An Enhanced Very Short-Term Load Forecasting Scheme Based on Activation Function

Adamu Sani Yahaya
COMSATS University Islamabad,
Islamabad 44000, Pakistan.
ayahaya.it@gmail.com

Nadeem Javaid *
COMSATS University Islamabad,
Islamabad 44000, Pakistan.
*Correspondence: nadeemjavaidqau@gmail.com
Website: www.njavaid.com

Kamran Latif
National Institute of Electronics,
Islamabad 44000, Pakistan.
klatif0077@gmail.com

Amjad Rehman
MIS Department COBA, Al Yamamah University,
Riyadh 11512, Saudi Arabia.
drrehman70@gmail.com

***Abstract*—In this paper, we proposed a framework for accurate load forecasting which consists of two stage processes; feature engineering and classification. Feature engineering consists of feature selection and extraction. Relevant features are selected by combining Decision Tree (DT) and Recursive Feature Elimination (RFE) techniques. Moreover, Linear Discriminant Analysis (LDA) technique is used to further improve the selected features in terms of redundancy and dimensionality reduction. To forecast the electricity load, an improved feedforward multilayer perceptron classifier is applied. Half a day ahead forecasting experiment is conducted by using the proposed framework. At the end, forecasting performance is examined by using Root Mean Square Error, Mean Absolute Error, Mean Square Error and Mean Absolute Percentage Error. Simulation results show higher accuracy of our proposed scheme with 1.397% as compared to the existing scheme.**

***Index Terms*—Electricity load forecasting, Feature selection, Features extraction, Feedforward multilayer perceptron, Accuracy.**

I. Introduction

Intelligent power grid, also known as smart grid (SG) is a power system that efficiently manages consumption, distribution and generation of energy through advanced technologies. The communication between the utilities and consumers in smart grid is done in two-way peer-peer i.e., from utility to consumer and vice versa [1]. In this world, utilization of energy is necessary and valuable asset. Increase in energy consumptions from various sectors such as commercial, residential, industrial and transportation escalated burden on traditional power grids. Increase in load consumption forces utility to migrate from traditional power grid to smart grid. With the implementation of different techniques on utility, power grid and demand side, Smart grid manages the consumption, distribution and generation.

In smart grid, data analysis and information communication technology are in use into every angle of the energy system such as energy distribution, generation, transmission and also in appliances of consumer. Smart grid technology is rapidly increasing in order to meet the increasing demand for energy delivering in a flexible, robust and in cost effective manner. As a result of the increase in data generated from smart grid, data analysis becomes the topic of discussion. Valuable information from the huge amount of data generated from smart grid can easily be extracted by data analytics techniques. The valuable information can be use in algorithms development and scheme designed.

List of Symbols

Symbols	Description
$f(x)_{improved}$	Improved activations function
$f(x)$	Logistic sigmoid function
$g(x)$	Rectify linear unit function
P_i	The forecasted electricity load value
R_i	The actual electricity load value

Machine learning and pattern recognition are among the data analytics techniques that been developed and adopted especially in the area of load and price forecasting, data mining of consumer behaviors and fault detection in energy networks [2]. The research in electricity load forecasting is not new, is begin in earlier 1965 [3]. Forecasting electricity load accurately is paramount important in decision-making process for economic load dispatch, power unit commitment, contingency, power system operation, scheduling, security and so on [3], [4].

Proper decision may results in minimization of cost of electricity and power loss [5]. According to [4] one percent (1%) reduction of mean absolute percentage error in load forecasting saved 10,000Mw electricity energy which may lead to saving approximately $1.6 million per year. With these problems, researchers paid more attention on how to resolve the power scheduling problem. Many techniques had been used to optimize power scheduling problem [6]–[10]. However, a precise load forecasting model is required to properly plan on how to effectively managed power grid resources. Redundancy and nonlinearity in the historical electricity load data causes huge challenges in load forecasting. The role of electricity load

