Signal Processing and Renewable Energy
December 2023, (pp. 69-86)
ISSN: 2588-7327
eISSN: 2588-7335

A Wavelet Transform-Based Hybrid Short Term Load Forecasting Method for Managing the Costs of EV Charging Stations and Parking Lots

Yashar Khanchoopani Ahranjani [1], Mojtaba Beiraghi [1*], Reza Ghanizadeh [1]

[1] Department of electrical engineering, Islamic Azad University, Urmia Branch, Urmia, Iran.

Received: 20-Jul-2023, Revised: 18-Oct-2023, Accepted: 18-Nov-2023.

Abstract

Advanced data processing methods are emerging on a daily basis and they can be used in many tasks. Short-term load forecasting (STLF) is an essential task for power distribution systems. Specially, for the electric vehicle (EV) charging stations and parking lots the STLF can give precise time tables in which the electrical power demand is at the lowest and the electric prices are low too. In this paper, we developed a STLF system that works based on wavelet transform of the load data collected from the previous 3 months. Next, we perform the feature selection using the Gram-Schmidt (GS) procedure. The selected features are then fed to the Radial basis function (RBF) network in order to predict the power usage one day ahead. The learning algorithm that is used for the RBF network is the hybrid "k-means, RLS" algorithm. It is indicated that the proposed RBF network works more accurately or at least equal to the previously presented support vector machine (SVM) predictor. Also, the computational complexity of the RBF network is much less than SVM and consequently, the time consumption of the presented system is far less than recently proposed methods. The numerical simulation results based on the real-world load data of the Urmia city showed that the proposed method based on the wavelet transform features and the hybrid RBF is more efficient than the previous SVM both in accuracy (more than 29%) and computational complexity (more than 30%).

Keywords: RBF Network, Short-Term Load Forecasting, RLS Algorithm, Hybrid Learning, Performance.

1. INTRODUCTION

In order to perform peak shaving for the load profile of the distribution systems, we must have the exact knowledge of the whereabouts

***Corresponding Authors Email:** mojtaba.beiraghi@iaurmia.ac.ir

of the peaks based on load forecasting. In recent years, there has been a growing interest in the application of hybrid intelligent methods for short-term load forecasting (STLF). The synergy of multiple techniques aims to capitalize on the strengths of individual models, offering a more robust and accurate forecasting framework. Notably, Mostafapour et al. [1] proposed a hybrid K-means and VSS-LMS learning method for Radial Basis Function (RBF) networks, demonstrating its efficacy in short-term load forecasting during the 9th IEEE International Conference on Electrical and Electronics Engineering in 2015. Building upon this approach, Abdoos et al. [2] further explored a hybrid intelligent method for short-term load forecasting, contributing to the knowledge base in the field of Knowledge-based systems in 2015. Fan and Chen [3] introduced an adaptive hybrid method based on IEEE Transactions on Power Systems in 2006, emphasizing adaptability as a key factor in improving forecasting accuracy. Moreover, advancements in distributed computing and information integration have played a role in refining forecasting models. Liu et al. [4] presented a distributed short-term load forecasting method leveraging local weather information, underscoring the importance of environmental factors in load prediction. The integration of neural networks, especially recurrent neural networks (RNNs), has also been a focus of recent research. Jiao et al. [5] introduced a Multiple Sequences LSTM Recurrent Neural Network for non-residential load forecasting, showcasing the potential of neural networks in capturing complex temporal dependencies. Furthermore, studies such as that by Amjady and Keynia [6] have explored innovative neural network approaches to short-term load forecasting, contributing to the diversification of methodologies in the field. In addition to hybrid methods, researchers have explored the application of traditional models like Radial Basis Function (RBF) networks, as seen in the work of Dongxiao et al. [7]. Additionally, Liu et al. [8] proposed a hybrid model combining FCW-EMD and KF-BA-SVM for short-term load forecasting, demonstrating the versatility of hybrid models in capturing diverse features for accurate predictions. Traditional statistical models, such as ARIMA, have not been overlooked. Lee and Ko [9] utilized the lifting scheme in conjunction with ARIMA models for short-term load forecasting in 2011, demonstrating the continued relevance of traditional methods. The exploration of innovative forecasting models has extended to the development of hybrid annual power load forecasting models, as evidenced by Li et al. [10], who proposed a hybrid model based on a generalized regression neural network with a fruit fly optimization algorithm. Finally, Demirorena and Ceylan [11] focused on regional load forecasting, using artificial neural networks for forecasting in the Middle Anatolian region, further diversifying the geographical applications of load forecasting models. Lopez et al. [12] applied Self-Organizing Map (SOM) neural networks to short-term load forecasting, presenting a case study of the Spanish electricity market. The study explored the adaptability and efficacy of SOM in capturing complex patterns inherent in electricity consumption. Chen et al. [13]

