C. Alexiadis and A. H. Maisiss. (1996) "A neural network short term load forecasting model for the Greek power system," in IEEE Transactions on Power Systems, vol. 11, no. 2, pp. 858-863, doi: 10.1109/59.496166.
[7] S. Fan and L. Chen. (2006) "Short-term load forecasting based on an adaptive hybrid method," in IEEE Transactions on Power Systems, vol. 21, no. 1, pp. 392-401, doi: 10.1109/TPWRS.2005.860944.
[8] Chunming Yuan, Yuanying Chi and Xiaojing Li. (2019) A combined forecasting method for short term load forecasting based on random forest and artificial neural network. IOP
Conference Series: Earth and Environmental Science, 252.

[9] Xuan Che. (2018) Application of Improved Grey Model in Medium and Long Term Load Forecasting. IOP Conference Series: Earth and Environmental Science, 128.

[10] R. K. Agrawal, F. Muchahary and M. M. Tripathi. (2018) "Long term load forecasting with hourly predictions based on long-short-term-memory networks," 2018 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, pp. 1-6, doi: 10.1109/TPEC.2018.8312088.

[11] Zheng, J., Xu, C., Zhang, Z., Li, X. (2017) Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: 51st Annual Conference on Information Sciences and Systems, pp. 1–6

[12] Yu, W., An, D., Griffith, D., Yang, Q., Xu, G. (2015) Towards statistical modeling and machine learning based energy usage forecasting in smart grid. ACM SIGAPP Applied Computing Review 15(1), 6–16

[13] Bansal, A., Rompikuntla, S.K., Gopinadhan, J., Kaur, A., Kazi, Z.A. (2015) Energy consumption forecasting for smart meters. arXiv preprint arXiv:1512.05979

[14] Cheng Y., Xu C., Mashima D., Thing V.L.L., Wu Y. (2017) PowerLSTM: Power Demand Forecasting Using Long Short-Term Memory Neural Network. In: Cong G., Peng WC., Zhang W., Li C., Sun A. (eds) Advanced Data Mining and Applications. ADMA 2017. Lecture Notes in Computer Science, vol 10604. Springer, Cham

[15] N. Amral, C. S. Ozveren and D. King. (2007) "Short term load forecasting using Multiple Linear Regression," 2007 42nd International Universities Power Engineering Conference, Brighton, pp. 1192-1198, doi: 10.1109/UPEC.2007.4469121.

[16] D. D. Belik, D. J. Nelson, and D. W. Olive. (1978) “Use of the Karhunen-Loeve expansion to analyze hourly load requirements for a powerutility,” IEEE Power Engineering Society Winter Meeting, vol. A78, pp. 225-230.

[17] D. Farmer and M. J. Potton. (1968) “Development of online load-prediction techniques with results from trials in the south-west region of the CEGB,” Proc. Institution of Electrical Engineers, vol. 115, no. 10, pp.1549-1558.

[18] M. T. Hagan and S. M. Behr. (1987) "The Time Series Approach to Short Term Load Forecasting," in IEEE Transactions on Power Systems, vol. 2, no. 3, pp. 785-791, doi: 10.1109/TPWRS.1987.4335210.

[19] S. Rahman and R. Bhatnagar. (1987) “An expert system based algorithm for short term load forecast,” IEEE Trans. Syst., vol. 3, no. 2, pp.392-399.

[20] G. Juberias, R. Yunta, J. Garcia Moreno and C. Mendivil. (1999) "A new ARIMA model for hourly load forecasting," 1999 IEEE Transmission and Distribution Conference (Cat. No. 99CH36333), New Orleans, LA, USA, pp. 314-319 vol.1, doi: 10.1109/TDC.1999.755371.

[21] H. He, T. Liu, R. Chen, Y. Xiao and J. Yang. (2012) "High frequency short-term demand forecasting model for distribution power grid based on ARIMA," 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE), Zhangjiajie, pp. 293-297, doi: 10.1109/CSAE.2012.6272958.

[22] J W Taylor. (2003) Short-term electricity demand forecasting using double seasonal exponential smoothing, Journal of the Operational Research Society, 54:8, 799-805, DOI: 10.1057/palgrave.jors.2601589

[23] smail Z, Mansor R. (2011) Fuzzy logic approach for forecasting half-hourly Malaysia electricity load demand. In: Int inst forecast 2011 proc; p. 1–17.

[24] Yu, Z., Haghhighat, F., Fung, B.C., Yoshino, H. (2010) A decision tree method for building energy demand modeling. Energy and Buildings 42(10), 1637–1646

[25] Bansal, A., Rompikuntla, S.K., Gopinadhan, J., Kaur, A., Kazi, Z.A. (2015) Energy consumption forecasting for smart meters. arXiv preprint arXiv:1512.05979

[26] Vapnik V. (1995) The nature of statistical learning theory,NewYork,USA:Springer.

[27] Nasr GE, Badr EA, Younes MR. (2001) Neural networks in forecasting electrical energy
consumption. FLAIRS-01. In: Proc., American association for artificial intelligence.

[28] Ryu, S.; Noh, J.; Kim, H. (2016) Deep neural network based demand side short term load forecasting. Energies, 10, 3

[29] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu and Y. Zhang. (2019) "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," in IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 841-851, Jan. doi: 10.1109/TSG.2017.2753802.

[30] Fattaheian-Dehkordi, S., Fereidunian, A., Gholami-Dehkordi, H., and Lesani, H. (2014) Hour-ahead demand forecasting in smart grid using support vector regression (SVR), Int. Trans. Electr. Energ. Syst., 24, pages 1650–1663, doi: 10.1002/etep.1791

[31] D. Weinberg. (2019) ‘Electrical power demand forecasting’, Dissertation.

[32] Yang, Yi; Wu, Jie; Chen, Yanhua; Li, Caihong. (2013) A New Strategy for Short-Term Load Forecasting. Abstr. Appl. Anal. Special Issue (2013). Article ID 208964, 9 pages. doi:10.1155/2013/208964. https://projecteuclid.org/euclid.aaa/1393449776

[33] S. Karthika, V. Margaret and K. Balaraman. (2017) "Hybrid short term load forecasting using ARIMA-SVM," 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), Vellore, pp. 1-7, doi: 10.1109/IPACT.2017.8245060.

[34] M. Krishnan, Y. M. Jung and S. Yun, "Prediction of Energy Demand in Smart Grid using Hybrid Approach. (2020) " 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, pp. 294-298, doi: 10.1109/ICCMC48092.2020.ICCMC-00055.

[35] Little, Todd D., and William W.S. Wei. (2013) "Time Series Analysis." In The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis. : Oxford University Press.

[36] Umass smart* dataset - 2017 release. http://traces.cs.umass.edu/index.php/Smart/Smart