978-1-5386-8125-1/19/$31.00©2019 IEEE

Abbreviation Table

Abbreviation	Description
AC	AutoCorrelation
CFS	Correlation-Based Selection
DE	Differential evolution
DT	Decision Tree
EMD	Empirical mode decomposition
ELM	Extreme learning machine
FOA	Fruit Fly Optimization
FS	Feature Selection
GPSO	Global Best Particle Swam Optimization
GCA	Grey Correlation Analysis
GWO	Grey Wolf Optimization
IMF	Intrinsic mode functions
KELM	kernel extreme learning machine
KPCA	Principle Component Analysis and Kernel function
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
LTLF	Long-Term Load Forecasting
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MTLF	Medium-Term Load Forecasting
MSE	Mean Square Error
PSO (SDPSO)	new switching delayed PSO.
QLD	New Queensland
RBFNN	Radial basis function neural network
RF	Random Forest
RF	RRelief
RFE	Recursive Feature Elimination
RMSE	Root Mean Square Error
SDA	Stacked denoising autoencoder
SG	Smart Grid
STLF	Short-Term Load Forecasting
SVM	Support Vector Machine
VSTLF	Very Short-Term Load Forecasting

forecasting is increasing in electricity generation, system operation, transmission and storage while forecasting accurately is becoming a major concern in the management [11]. Therefore, many researchers focused on how to promote the accuracy and robustness of the electricity load forecasting. According to [3], [12] and [13] electricity forecasting horizons are categorized into Very Short-Term Load Forecasting (VSTLF), Short-Term Load Forecasting (STLF), Medium-Term Load Forcasting (MTLF) and Long-Term Load Forecasting (LTLF) as shown in Table I. In this research, very short-term load forecasting horizon will be the main focus. Modified feedforward multilayer perceptron will be used as classifier to forecast the electricity load.

A. Motivation

A lot of researches had been conducted in electricity load forecasting as in [3], [11]–[14]. Although, optimal result from each model is obtained. However, forecasting accuracy of load and price needs further improvement. As consequence, this motivated us to propose a model that provides better accuracy as well as to improve the efficiency of the existing models in literature.

B. Problem Statement

In [13], recurrent deep neural network and feedforward deep neural network models performance are compared on the basis of accuracy and computational performance in which they consider short-term forecasting. Also [14], Global Best Particle Swam Optimization (GPSO) is proposed to optimized Neural Network bias value in training processes which decrease the error and increase accuracy.

In this paper, very short-term load forecasting is considered to predict the variations in load and activation function is modified to increase accuracy and decrease error which are not considered in [13], [14]. Although, the dataset used in this research is not the same as that of the benchmark models, However a reasonable result is obtained. The neural network is improved by combining logistics sigmoid and rectify linear unit activation functions.

C. Contribution

The main contributions of this research is as follows:

- Proposing an improved electricity load forecasting system.
- Combination of Decision tree and Recursive feature elimination are used in features selection to find the most relevant features.
- To reduce features redundancy and dimensionality, linear discriminant analysis is applied.
- Activation function of Feedforward multilayer perceptron is improved by combining logistics sigmoid and rectify linear unit functions which enhanced the accuracy and decreases the errors in the forecasted result.

The term prediction and forecasting are alternatively used throughout this paper. The remaining part of this paper is planned as follows: Section 2 contained the related works, section 3 gives details of the techniques (Methodology) used to build up the model and section 4 describes system model, dataset and their experimental setup. Section 5 describes the performance estimation while the simulations result obtained are discussed in section 6. Finally, The conclusion of the research is discussed in Section 7.

II. Related Work

Lots of researches had been done in STLF in order to improve forecasting accuracy. Kalman filtering, linear regression methods, exponential smoothing, gray forecasting and ARIMA [16]–[18] were proposed in the primal stage. These forecasting models are good in forecasting linear problem however are not sufficient when processing more complex non-linear load problem forecasting. Due to the limitations of these models, accurate forecasting cannot be achieved. The aforementioned models are based on mathematical statistics. Automated and more intelligent artificial neural network techniques were developed to overcome the complex relationship and non-linearity in electricity time series. When an ANN was developed, fewer parameters were needed to step up the model. This directly affects the forecasting performance. This reason influence the researchers to come up with hybrid models [19] which leads to the addition of numerous smart optimizations techniques like Fruit Fly Optimization Algorithm (FOA), Grey Wolf Optimization Algorithm (GWO) e.tc [20], [21]. In [5] cuckoo search is incorporated with Support Vector Machine (SVM) to improve its parameters. The performance of the

