Artificial Intelligence Techniques for Electrical Load Forecasting in Smart and Connected Communities

Victor Alagbe[1], Segun I. Popoola[1,2(✉)], Aderemi A. Atayero[1],
Bamidele Adebisi[2], Robert O. Abolade[3], and Sanjay Misra[1]

[1] IoT-Enabled Smart and Connected Communities (SmartCU) Research Cluster,
Covenant University, Ota, Nigeria
victor.alagbe@stu.cu.edu.ng, atayero@covenantuniversity.edu.ng
[2] Department of Engineering, Manchester Metropolitan University,
Manchester M1 5GD, UK
segun.i.popoola@stu.mmu.ac.uk, b.adebisi@mmu.ac.uk
[3] Department of Electronic and Electrical Engineering,
Ladoke Akintola University of Technology, Ogbomoso, Nigeria
roabolade@lautech.edu.ng

Abstract. Electricity consumption has been on a rapid increase world-wide and it is a very vital component of human life in this age. Hence, reliable supply of electricity from the utility operators is a necessity. However, the constraints that electricity supplied must be the same as electricity consumed puts the burden on the utility operators to make sure that demand is equal to supply at any point in time in smart and connected communities. Load forecasting techniques, therefore, aim to resolve these challenges for the operators by providing accurate forecasts of electrical load demand. This paper reviews current and mostly used short term forecasting techniques, drawing parallels be-tween them; and highlighting their advantages and disadvantages. This paper concludes by stating that there is no one-size-fits-all technique for load forecasting problems, as appropriate techniques depend on several factors such as data size and variability and environmental variables. Different optimization techniques can be used whether to reduce errors and its variations or to speed up computational time, hence resulting in an improved model. However, it is imperative to consider the tradeoffs between each model and its different variants in the context of smart and connected communities.

Keywords: Artificial Intelligence · Load forecasting · Smart city · Neural network · Support Vector Machine

1 Introduction

Electricity consumption has been on a rapid increase worldwide and it is a very vital component of human life in this age. Several studies have shown the presence of direct correlation between growth in Gross Domestic Product (GDP) and

increase in energy consumption. A study of seventeen countries in Africa over a thirty-year period found a correlation between GDP and electricity consumption with varying amount of causality [1]. A seventeen-year study period of seventeen industries in Taiwan revealed a bi-directional causality effect and showed that increase in electricity consumption by 1% resulted in a corresponding 1.72% increase in GDP [2]. Electricity generation at any time need to be equal to the electricity consumption to prevent huge losses. However, energy demand can be highly infrequent and varies greatly; hence, there is a need to accurately forecast load demand of users to allow for adequate planning. Load forecast can be divided into three forecast periods, the short term which varies from less than one hour to one-week, medium term from one month to one year, and long term forecast from one year to ten years. Short Term Load Forecast (STLF) is very important in energy planning and management for ensuring optimum supply and consumption of data without wasting resources. However the load series is characterized by non-linearity, non-stationary and seasonal variations, hence making accurate forecast hard to achieve [3]. Also, STLF models are found to be affected by weather conditions due to the strong correlation between weather variables and electricity consumption [4]. In this review, we discussed relevant sections of literature which attempted to solve the short-term load forecast challenge with diverse models, drawing up parallels and similarities between them.

Load forecast techniques generally can be classified into two: parametric and non-parametric techniques. The parametric techniques are based on the assumptions that sampled data follow a probability distribution based on a fixed set of parameters and include linear regression, exponential smoothing, auto regressive moving average (ARMA) etc. While non-parametric techniques – which consist mainly of Artificial Intelligence – have flexible parameter sets which may increase or decrease depending on new information. Such examples include Support Vector Machine (SVM), Artificial Neural Networks (ANN), Extreme Learning Machine (ELM) etc. These techniques however have different strengths and weaknesses as pertaining to the research problem at hand.

This paper reviews current and mostly used short term forecasting techniques, drawing parallels between them; and highlighting their advantages and disadvantages. This paper concludes by stating that there is no one-size-fits-all technique for load forecasting problems, as appropriate techniques depend on several factors such as data size and variability and environmental variables. Different optimization techniques can be used whether to reduce errors and its variations or to speed up computational time, hence resulting in an improved model. However, it is imperative to consider the tradeoffs between each model and its different variants in the context of smart and connected communities.