delved into the application of Support Vector Machines (SVMs) for load forecasting, specifically examining the EUNITE competition in 2001. Their work shed light on the potential of SVMs in handling complex load patterns and their competitive performance in a benchmark competition. Ko and Lee [14] proposed a Short-Term Load Forecasting model using a Support Vector Regression (SVR)-based radial basis function neural network with dual extended Kalman filter. This model aimed to improve accuracy by integrating SVR with Kalman filtering techniques. Wu [15] introduced a novel approach by combining Least Square Support Vector Machines with chaos theory for power load forecasting. This fusion of methodologies aimed to harness the chaotic nature of power consumption patterns. Hybrid models have been a focus of exploration, as seen in Wu's work [16], which integrated Particle Swarm Optimization (PSO) and Wavelet Multiscale Support Vector Machines (Wm-SVM) for power load forecasting. The study highlighted the potential of hybrid models in improving forecasting accuracy. Tomonobu et al. [17] explored the integration of wavelet analysis and neural networks for next-day load curve forecasting. This study showcased the ability of wavelet analysis to enhance feature extraction in load forecasting models. Bashir and El-Hawary [18] applied wavelet to short-term load forecasting using PSO-based neural networks. This approach demonstrated the effectiveness of wavelet analysis in capturing the intricate temporal patterns in power consumption. Chen and Luh [19] introduced a similar day-based wavelet neural network approach for short-term load forecasting, emphasizing the significance of leveraging historical patterns for accurate predictions. Hybrid intelligent algorithms were further explored by Hooshmand et al. [20], who proposed an approach based on a combination of techniques. Their work emphasized the adaptability of hybrid models in handling the dynamic nature of load forecasting. Wu [21] presented a hybrid model combining a wavelet support vector machine and a modified genetic algorithm to penalize Gaussian noises in power load forecasts. This study contributed to the ongoing exploration of hybrid methodologies in load forecasting. Ulagammai et al. [22] applied a bacterial foraging technique trained with artificial and wavelet neural networks in load forecasting. The study showcased the potential of bio-inspired algorithms in enhancing the learning capabilities of neural networks for load prediction. Chauhan and Hanmandlu [23] introduced a novel approach by combining wavelet transformation and fuzzy neural networks for load forecasting. The integration of wavelet analysis and fuzzy logic demonstrated the potential to capture both temporal patterns and uncertainties in load behavior. Hossein et al. [24] delved into the realm of deep neural networks (DNN) for short-term load forecasting, exploring the capabilities of deep learning in handling complex and nonlinear relationships within load data. Their work, presented at the North American Power Symposium in 2017, shed light on the emerging role of deep learning in load prediction. Foundational to the understanding of neural networks is Simon Haykin's work [25] on neural networks and learning machines. His contributions laid the groundwork for the application of neural networks in various domains, including load forecasting. Data sources play a crucial role in training and validating forecasting models. The study by Zhang [28] leveraged fuzzy clustering and least squares support vector machines optimized by a wolf pack algorithm for short-term load forecasting in electric bus charging stations. The incorporation of fuzzy clustering showcased the adaptability of models to diverse operational scenarios. The integration of weather data into load forecasting models was explored by obtaining data from www.weather.com [27]. The interplay between weather patterns and load demand is a critical factor, and incorporating this information enhances the accuracy of predictions. Recent advancements in energy forecasting include Peng et al.'s work [29], which utilized empirical wavelet transform and long short-term memory for effective energy consumption forecasting. The integration of wavelet transform showcased the ability to capture multiscale patterns inherent in energy consumption data. The broader landscape of machine learning applications in energy forecasting is evident in studies such as Pugliese et al. [30], Ahmad et al. [31], Ayub et al. [32], and Irfan et al. [33]. These works explore improved support vector machines, ensemble techniques, and deep learning models for short and medium-term electricity load forecasting, reflecting the ongoing evolution of methodologies to meet the challenges of diverse and dynamic load patterns. Rafi et al. [34] presented a unique method by integrating Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for short-term load forecasting. This integrated approach demonstrated the synergy between spatial and temporal dependencies in load data, providing a comprehensive model for accurate predictions. Therefore, the amalgamation of wavelet transformation, neural networks, deep learning, and machine learning techniques reflects the multifaceted approach to short-term load forecasting. The integration of diverse methodologies aims to enhance the adaptability, accuracy, and robustness of forecasting models in the face of evolving energy landscapes. Based on these methods, in this paper, we proposed the hybrid learning method for the RBF network in STLF using wavelet transform extracted features. Our proposed hybrid algorithm consists of the k-means algorithm for the hidden layer and the adaptive Recursive Least Square (RLS) algorithm for the output layer of the RBF network. Furthermore, to reduce and select the extracted features with the wavelet transform, we applied the GS algorithm. This algorithm is shown to be less computationally complex in comparison with the equally powerful feature reduction methods. For the prediction part, a similar hybrid RBF network has been used for load forecasting in [1]. However, the used adaptive algorithm in the final layer of that RBF network was the variable step-size least mean square (VSS-LMS) algorithm which is the recursive least square (RLS) in this paper. Also, the data that is used in [1] is based on the time features of the load data, while in this paper we used the wavelet-based features of the load data in order to improve our forecasting results. The data that is used in this paper is fed to the wavelet transform and then the frequency features are used for the learning process of the Hybrid RBF-RLS network [25]. As mentioned, in [28] a method based on SVM networks is used for finding the best charging time for the electric buses, and based on the findings in [28] we proposed a method for reducing the costs of electric vehicle (EV) charging stations and parking lots.