TABLE I: Forecasting Horizons

S/N	Category	Period range	Aim
1	VSTLF	The range is between Few minutes, 30 minutes to approximately one day.	To control and adjust demand load and price in real-time
2	STLF	Started with one day to a week and upto a month	To makes a real-time plan for optimal generator unit commitment and economic dispatch
3	MTLF	start with one month upto approximately one year	To maintain the balance between generation and demand for few months to one year
4	LTLF	More than one year	To plan for the future electricity network conditions

model significantly increase in the experimental results. In [22] New Switching Delayed PSO (SDPSO) is used to enhance the Extreme Learning Machine (ELM) techniques. The experimental result of the model outperforms Radial Basis Function Neural Network (RBFNN) model. However, experiment focus only in STLF. Finding shows that combination of two or more models can obtain better forecasting result than the single ANN model. A robust combination model is proposed by many researchers to integrate more than two forecasting models [23]. This approach balance the distribution of forecasting risk and gain benefits of individual models.

Some researchers observed that best forecasting result is not guarantee with models that has unavoidable deficiency. As a result of that combining two or more forecasting model is proposed to resolve the limitations. A hybrid model of Boosting algorithm and a multistep forecast are combined to produce accurate forecasting results in electricity market [24]. In [25], Auto Regressive Moving average and Self-adapting Particle Swarm Optimization improve the hybrid combination of the Kernel Extreme Learning Machine (KELM) and wavelet transform. The result produce by the model was remarkable. however, computational time is not considered. A short-term electricity price was forecasted in [26] using Stacked Denoising Autoencoder (SDA). The result of the model was compared with multivariate adaptive regression splines, classical NN, SVM and least absolute shrinkage and selection operators and better result is obtained. The price forecasting accuracy decreases in [27] when forecasting wider range i.e. weeks, month. Four deep learning techniques were used to predict the electricity price. The techniques used are DNN, LSTM-DNN, GRU-DNN and CNN model.

Features Selection (FS) contributed immensely in machine learning to establish a powerful and reliable model. Feature selection is a process whereby a set of important features with high correlation with output are selected [28]. Researchers used different set of feature selection method. For example, in [29] four feature selection techniques: RRelief (RF), Mutual Information, Correlation-Based Selection (CFS), and Autocorrelation (AC) were evaluated. The corresponding experimental results show that AC-NN and RF-NN gives a better performance. In [30] Grasshopper and evolutionary population dynamics techniques are used to select more relevant features. This method performed very well when handling a larger set of features and on the other hand, its performance reduced drastically with a smaller set of input features. Forecasting performance increasing with more relevant features and vice versa. In [31], K-means and gradient boosting-based weighted techniques are used to find the similarities between the actual and predicated days. Empirical Mode Decomposition (EMD), decompose the similar days to residual and several Intrinsic Mode Functions (IMFs) and LSTM were used to forecast each decomposed result. New forcasting model is established in [32]. a hybrid feature selection is proposed by combining Relief-F algorithm and Random Forest (RF) together with Grey Correlation Analysis (GCA) in order to minimized the redundancy. Principle Component Analysis and Kernel function (KPCA) are integrated to reduced dimensionality. In the end, SVM is improved by using differential evolution (DE) to forecast the electricity price.

III. Methodology

A. Decision Tree

Decision tree aim to find the purest child possible nodes with minimum split in order to classified the instances. In this model "Information gain" is used as the attribute selection measures to find purest child nodes.

Algorithm 1 Decision Tree

S, where S= set of classified instances
Require: $S \neq \emptyset, numattribute > 0$
procedure Decision Tree:
    Repeat:
    $maxGain \leftarrow 0$
    splitA $\leftarrow$ *null*
    $e \leftarrow$ *Entropy(attributes)*
    for all a in S do
        gain $\leftarrow$ *InformationGain(a,e)*
        if gain $>$ *maxGain then*
            maxGain $\leftarrow$ *gain*
            splitA $\leftarrow a$
        end if
    end for
    Partition(S, splitA)
    Untill all partition processed

B. Recursive Feature Elimination

Description on how iterative Recursive Feature Elimination (RFE) works is provided as follows:

1) The classifier will be trained.
2) The ranking criterion for all features will be calculated.
3) Eliminate all the features with the lowest ranking criterion.