2   State-of-the-Art Computational Intelligence Techniques

2.1   Artificial Neural Networks

ANN draws its inspirations from the human nervous system. It consists of large array of artificial neurons interconnected with each other. Each neuron has an

input, output, weight and threshold in a bio-mimicry of the brain neurons. ANN is defined as a large array of parallel combinations of simple processing units which are highly capable in modifying their parameters through a learning process in response to their environment to capture information [5]. ANN has advantage in not being explicitly programmed but in being trained hence, allowing it to capture and recognize invariance between variables which it was not trained for [6].

The neuron in an artificial neural network serves as the basic computing element. The neuron has a set of inputs which are passed into it with their respective weights which are then summed together and passed through an activation function and bias which determines its output. This output can be the final output or input to another neuron in the next layer.

A neural network primarily comprises of three layers which are: input, hidden and output layer. A network with no hidden layer is referred to as Single Layer Neural Network (SLNN) while a network with one or more hidden layer is Multilayer Perceptron (MLP). Neural networks, however, has different forms which vary largely from Feedforward NN, Recurrent NN, Convolutional N etc. and are usually trained using Backpropagation algorithm, Levenberg Marquardt algorithm etc.

2.2   Support Vector Machine

SVM was first introduced in by Boser et al. [7] as a training algorithm which maximizes the margin between the decision boundaries and the training patterns. The SVM were mainly inspired by statistical learning algorithm, based on risk minimization principles [8]. SVM is used to create separating hyperplane optimally with the feature space in higher dimension and in which subsequent findings can be grouped into different subsets [9]. SVM can also be used with non-linear boundaries as stated in [10], by transforming them to linear boundaries on a higher dimensional plane.

The SVM has found increased use in research and industry as it very effective in solving non-linear problems [11], and have demonstrated good performance in regards to preventing over-fitting leading to improved performances in solving time-series problems [12]. Over-fitting occurs when the learning algorithm learns too much from the data and become unable to generalize with new data sets. However, the SVM suffers from computational inefficiency as large computational resources are needed for higher dimensional feature space which is resolved using "kernel trick" [13]. Kernels are functions which modify how similarities are calculated between functions by representing inner products of observation rather than the actual observation.

The Support Vector Regression (SVR) however, is a form of SVM centered on regression, built and developed to minimize structural risks [14]. This is achieved by reducing the probability of the model from the training examples will achieve minimal error when introduced to new examples. The best solution to this risk minimization is achieved when the convex criterion function is reduced.

2.3   Fuzzy Logic

Fuzzy Logic can be said to be a form of Boolean logic that can be used to handle partial truth values i.e. values that are not completely true or false. This is derived from humans' reasoning methods which are not exact in nature but approximate. The theory of fuzzy sets, which is a generalization of classic theory of sets serves as the fundamental mathematical basis of fuzzy logic [15]. The switch of an object from classic theory to fuzzy logic theory is realized through the membership function, which can be triangular, Gaussian, trapezoidal etc. A fuzzy application is made up of four distinct layers: input; rule evaluation or inference; composition or aggregation of outputs; and defuzzification or output layers [16].

2.4   Extreme Learning Machine

ELM is a modern learning algorithm which was proposed in [17] by Huang et al. for feed forward neural networks with single hidden layer (SLFN) as an improvement in learning speed over traditional algorithms with the model exhibiting better generalization performance. In ELM, the input weights and hidden biases of the network are not tuned but randomly generated; hence, enabling the transformation of nonlinear system to linear system [18]. ELM has been implemented in several applications such as regression, classification, clustering, feature selection etc. The network layout of an ELM consist of the input layer as the first layer; a hidden layer which is activated by weighted projections of the input to non-linear sigmoid units as the second layer; and the output layer which comprises of units wi [19].