Our contributions to this work are as follows:

1) The addition of the feature extraction part using the Wavelet transform and the feature reduction method using the Gram-Schmidt (GS) to the previous work in [1].

2) Improving the hybrid learning by adding the recursive least squares (RLS) algorithm to it.

3) Improving and comparing our load forecasting results with respect to the recently presented forecasting methods.

4) Reducing the computational complexity and processing time of the STLF.

5) Finding the suitable times for charging and discharging the EVs for reducing the charging stations and parking lots.

The rest of this paper is designed as follows:In section 2, we introduced the signal processing tools that are needed for implementing and comparing the STLF results. In section 3, we presented the prevalently used STLF networks. In section 4, we explained the overall STLF method in this paper and then presented the simulation results and comparisons. Finally, in section 5, we presented a brief concluding remarks to our work.

2. THE LOAD FORECASTING TOOLS

Load forecasting is performed using the data from the past. This data consists of several parameters that each can affect the load amount directly. First of all, the future load amount is forecasted based on the load amount of the recent past. For example, we acquire the load amount of three previous months and forecast the load amount of the next day or week. However, several factors might change the load usage amount

Table 1. ***List of the used abbreviations*.**

Abbreviations	Descriptions
RBF	Radial basis function
SVM	Support vector machine
RLS	Recursive least squares
GS	Gram-Schmidt
STLF	Short-time load forecasting
WT	wavelet transform
DNN	Deep neural network
VSS-LMS	Variable step-size LMS
MAPE	Mean absolute percentage error

drastically. Weather is the most important part of these factors. A cloudy sky will most certainly cause the lights on the streets and buildings to be turned on sooner than usual. Economic factors like various electricity prices for different times of the day may cause the usage reduction or increase. Holidays also make a large change in the electric load demand. There are other random disturbances affecting the demanded load level that we do not take into consideration in his paper [2]. Most of the data and signal processing tasks include the feature extraction, feature selection, and classifier or the predictor parts based on these features. Here, we consider the load feature extraction and selection tools.