The algorithm of SVM-RFE is described in algorithm 2

Algorithm 2 Recursive Feature Elimination

procedure RFE:
classifier training : $\alpha = train - SVM(x, y)$
weight vector is computed of the demension legth :
$w = \sum_k \alpha_k x_k y_k$
ranking criterias is computed : $c_i = (w_i)^2$
Find the feature with the smallest ranking criteria :
$f = argmain(c)$
Eliminate feature with c=f

The stop condition can be based on the number of features or the classification accuracy.

C. Linear Discriminant Analysis

The pseudocode for LDA is depicted in Algorithm 3. Where 1_n is a vector of all ones and $1_n \in \mathbb{R}$, with the appropriate dimension for each class i = 1, 2. Linear Discriminant Analysis proceeds to compute the in-between and within class scatter matrices, B and S, after The data is dividing into two groups G1 and G2. The best LD vector is obtained as the dominant eigenvector of $S^{-1}B$.

Algorithm 3 Linear Discriminant Analysis

procedure LDA$\big($ G=$\{(x_i^T, y_i)\}_{i=1}^n\big)$:
specific-class subsets :
$G_i = \{(x_j^T | y_j = c_i, j = 1, .., n\}, i = 1, 2$
class means :
$\beta_i = mean(G_i), i = 1, 2$
scatter in-between matrix class :
$B = (\beta_1 - \beta_2)(\beta_1 - \beta_2)^T$
Centeric matrices class :
$z_i = D_i - 1_n \beta_i^T, i = 1, 2$
scatter Class matrices :
$P_i = z_i^T z_n, i = 1, 2$
class within scatter matrix :
$P = P_1 + P_2, i = 1, 2$
computed eigenvector is obtained :
$\lambda_1, w = eigen(P^{-1}B)$

D. Feed Forward Neural Network

The Multilayer feedforward NN is divided into three layers. The layers are input layer, hidden layers and output layers (Figure 1) [14].

Furthermore, relevant features with lowest error forecast were used as input of feedforward neural networks. In this research, an enhanced multilayer perceptron was used to train the network. The activation function of the model is improved by combining the logistic sigmoid function (equ. 1) and rectify linear unit (equ. 2). Both activations function happened to produced good result compared to the remaining functions. The new activation function $f(x)_{improved}$ is used to improve accuracy of electricity load forecasting.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

$$g(x) = \begin{cases} 0, \text{ if x} < 0 \\ 1, \text{ if } x \geq 0 \end{cases} \tag{2}$$

Combining equation 1 and 2 we have the improved function

$$f(x)_{improved} = 2(f(x) * g(x)) \tag{3}$$

where the constant value 2 is a scaling factor

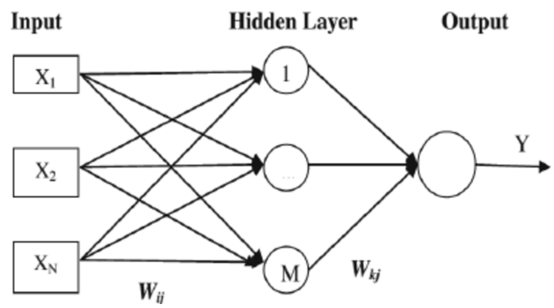

Fig. 1: Feedforward Neural Network [15]