3   Current Trends and Research Directions

3.1   Neural Network for Electrical Load Forecasting

In using ANN for sub-hourly electricity usage forecast in commercial buildings, Chae et al. [20] used time indicator, environmental and operational data gotten from the metering infrastructure for three buildings with one utility billing system. The variable selection analysis was done with Random Forest yielding five variables: day type, outdoor dry-bulb temperature, outdoor relative humidity, operational condition and time indicator, together with previous electricity usage which were used as the predictor variables after which the model was built with Bayesian regularized neural network with Levenburg-Marquart (LM) back propagation algorithm. The result revealed that additional hidden layers and increased time delay resulted into better performance of the model with reduced average MSE and the min-max error range. In predicting day-ahead peak load demand for the Iranian National Grid, authors in [21] proposed an hybrid method comprising of wavelet decomposition and ANN. The high and low frequency components of the data were first captured using wavelet decomposition while the ANN which was optimized with genetic optimization was then

used for each component. The peak demand was then determined by reconstructing the low and high components. Results demonstrated better MAPE forecast errors and applicability of the model in real-time applications due to acceptable processing time.

In [22], Generalized Regression Neural Network (GRNN) was implemented to solve the non-linearity problem in which STLF is one of such. To select the spread parameter which determines the performance of GRNN, the fruit fly optimization decreasing step (SFOA) was employed; and this was integrated with weather variables and the periodicity of the short-term load to build a credible model. The model when com-pared to a Back Propagation NN yielded better performance in accuracy (SFOA-GRNN RMSE = 0.0018 vs. BPNN RMSE = 0.024), stability and convergence speed.

In [23], Khwaja et al. used Bagged Neural Network, which creates multiple sets of a dataset by sampling randomly with dataset, trains on each dataset and averages the results with an aim to improve on the performances of single neural network by reducing errors variation and estimation errors. The data used was collected from the New England Pool region using historical temperature data, load pattern history, hour of the day, day type etc. The model demonstrated consistent results when the multiple sets of ANN are more than or equal to 50, with appreciable performance regarding reduced MAPE compared to other techniques like single ANN, bagged regression trees, ARMA etc. However, a similar STLF in [24] using Boosted Neural Networks, BooNN – a process which uses several iterations of multiple ANN models and then minimizes the error in each iteration – which was trained with the Levenberg–Marquardt backpropagation (LMB) and the same datasets from New England Pool showed a better forecasting error and minimum variation when benchmarked with ANN, Bagged NN and other techniques; with reduction in computational time compared to BNN.

The backpropagation training algorithm is commonly used for feed-forward neural network, however, it has high energy, low rates of convergence, and poor generalization of the neural network [25]. Ozerdem et al. in [26] forecast hourly load supply using feedforward neural network. The authors developed two models, the first a feedforward neural network with particle swarm optimization and the other trained with back propagation learning algorithm using data from an energy company based in North Cyprus. The results demonstrated the suitability of the two networks for modelling energy demand with the back-propagation network achieving higher performance on MAE and MSE while the particle swarm optimized model achieved faster convergence with twice as fast training speed. The authors concluded on the choice of particle swarm optimized models for faster development of models limited impacts on error metrics performance.

Authors in [27] proposed a novel deep feedforward neural network for STLF. Evaluation of the model with three case studies of daily electricity consumption of cities in China showed the better forecasting accuracy of the proposed model com-pared to gradient boosting and random forest. The result also demonstrated the influence of weekly, monthly, and weather-related variables on electricity consumption of households. Also, He in [28] developed a deep neural network

model for STLF by processing the multiple types of input features individually and extracting the information using convolution neural networks (CNN), and modelling the implicit dynamics using recurrent components. Results on hourly loads of a city in North China showed flexibility and superiority of the method.

In predicting energy demand for smart grid, the authors in [29] proposed the use of neural network optimization approach. Using real time data from Pecan Street Inc, CNN was first used in predicting the energy demand and then Particle Swarm Optimization and Genetic Algorithm were used to optimize the results of the CNN model. The authors discovered from observing the results that the NNGA was better suited for short term energy predictions (achieving MSE of 0.391 compared to 0.495 of NNPSO) while the NNPSO was suited for long term (MSE of 0.408 compared to 0.429 of NNGA).

Authors in [30] proposed a new prediction method using Self-Recurrent Wavelet Neural Networks (SRWNN) as the forecast engine and Levenberg-Marquardt as the learning algorithm. The model was used to forecast hourly load demand of a building in a micro-grid. Results of the forecast showed that the SRWNN outperformed other forecasting models and demonstrated the ability of SRWNN to effectively adapt to variations and non-smooth behavior of the time series. The authors in [31] made use of the Advanced Wavelet Neural Network (AWNN) for very short term load forecast. The AWNN decomposes the complex load series into different frequencies and predict them separately. Evaluation of the model with Australian and Spanish electricity load data revealed that AWNN was the most accurate model for both datasets when compared with other models such as NN, LR, MTR etc.