2.1. The Discrete Wavelet Transform (DWT) for Feature Extraction

The discrete wavelet transform (DWT) can be used for extracting the features of various signals from biological to economic, environmental, and electrical load measurements. The outputs of this transform are the low-frequency approximation components and the detailed high-frequency signal components [2]. The electric load can be thought of as the sum of the lines of different frequency components. Each component can be represented by one or more frequencies. The load curve is divided into several curves with different frequencies by the wavelet conversion, which paves the way for predicting total and partial load. DWT is more suitable for processing and analyzing time series than continuous wavelet converters because DWT data have no additional components and therefore reverse conversion can be used for any set of time-frequency data. Here, we briefly explain the DWT:

Various wavelet functions can be used for this transformation. If we name the wavelet function as $\phi(x)$, then its Fourier transform $\Psi(\omega)$ must satisfy the following inequality:

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|}{|\omega^2|} d\Psi < +\infty \qquad (1)$$

There are several functions that satisfy this condition. The above condition is known as admissibility for $\psi(x)$ that can be written in duality format as follows:

$$\Psi(0) = \int_{-\infty}^{+\infty} \psi(x) dx = 0 \qquad (2)$$

There are several functions that satisfy this condition: Haar, Dabeches, Coiflet, Symlet, and Mexican hat. Also,

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-a}{b}\right) \qquad (3)$$

The wavelet coefficients can then be expressed as:

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi\left(\frac{x-a}{b}\right) dx = \int_{-\infty}^{+\infty} f(x) \psi_{a,b}(x) dx \qquad (4)$$

After extracting features from the load data using the DWT, we must reduce them in order to reduce the computational complexity of the load forecasting process. Also, feature selection is needed to improve the forecasting quality based on the most suitable features. Here, we explain one of the feature selection methods that best suit our forecasting purposes:

2.2. Feature Selection Using the Gram-Schmidt Algorithm

The Gram-Schmidt (GS) feature selection algorithm was proposed in [2] for the reduction of the wavelet-extracted features of the load data. We start with N samples $x(1)$, $x(2)$,…,$x(N)$. Each sample is a vector $x(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T$. We have [2]:

$$X = [x_1, \dots, x_n] = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(N) & x_2(N) & \dots & x_n(N) \end{bmatrix},$$

$$x_i = [x_i(1), x_i(2), \dots x_i(N)]^T \tag{5}$$

We decompose the feature matrix as:

$$X = QR \tag{6}$$

In this relation, Q is an orthogonal matrix and R is an upper triangular matrix. Also:

$$R = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{nm} \end{bmatrix},$$

$$Q = [q_1, q_2, \dots, q_n] = \begin{bmatrix} q_1(1) & q_2(1) & \dots & q_n(1) \\ q_1(2) & q_2(2) & \dots & q_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ q_1(N) & q_2(N) & \dots & q_n(N) \end{bmatrix} \tag{7}$$

In this relation q_i is the new feature vector. The Q matrix consists of:

$$q_1 = x_1 \tag{8}$$

$$q_i = x_i - \sum_{j=1}^{i-1} \alpha_{ji} q_j \tag{9}$$

where

$$\alpha_{ji} = \begin{cases} \frac{q_j^T x_i}{q_j^T q_i} & for\ j = 1,2,, \dots, i-1 \\ 1 & for\ j = i \end{cases} \tag{10}$$

The GS feature selection method maps the input features from space X to space Q. The feature vector q_i can be expressed as sample allocation in the direction of the feature q_k in the orthogonal space. The rank of these features can be determined based on their value in the class separability. In the context of load forecasting, this ability translates to the selection of the features that are most capable of explaining the load pattern. The Gram-Schmidt method can be used for feature selection based on orthogonalization as mentioned [2]. The Gram-Schmidt method ensures that the resulting features are orthogonal, simplifying interpretations and potentially aiding in avoiding multicollinearity issues. It is conceptually straightforward and can be relatively simple to implement for a small number of features. Also, Gram-Schmidt provides direct control over the dimensionality of the resulting feature set, allowing for flexibility in choosing the number of features. In comparison with the principal component analysis (PCA) method, while PCA is a powerful and widely-used technique for dimensionality reduction, the Gram-Schmidt method offers explicit orthogonality and interpretability, making it a viable choice in situations where these properties are crucial. The benefits of Gram-Schmidt lie in its simplicity, explicit control over orthogonality, and potential for a more interpretable feature set.