IV. System model

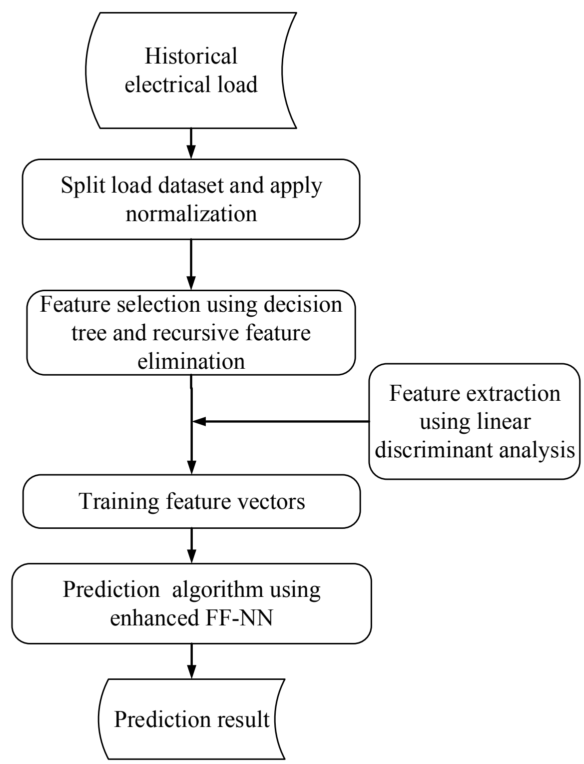

Fig. 2: Proposed System Model

The electricity load demand data set was collected from New Queensland (QLD) Australia [33]. The collected load data are used to verify the authenticity of the proposed model performance. The load data are sampled for every half an hour, which means there are 336 observations per week and 48 observations per day. The data set is covered from 11-10-2018 21:00–14-10-2018 4:00 in QLD. The collected data are used as input variable of the forecast model. After carrying out some preprocessing, predictive analysis is performed with enhanced feedforward neural network method as shown in figure 2.

In the preprocessing step, data is normalized and then divided into three parts: train, test and validation. Following the preprocessing, the data is then given to hybrid features selection. The hybrid features selection is the combination of decision tree and recursive feature elimination, which are used to select the relevant features. The features selected by the hybrid feature selection are assumed to have fewer irrelevant features; however, dimensionality and redundancy are further reduced to improve the forecasting accuracy. Linear discriminant analysis (LDA) is applied to remove the redundancy and reduce dimensionality of the dataset.

Low weighted and less redundant features have been filtered after the processes with feature selector and feature extraction. Finally, the filtered data will be sent to the enhanced feed forward neural network to forecast the electric load.

V. Performance Estimation

In order to measure the forecasting performance of the model, four standard evaluating metrics are used: Root Mean Square Error (RMAE),Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Mean Square Error (MSE). They are defined in equations (1)-(4) [3], [12], [13].

$$MAPE = \frac{1}{w}\sum_{i=1}^{w} |\frac{R_i - P_i}{A_i}| \tag{4}$$

$$MAE = \frac{1}{w}\sum_{i=1}^{w} |R_i - P_i| \tag{5}$$

$$RMSE = \sqrt{\sum_{i=1}^{w} \frac{(R_i - P_i)^2}{R_i}} \tag{6}$$

$$MSE = \sum_{i=1}^{w} \frac{(R_i - P_i)^2}{R_i} \tag{7}$$

where P_i is the forecasted electric load value at point i, R_i is the actual electric load value at point i; w is the total number of data.

VI. Simulations Result

Numerical results are provided in this section to validate the new forecasting model. Figure 3 depicted the normalized loads of QLD for three days (half an hour basis). The normalization graph shows different variations across the hours. In the implementation, data are divided in the ratio 75:25. Where training data has 41.5 hours and testing data has 14 hours. The processed data from feature selection and extraction techniques are then forwarded to the forecasting engine for predictions as shown in figure 5. The data of features selection and dimensionality reduction is given in Table II.

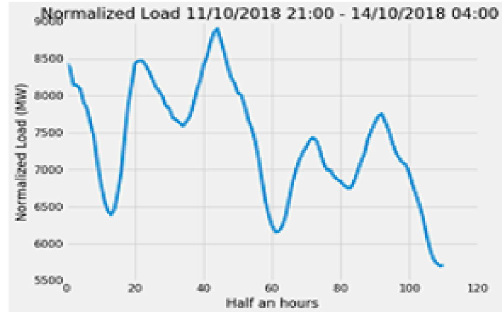

Fig. 3: Normalized load of QLD from 11-10-2018 21:00–14-10-2018 4:00.

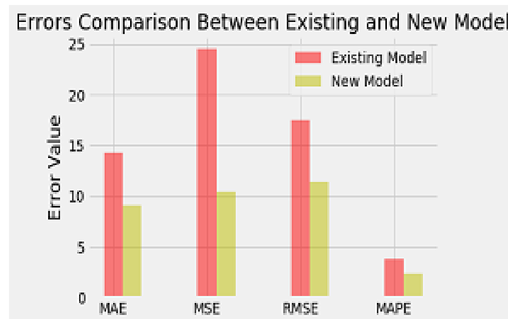

Fig. 4: Comparison of Performances Evaluators.