In predicting the medium to long term load demand of commercial and residential buildings at 1 h resolution, Rahman et al. in [32] proposed two deep Recurrent Neural Network (RNN) models, which are also used for inputting missing data. Results predicting the load demand of the Public Service Building in Utah showed that the RNN models performed better in predicting the electric load profiles of buildings than a three-layer multi-layered perception model with the data imputation scheme performing better too with higher accuracy. In comparing the performances of SVM and ANN, the authors in [33] explored the data driven performance of ε-SVM Regression (ε-SVM-R) based on Radial Basis Function (RBF) and polynomial kernel; and two Nonlinear Autoregressive Exogenous Recurrent Neural Networks (NARX RNN) of different depths. Result using historical data of heating and cooling load demand of a non-residential district in Germany demonstrated the advantage of NARX RNNs over ε-SVM-R using computational time and accuracy as metrics.

Ruiz et al. in [34] proposed an Elman Neural Network model, a form of RNN together with genetic algorithm which optimized the weights of the model to forecast electricity consumption of public buildings with the aim of increasing energy efficiency and hence result in energy savings. The proposed model was based on electricity consumption data collected from buildings in University of Granada. The test results of the model showed a 61% improvement benchmarked

against the NAR and NARX models, with MSE of 0.005085 for model without temperature and 0.004413 for model including temperature.

Furthermore, Ko et al. in [35] proposed a hybrid method comprised of RBF Neural Network (RBFNN), SVR and Dual Extended Kalman Filter (DEKF). The SVR is firstly used to deduce the initial parameters and the neural network's structure while the DEKF is the learning algorithm used to optimize the parameters determined by the SVR; the optimized RBFNN is then finally used to perform the forecast. Using da-tasets form the Taipower Company with three case scenarios to evaluate the multi-day ahead forecast of the hybrid model, the proposed model SVR-DEKF-RBFNN demonstrated better forecasting performances in robustness, stability and accuracy when compared to other hybrid methods such as DEKF-RBFNN and gradient descent RBFNN (GRD-RBFNN).

3.2   Support Vector Machine for Electrical Load Forecasting

Kernel-based methods such as the SVR has shown tremendous success in STLF applications. However, the performance of such methods depends on choosing suitable kernel functions for the learning target. Che and Wang in [36] proposed a combinational method in addressing this issue using datasets from New South Wales and California with differing characteristics to compare the proposed model with other individual SVR models that are kernel-based. The result showed that the combined model, P-KSVR-CM, a combination of four kernel functions – linear, Gaussian, tanh and polynomial kernels resulted in an increased in forecasting accuracy when compared to the best performing individual SVR models.

Chen et al. in forecasting hourly load demand of a non-stationary operated hotel developed hybrid support vector regression model combined with multi-resolution wavelet decomposition (MWD) in [37]. The WMD which was used to remove random noises from the load series and to better illustrate the special periodic features. Results of the model with and without MWD were compared, and the MWD was found to reduce the deviations slightly only when ε the non-sensitive loss function is higher than 0.1. In comparison, operating under the belief that good feature selection is very import in influencing prediction accuracy, Yang et al. in [38] used the Auto Correlation Function (ACF) for its feature selection while the Least Squares Support Vector Machines (LSSVM) was used for the forecast and was optimized using the Grey Wolf Optimization (GWO) and Cross Validation (CV). The proposed model, AS-GCLSSVM was used for the week ahead half hourly load forecast and results demonstrated the effectiveness of the approach in improving forecasting accuracy compared with other benchmark models though, the algorithm is time consuming and complicated.

Meanwhile, Niu et al. in [39] proposed the ant colony optimization so as to reduce the processing of large datasets and the resultant slow processing speed. The optimization technique is employed to discover optimal feature subsets in the data resulting in more accurate selection compared to other techniques like PCA, entropy-based feature selector etc. Using the selected features in forecast-

ing short term load with Support Vector Regression, the novel method achieved better forecasting accuracy compared to single SVM and BPNN highlighting the importance of using data mining techniques for SVM-learning system.