3. THE PROPOSED FORECASTING NETWORKS

As expressed in the introduction part, several networks have been used for the implementation of load forecasting. The most prevalently used ones are the SVM and RBF networks [8]. Here we explain both of these networks briefly for the sake of comparison. In reference [2] as in many other references, the proposed STLF network is SVM. Also, in [25] the hybrid RBF network is directly compared to the SVM network.

3.1. SVM

The SVM network is recommended for many applications such as machine learning, regression analysis, and prediction. As in the load forecasting our problem can be categorized as a regression analysis one, we introduce the training sets as the pairs of the input and outputs (x_i, y_i). The goal is to estimate the relation of the inputs and the outputs defined as (11) with the lowest error level.

$$f(x,w) = w^T x + b \tag{11}$$

The error value for this case can be measured using the following equation:

$$E(x,y,f) = |y - f(x,w)|_\varepsilon = \begin{cases} 0 & if\ |y - f(x,w)| \le \varepsilon \\ |y - f(x,w)| - \varepsilon & otherwise \end{cases} \tag{12}$$
$$e(x,y,f) = max(0, |y - f(x,w)| - \varepsilon)$$

The total error value in this case is then given using the empirical risk value:

$$R_{emp}^{\varepsilon}(w,b) = \frac{1}{l}\sum_{i=1}^{l} |y_i - w^T x_i - b|_\varepsilon \tag{13}$$

Our goal is to minimize the following objective function:

$$R = \frac{1}{2}(|w|)^2 + C\sum_{i=1}^{l} |y_i - f(x_i,w)|_\varepsilon \tag{14}$$

$$R_{w,\xi,\xi^*} \tag{15}$$
$$= \left[\frac{1}{2}\|w\|^2 + C\left(\sum_{i=1}^{l} \xi_i + \sum_{i=1}^{l} \xi_i^*\right)\right]$$

Our problem constraints are:

$$y_i - w^T X_i - b \le \varepsilon + \xi_i, \quad i = 1, \dots, l \tag{16}$$

$$w^T x_i + b - y_i \le \varepsilon + \xi_i^*, \quad i = 1, \dots, l \tag{17}$$

$$\xi_i \ge 0, \qquad \xi_i^* \ge 0 \qquad i = 1, \dots, l \tag{18}$$

where the parameters ξ_i and ξ_i^* are the slack variables. There are two main paths for the STLF. One is the usage of feature extraction and selection with non-deep neural networks. The other path is to use deep learning that does not need feature extraction as it is done in the network itself. We compared our results mainly with the SVM and RBF networks that are not deep learning. Therefore, with explaining SVM we wanted to emphasize the importance of this comparison.

3.2. The Hybrid RBF Neural Network

With recent advances in artificial intelligence, new techniques for prediction have been proposed that are more accurate than traditional techniques. The most common of these techniques are neural network algorithms. The main capability of the neural networks that have led to their use in load forecasting is the high ability of these networks to detect patterns and learn the relationship between inputs and outputs. We have various types of neural networks, and different methods have been used to train these networks to predict time series with varying accuracy. On the other hand, the experience of recent studies shows that hybrid models have a higher ability to predict than to use only one predictive model. Our improved STLF method is based on the previous work of [1] and the results will confirm the correctness of our approach. Neural networks can be used to model different functions and predict their future values. The most important task in building a neural network for load prediction is to select input variables. There is no specific general rule for this purpose. This is more dependent on engineering judgment and experience, and in almost all cases is accompanied by trial and error. However, some statistical studies can be effective in determining variables that have a greater impact on system load. The main variables used in short-term forecasting of electric charge are the day of the week, hour, temperature, and recorded measurements of the load over the past few days. To perform STLF, we used the hybrid radial basis functions (RBF) networks that use certain basis functions to do this task. We assumed that the basis functions of the RBF network are Gaussian, however, other types of basis functions are available for the RBF networks. These functions are implemented in the hidden layer of the network and they are used to model regression functions. The learning algorithms of the RBF network can be improved separately. For the hidden layer, we used the improved k-means algorithm. The output layer of this network operates based on the weights that are identified using the adaptive algorithms. In this paper, for the output layer we used the powerful adaptive RLS adaptive algorithm.