TABLE II: Attribute of Data

S/N	Features	Status
1	Date	R
2	Net Import (MW)	R
3	Spot Price ($/MWh)	S
4	Scheduled Generation (MW)	S
5	Semi Scheduled Generation (MW)	S
6	Type	R

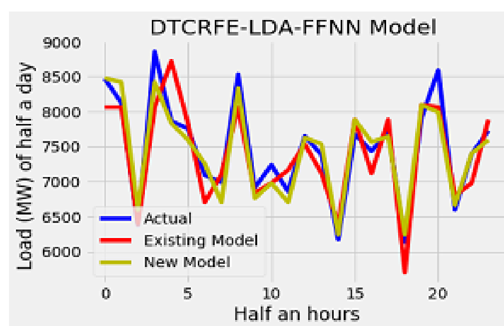

Fig. 5: Comparison of the Forecasted and Actual Load in QLD.

NOTE: The 'S' means that the feature is relevant and should be selected, the 'R' means the feature is irrelevant and should be rejected. The actual and forecasted loads for the enhanced and existing model of the QLD is presented in figure 5. In this figure, blue curves represent the actual load, the red curve is the forecasted load for the existing model while the yellow curve is for the enhanced model. As we can see from figure

5, the performance of the enhanced model shows good result, since the load curve of the enhanced model is more closely to the actual value as compared to the existing model. Error analysis performance result can be obtained from table III. Figure 4 shows the comparison within performance evaluators. Table III records four forecasting errors about the proposed and existing model in the experiment. As shown in table III, the result of performance apparently shows that the proposed model is reasonably good because it has the lowest error in MAPE, MAE, MAE, RMSE and MSE. For example, Value of MAPE is 2.366%, which is apparently smaller than that of the existing model. This indicates that the new model outperforms its counterpart model.

TABLE III: Errors evaluation metrics for the QLD market for the days 11-10-2018 21:00–14-10-2018 4:00

Method	MSE	MAE	RMSE	MAPE
Existing Model	24.505	14.289	17.502	3.745
New Model	10.455	9.05	11.432	2.366

VII. CONCLUSION

This research proposed a modified electricity load forecasting model by improving the activation function of the existing model. The corresponding results of the experiment show that the new model outperforms the existing one with 1.397%. Feature selection and extraction are used to reduce dimensionality and redundancy in order to give a more relevant features to predictor.

REFERENCES

[1] Neupane, Bijay, Wei Lee Woon, and Zeyar Aung. "Ensemble prediction model with expert selection for electricity price forecasting." Energies 10.1 (2017): 77.

[2] Ma, Zhanyu, et al. "The role of data analysis in the development of intelligent energy networks." IEEE Network31.5 (2017): 88-95.

[3] Yang, Ailing, Weide Li, and Xuan Yang. "Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines." Knowledge-Based Systems (2018).

[4] Dong, Yongquan, Zichen Zhang, and Wei-Chiang Hong. "A Hybrid Seasonal Mechanism with a Chaotic Cuckoo Search Algorithm with a Support Vector Regression Model for Electric Load Forecasting." Energies 11.4 (2018): 1009.

[5] X. Zhang, J. Wang, K. Zhang, et al., Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm, Electr. Power Syst. Res. (2017) 270–285.

[6] Moon, Seokjae, and Jang-Won Lee. "Multi-residential demand response scheduling with multi-class appliances in smart grid." IEEE transactions on smart grid 9.4 (2018): 2518-2528.

[7] Ahmad, Adnan, et al. "An optimized home energy management system with integrated renewable energy and storage resources." Energies 10.4 (2017): 549.

[8] Ahmad, Ashfaq, et al. "An accurate and fast converging short-term load forecasting model for industrial applications in a smart grid." IEEE Transactions on Industrial Informatics 13.5 (2017): 2587-2596.

[9] Moon, Seokjae, and Jang-Won Lee. "Multi-residential demand response scheduling with multi-class appliances in smart grid." IEEE transactions on smart grid 9.4 (2018): 2518-2528.

[10] Javaid, Nadeem, et al. "A new heuristically optimized Home Energy Management controller for smart grid." Sustainable Cities and Society 34 (2017): 211-227.