Tong et al. in [40] proposed a deep learning theory in order to handle massive data accumulated from different sensors. The model firstly processes the features from the historical load and temperature datasets using stacked denoising auto-encoders (SDA). The model then trains a support vector regression for day-ahead load fore-cast. Observations from the model showed the capabilities of the model to describe and forecast the load tendency with better accuracy and minimal error. Comparison of the model with SVR and ANN also revealed better performance with low MAPE values. In demonstrating the applicability of online support vector for STLF, Vrablecová et al. in [41] compared the accuracy performances of ten state of the art model using Irish CER dataset. The results showed that tree-based ensemble methods such as random forests, bagging etc. achieved similar or superior forecasting accuracy than the online SVR. Furthermore, online SVR had comparable accuracy result with other online load forecast methods. To improve efficiency and computational accuracy, Li et al. in [42] proposed a the use of sub-sampled support vector regression ensemble (SSVRE). The SSVRE was also combined with swarm optimization learning thereby ensuring that each individual SVR ensemble has enough diversity for STLF. The results showed the superior performance and reduced uncertainties of the SSVRE model. Furthermore, a Guassian SVM short term load forecasting was developed by Zheng et al. in [43] operating on the basis that of the admissible translation-invariant function of SVM of pth-derivative Gaussian wavelet when p is an even number. The method was constructed with the wavelet kernel function and the parameters were optimized using stochastic focusing search (SFS) algorithm. The results showed a better forecasting accuracy compared to Morlet wavelet SVM and Gaussian SVM with the lowest MAPE and MRE.

To address the non-linearity of electric load caused by seasonal variations. The authors in [44] proposed a novel decomposition-ensemble model containing SVM, Single Spectrum Analysis (SSA), Autoregressive Moving Average (ARIMA) and cuckoo search algorithm. Using half-hourly datasets from New South Wales and Singaporean hourly load datasets, the proposed model resulted into higher forecasting accuracy compared to eight other models such as SVM, CS-SVM, BNN, SSA-SVM etc.; with all performance metrics of half-hourly loads better than the hourly loads, indicating the strength of SSA-SVM-CA in its robustness.

3.3   Fuzzy Logic for Electrical Load Forecasting

The authors in [45] used the triangular fuzzy-number models for forecasting. The first being the triangular fuzzy-number grey model (TFGM) which is used for TF series having poor fluctuation, and others being an amended TFGM with BP neural networks (NNTFGM) and the SVMTFGN. The result of load forecast in a district in China showed a better performance of the amended models with

smaller MREs (NNTFGM = 7.54%, SVMTFGN = 7.99%) compared to the 23.74% of the TFGM.

Coelho et al. in [46] proposed a self-adaptive evolutionary model for STLF in a micro-grid environment, by applying a bio-inspired optimizer called GES to determine the weights' optimal values and fuzzy rules. The result of the meta-heuristics model showed better forecast accuracy with lesser computation time compared with other hybrid model in [47] with low variations in the forecast errors, with the model being suitable for micro and large grids.

In forecasting the short term thermal power demand of HVAC, the authors in [48] proposed an estimation method for the usage activity pattern of the HVAC. The authors used Recurrent NN for the dynamic activity prediction together with ANFIS for the demand prediction model. The use of specialized modeling structure in instances where power demand is not readily available was validated by the results in respect of accuracy and performance of the model.

3.4   Extreme Learning Machine for Electrical Load Forecasting

Chen et al. in improving the accuracy of forecasts of the ELM model in [49] used a novel method optimized with the empirical mode decomposition which removed noises and decomposed the load series together with mixed method of RBF and UKF kernel for optimum selection of kernels which greatly influence the performance of ELM. The novel method resulted in better accuracy than RBF-ELM, UKF-ELM, mixed ELM etc.; and verification of the model using three datasets showed similar results.

In order to solve the challenge caused by small capacity and higher randomness in micro-grids, the authors in [47] proposed a hybrid model with parameter optimization. The model which included EMD for time series decomposition, EKF and KELM for prediction algorithms and PSO for optimization polled results with high accuracy and efficiency using four different datasets. However to improve accuracy and achieve low reduction rates, the authors in [50] developed a Data Framework Strategy to construct features pool and the genetic algorithm binary improved cuckoo search for lowest reduction rates. The authors then used the ELM as the forecasting model and the hybrid model achieved a high and robust accuracy with minimum number of effective features used.