As we mentioned the Gaussian function is used as the radial basis functions of the RBF network. In (1) we have this function:

$$\phi(x_i) = exp\left[-\frac{1}{2\sigma_i^2}\|x_i - t_i\|\right] \qquad (19)$$

here x_is are showing the training samples, t_is show the central points of the radial functions. Also, the variance of these

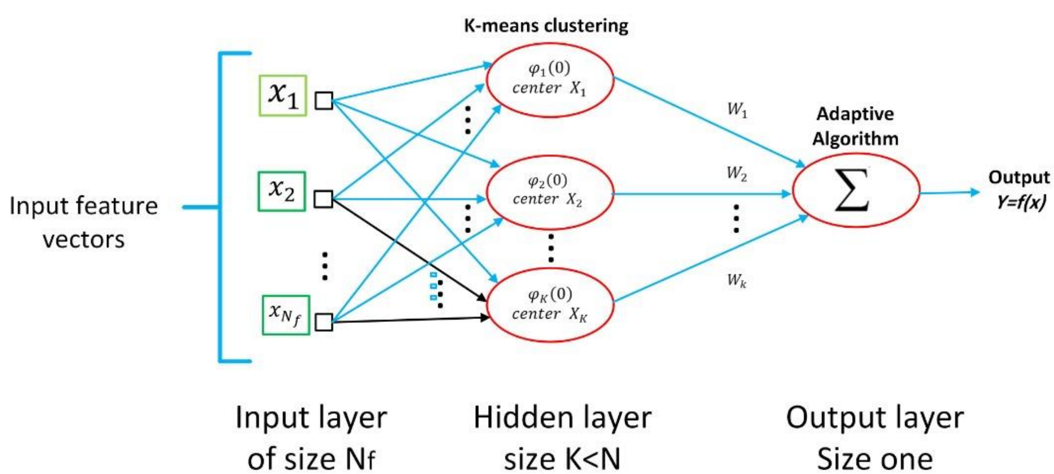

Fig. 1. The RBF network and its learning algorithms.

functions is assumed to be σ_i. Using these functions, the RBF network performs a non-linear mapping from the inputs to the hidden layer. Finding the proper centers and variances for them is the task of the K-means algorithm. Next, we have the mapping from the hidden layer to the output layer which is a linear one. We have:

$$Y = \sum_{i=1}^{m} w_i \phi(x_i) \tag{20}$$

The weights in this relation (w_i) are acquired with the usage of the adaptive algorithms which in this paper is the recursive least squares (RLS) algorithm.

The typical RBF network learning begins by choosing the suitable radial function (which is Gaussian here) and finding its variance by the following relation:

$$\sigma = \sqrt{\frac{d_{max}}{m+1}} \tag{21}$$

where d_{max} is the highest distance among the inputs and (m) is the number of neurons in the hidden layer. Up until now, many methods have been proposed for training the hidden layer of the RBF network including self-organizing map algorithms, randomized algorithms, and the famous K-means algorithm [25]. Self-organizing and randomized algorithms are appropriate only for static and dynamic training. Here, we explain the K-means and RLS algorithms briefly:

3.2.1. K-Means Algorithm

As we mentioned, the K-means algorithm is responsible for the non-linear mapping of the RBF network and does this by pointing out the centers of the basis functions. The K-means algorithm is a clustering procedure that divides the input data into k clusters according to the value of k. The method performs the iterative updates: First, in each level, it divides the data points in the vicinity into k clusters based on the k value, and the center of each cluster (the average of all points in this cluster, namely, the geometric Center) will be utilized as the new reference point for the next level of iteration. Iteration makes the selected reference point get closer and closer to the true cluster centroid so that clustering results improve in time. The stages of this method are as follows:

1) Choose a random k value as the initial cluster number for n data points;
2) Measure the distances between each data point and each cluster center, and put the input data point into the closest cluster;
3) Re-calculate the centers of k clusters following the assignment of all the data points to the clusters;
4) Match the results with the previous k-cluster centers, if the new cluster centers differ, turn to 2), Otherwise turn to 5);
5) the clustering results are authenticated.