[11] Jiang, Huaiguang, et al. "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization." IEEE Transactions on Smart Grid 9.4 (2018): 3341-3350.

[12] Jiang, Ping, Feng Liu, and Yiliao Song. "A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting." Energy 119 (2017): 694-709.

[13] Din, Ghulam Mohi Ud, and Angelos K. Marnerides. "Short term power load forecasting using deep neural networks." Computing, Networking and Communications (ICNC), 2017 International Conference on. IEEE, 2017.

[14] Raza, Muhammad Qamar, et al. "An intelligent hybrid short-term load forecasting model for smart power grids." Sustainable Cities and Society 31 (2017): 264-275.

[15] Yılmaz, Ali Can, Cigdem Inan Aci, and Kadir Aydin. "MFFNN and GRNN Models for Prediction of Energy Equivalent Speed Values of Involvements in Traffic Accidents/ Trafik Kazalarında tutulumunun Enerji Esdger Hız Degerleri Tahmininde MFFNN ve GRNN Modelleri." International Journal of Automotive Engineering and Technologies 4.2 (2015): 102-109.

[16] Rafiei, Mehdi, Taher Niknam, and Mohammad-Hassan Khooban. "Probabilistic forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network." IEEE Transactions on Industrial Informatics 13.1 (2017): 71-79.

[17] Singh, Nitin, Soumya Ranjan Mohanty, and Rishabh Dev Shukla. "Short term electricity price forecast based on environmentally adapted generalized neuron." Energy 125 (2017): 127-139.

[18] Yu, Chun-Nam, Piotr Mirowski, and Tin Kam Ho. "A sparse coding approach to household electricity demand forecasting in smart grids." IEEE Transactions on Smart Grid 8.2 (2017): 738-748.

[19] W. Li, X. Yang, H. Li, et al., Hybrid forecasting approach based on GRNN neural network and SVR machine for electricity demand forecasting, Energies 10 (1) (2017) 1–17.

[20] Mirjalili, Seyedali. "Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems." Neural Computing and Applications 27.4 (2016): 1053-1073.

[21] Saremi, Shahrzad, Seyedali Mirjalili, and Andrew Lewis. "Grasshopper optimisation algorithm: theory and application." Advances in Engineering Software 105 (2017): 30-47.

[22] Zeng, Nianyin, et al. "A switching delayed PSO optimized extreme learning machine for short-term load forecasting." Neurocomputing 240 (2017): 175-182.

[23] Xiao, Liye, et al. "Research and application of a combined model based on multi-objective optimization for electrical load forecasting." Energy 119 (2017): 1057-1074.

[24] Jiang, Yu, et al. "Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm." Journal of Modern Power Systems and Clean Energy 5.1 (2017): 126-133.

[25] Yang, Zhang, Li Ce, and Li Lian. "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods." Applied Energy 190 (2017): 291-305.

[26] Wang, Long, Zijun Zhang, and Jieqiu Chen. "Short-term electricity price forecasting with stacked denoising autoencoders." IEEE Transactions on Power Systems 32.4 (2017): 2673-2681.

[27] Lago, Jesus, Fjo De Ridder, and Bart De Schutter. "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms." Applied Energy 221 (2018): 386-405.

[28] Abedinia, Oveis, Nima Amjady, and Hamidreza Zareipour. "A new feature selection technique for load and price forecast of electrical power systems." IEEE Transactions on Power Systems 32.1 (2017): 62-74.

[29] Koprinska, Irena, Mashud Rana, and Vassilios G. Agelidis. "Correlation and instance based feature selection for electricity load forecasting." Knowledge-Based Systems 82 (2015): 29-40.

[30] Mafarja, Majdi, et al. "Evolutionary population dynamics and grasshopper optimization approaches for feature selection problems." Knowledge-Based Systems 145 (2018): 25-45.

[31] Zheng, Huiting, Jiabin Yuan, and Long Chen. "Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation." Energies 10.8 (2017): 1168.

[32] Wang, Kun, et al. "Robust big data analytics for electricity price forecasting in the smart grid." IEEE Transactions on Big Data (2017): 1-12.

[33] https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Data-dashboard. accessed date: 10/11/2018.