4   Conclusion

Computational Intelligence techniques have proved important and extremely useful in forecasting energy demand in the short term for effective operation of the grid by the utility operators. Review of past works have shown that there is no one coat fits all in the short term forecast problem as each model has its own merits and demerits; and different datasets have different peculiarities. However, basic fundamentals for solving the short term forecast problems for future works can be inferred.

The reviewed works in this paper has expounded on the importance of data pre-processing which includes imputation of missing variables and treating of

outlier values, as these errors in the data which are regarded to as noise may affect severely the performance of the model. Furthermore, there is also the need feature engineering. This is a process whereby variables which are very significant and with high correlation to the variable to be forecast are chosen to improve the learning abilities of the model and hence its performance. Also, there are evidences of weather variables in improving the accuracy of load forecast models, as there are correlations between electricity consumption and weather conditions.

Furthermore, different optimization techniques can be used whether to reduce errors and its variations or to speed up computational time, hence resulting in an improved model. However, it is imperative to consider the trade-offs between each model and its different variants in the specific context it will be used in before committing to a short-term load forecast model.

References

1. Wolde-Rufael, Y.: Electricity consumption and economic growth: a time series experience for 17 African countries. Energy Policy **34**(10), 1106–1114 (2006)
2. Lu, W.-C.: Electricity consumption and economic growth: evidence from 17 Taiwanese industries. Sustainability **9**(1), 1–15 (2016)
3. Mohan, N., Soman, K.P., Sachin Kumar, S.: A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. Appl. Energy **232**, 229–244 (2018)
4. Amara, F., et al.: Household electricity demand forecasting using adaptive conditional density estimation. Energy Build. **156**, 271–280 (2017)
5. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall, Upper Saddle River (1999)
6. Luger, G.F.: Artificial Intelligence: Structures and Strategies for Complex Problem Solving, 6th edn. Pearson (2009)
7. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory (1992)
8. Vapnik, V.N.: An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**(5), 988–999 (1999)
9. Kuster, C., Rezgui, Y., Mourshed, M.: Electrical load forecasting models: a critical systematic review. Sustain. Cities Soc. **35**, 257–270 (2017)
10. Auria, L., Moro, R.A.: Support Vector Machines (SVM) as a Technique for Solvency Analysis. DIW Berlin, German Institute for Economic Research (2008)
11. Zhao, H.X., Magoulés, F.: A review on the prediction of building energy consumption. Renew. Sustain. Energy Rev. **16**(6), 3586–3592 (2012)
12. Hong, T., Fan, S.: Probabilistic electric load forecasting: a tutorial review. Int. J. Forecast. **32**(3), 914–938 (2016)
13. Adhikari, R., Agrawal, R.K.: An introductory study on time series modeling and forecasting (2013)
14. Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**(3), 199–222 (2004)
15. Elias, C.N., Hatziargyriou, N.D.: An annual midterm energy forecasting model using fuzzy logic. IEEE Trans. Power Syst. **24**(1), 469–478 (2009)