3.2.2. The Hybrid "K-Means, RLS" Learning

The aim of adaptive algorithms is to iteratively converge to the desired weights. The weights of the output layer in the RBF network are tuned in the training step. The RLS algorithm is one of the most prevalent used and promising adaptive algorithms in context of the adaptive signal processing. Here, we describe the objective of the RLS algorithm in the output layer of the RBF network [8]. Let the $K \times 1$ vector:

$$\boldsymbol{\Phi}(\boldsymbol{x}_i) = \begin{bmatrix} \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_1) \\ \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_2) \\ \vdots \\ \varphi(\boldsymbol{x}_i, \boldsymbol{\mu}_K) \end{bmatrix} \quad (22)$$

Denote the outlets of the K neurons in the hidden layer (Radial functions). This vector is produced in response to the inputs $\boldsymbol{x}_i, i = 1,2,\dots,N$. For the supervised training of the output layer, we present the training sample by $\{\boldsymbol{\Phi}(i), d(i)\}_{i=1}^{N}$, where d_i is the desired output of the RBF network for input $\boldsymbol{x}_i$.This output is achieved by the utilization of the RLS algorithm as follows [25]:

By assuming that the training sample is shown with $\{\boldsymbol{\Phi}(i), d(i)\}_{i=1}^{N}$, perform the next computations for iterations $n = 1,2,\dots,N$:

$$\boldsymbol{P}(n) = \boldsymbol{P}(n-1) - \frac{\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^{\boldsymbol{T}}(n)\boldsymbol{P}(n-1)}{1+\boldsymbol{\Phi}^{\boldsymbol{T}}(n)\boldsymbol{P}(n-1)\boldsymbol{\Phi}(n)} \quad (23)$$

$$\boldsymbol{g}(n) = \boldsymbol{P}(n)\boldsymbol{\Phi}(n) \quad (24)$$

$$\alpha(n) = d(n) - \widehat{\boldsymbol{w}}^{T}(n-1)\boldsymbol{\Phi}(n) \quad (25)$$

$$\widehat{\boldsymbol{w}}(n) = \widehat{\boldsymbol{w}}(n-1) + \boldsymbol{g}(n)\alpha(n) \quad (26)$$

To begin the algorithm, we assume that the weights are $\widehat{\boldsymbol{w}}(0) = \boldsymbol{0}$ and, $\boldsymbol{P}(0) = \lambda^{-1}\boldsymbol{I}$. where λ is the forgetting factor of the RLS algorithm.

In [25] a complete numerical and simulation-based performance analysis has been performed to present the advantages of hybrid RBF to the SVM network, based on accuracy. Furthermore, a minimum of 30 percent reduction in computational complexity along with the time of convergence is achieved using the hybrid RBF in comparison with SVM [25]. In the next part, we presented the STLF simulations using the hybrid RBF network based on the DWT features.

4. NUMERICAL SIMULATIONS

In order to present our simulation results, first we consider the data processing part of our proposed STLF method and then proceed to explain the simulation frame that we used. It is important to mention that the historical load and weather data that are gathered from the websites usually contain bad data.

4.1. Data Preprocessing

The quality of the data is very effective in extracting the desired results and real information. Large databases contain annoying, missing, and unstable data. Therefore, in order to improve the quality of the data, it is necessary to process it at the beginning of the work. Raw data usually has problems such as noise, bias, and drastic changes in dynamic and sampling rate, and using them in this way will undermine subsequent designs. In short, the data processing process includes all the processing that is done on the raw data, which makes it easier and more efficient to process the next data. There are various tools and methods for the pre-processing load data, such as normalization, which converts data into new data with appropriate modification intervals or distributions. Dimension reduction, which is used to delete duplicate, add, or irrelevant data for categorization. These are necessary parts of load forecasting and we applied them on our raw data. The Load and weather data of this paper is available and we can share it upon request. In some parts, this data is pre-processed by data augmentation for the empty parts and none non-compatible data is removed from the data using thresholding.