16. Chen, S.X., Gooi, H.B., Wang, M.Q.: Solar radiation forecast based on fuzzy logic and neural networks. Renew. Energy **60**, 195–201 (2013)
17. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K.: Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE International Joint Conference (2004)
18. Zhu, Q.-Y., et al.: Evolutionary extreme learning machine. Pattern Recogn. **38**(10), 1759–1763 (2005)
19. Tissera, M.D., McDonnell, M.D.: Deep extreme learning machines: supervised autoencoding architecture for classification. Neurocomputing **174**, 42–49 (2016)
20. Chae, Y.T., et al.: Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. Energy Build. **111**, 184–194 (2016)
21. Moazzami, M., Khodabakhshian, A., Hooshmand, R.: A new hybrid day-ahead peak load forecasting method for Iran's National Grid. Appl. Energy **101**, 489–501 (2013)
22. Hu, R., et al.: A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm. Neurocomputing **221**, 24–31 (2017)
23. Khwaja, A.S., et al.: Improved short-term load forecasting using bagged neural networks. Electr. Power Syst. Res. **125**, 109–115 (2015)
24. Khwaja, A.S., et al.: Boosted neural networks for improved short-term electric load forecasting. Electr. Power Syst. Res. **143**, 431–437 (2017)
25. Zhang, J., et al.: Enhancing performance of the backpropagation algorithm via sparse response regularization. Neurocomputing **153**, 20–40 (2015)
26. Ozerdem, O.C., Olaniyi, E.O., Oyedotun, O.K.: Short term load forecasting using particle swarm optimization neural network. Procedia Comput. Sci. **120**, 382–393 (2017)
27. Guo, Z., et al.: A deep learning model for short-term power load and probability density forecasting. Energy **160**, 1186–1200 (2018)
28. He, W.: Load forecasting via deep neural networks. Procedia Comput. Sci. **122**, 308–314 (2017)
29. Muralitharan, K., Sakthivel, R., Vishnuvarthan, R.: Neural network based optimization approach for energy demand prediction in smart grid. Neurocomputing **273**, 199–208 (2018)
30. Chitsaz, H., et al.: Short-term electricity load forecasting of buildings in microgrids. Energy Build. **99**, 50–60 (2015)
31. Rana, M., Koprinska, I.: Forecasting electricity load with advanced wavelet neural networks. Neurocomputing **182**, 118–132 (2016)
32. Rahman, A., Srikumar, V., Smith, A.D.: Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. Appl. Energy **212**, 372–385 (2018)
33. Koschwitz, D., Frisch, J., van Treeck, C.: Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX Recurrent Neural Network: a comparative study on district scale. Energy **165**, 134–142 (2018)
34. Ruiz, L.G.B., et al.: Energy consumption forecasting based on Elman neural networks with evolutive optimization. Expert Syst. Appl. **92**, 380–389 (2018)
35. Ko, C.-N., Lee, C.-M.: Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. Energy **49**, 413–422 (2013)
36. Che, J., Wang, J.: Short-term load forecasting using a kernel-based support vector regression combination model. Appl. Energy **132**, 602–609 (2014)

37. Chen, Y., Tan, H., Song, X.: Day-ahead forecasting of non-stationary electric power demand in commercial buildings: hybrid support vector regression based. Energy Procedia **105**, 2101–2106 (2017)
38. Yang, W., Kang, C., Xia, Q., et al.: Short-term probabilistic load forecasting based on statistics of probability distribution of forecasting errors. Autom. Electr. Power Syst. **30**(19), 47–52 (2006)
39. Niu, D., Wang, Y., Wu, D.D.: Power load forecasting using support vector machine and ant colony optimization. Expert Syst. Appl. **37**(3), 2531–2539 (2010)
40. Tong, C., et al.: An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. J. Parallel Distrib. Comput. **117**, 267–273 (2018)
41. Vrablecová, P., et al.: Smart grid load forecasting using online support vector regression. Comput. Electr. Eng. **65**, 102–117 (2018)
42. Li, Y., Che, J., Yang, Y.: Subsampled support vector regression ensemble for short term electric load forecasting. Energy **164**, 160–170 (2018)
43. Zheng, Y., Zhu, L., Zou, X.: Short-term load forecasting based on Gaussian wavelet SVM. Energy Procedia **12**, 387–393 (2011)
44. Zhang, X., Wang, J.: A novel decomposition-ensemble model for forecasting short-term load-time series with multiple seasonal patterns. Appl. Soft Comput. **65**, 478–494 (2018)
45. Zeng, X.-Y., et al.: Triangular fuzzy series forecasting based on grey model and neural network. Appl. Math. Model. **40**(3), 1717–1727 (2016)
46. Coelho, V.N., et al.: A self-adaptive evolutionary fuzzy model for load forecasting problems on smart grid environment. Appl. Energy **169**, 567–584 (2016)
47. Liu, N., et al.: A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids. Appl. Energy **129**, 336–345 (2014)
48. Sala-Cardoso, E., et al.: Activity-aware HVAC power demand forecasting. Energy Build. **170**, 15–24 (2018)
49. Chen, Y., et al.: Mixed kernel based extreme learning machine for electric load forecasting. Neurocomputing **312**, 90–106 (2018)
50. Jiang, P., Liu, F., Song, Y.: A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting. Energy **119**, 694–709 (2017)