4.2. Data Processing

The feature extraction and selection parts along with the RBF network are all implemented with MATLAB software version 16a. First, the historical load and weather data of the past months are collected and then the WT transform is applied to them. Next, the GS feature extraction is applied to the extracted features to select 30 prominent ones. The historical load data is collected from the regional power management office and the website in [26]. Also, the weather data can be collected from the weather website [27]. One important aspect of the proposed STLF methods in literature is that they usually neglect the signal processing part of the load forecasting data and mostly apply the raw and unintended load and weather data to their forecasting method.
The RBF network is used for predicting both the detailed and approximation parts of the transformed load data. All the necessary actions regarding the separation of the weekdays and weekends and the impact of holidays in the STLF are explained in [2] and we adopted them in this paper accordingly.

Finally, it should be noted that many parameters can affect the STLF performance and this is why there are various error levels that are reported in the literature. Even the used power levels of the load data can lead to different error levels. Our robust STLF method in this paper is compared directly to the STLF performances of [1] and [7], and our adopted method is an improved and combined version of the algorithms that are used in [1] and [2]. The Overall flowchart of the proposed load forecasting system is given in Fig. 2:

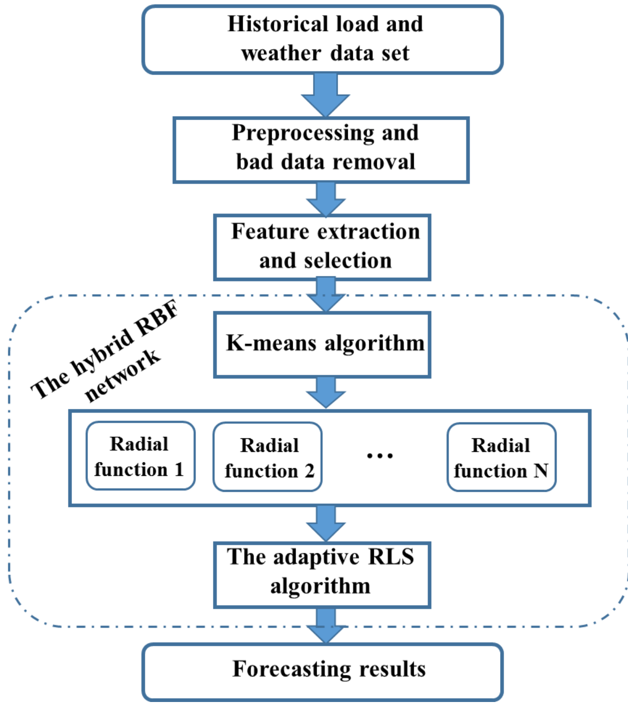

Fig. 2. Flowchart of the proposed load forecasting scheme.

In [2] it is mentioned the best number of features is 30 that are collected from the load data of 50 consecutive days.

4.3. Simulation results

As in [1] for our simulations, we used the hourly load data belonging to the 50 days of Urmia City, Iran, from the 2nd of October to the 21st of November as a training set and performed the prediction for a load of 22nd of November. Urmia city is located in the West Azarbayjan province of Iran and has an average load demand of about 500 MWs. Also, Urmia load demand has a peak of about 750 MWS in the summers. The load demand in this city never goes below 300MWs. Based on the data of Urmia city we want to manage the charging and discharging of EVs in the charging stations and parking lots of this city.

The STLF is done using the hybrid RBF network with a hybrid learning scheme using the "K-means, RLS" algorithm of [1]. Relative accuracy in load forecasting is of particular importance in the electricity industry. On the other hand, the plans for the development of the network and the capacity of the power plants are prepared based on the forecast of the maximum simultaneous load for the planning periods. Therefore, it is considered that load forecasting can be the basis for the development and optimal operation of power plants and power grids to provide reliable electricity and electricity to consumers in the most economical way possible in network conditions. The metric for comparing the performances of STLF methods is MAPE (Mean absolute percentage error) that can be achieved as follows:

$$MAPE = \left(\frac{|x_i - y_i|}{x_i}\right) \times 100 \tag{27}$$

where x_i is the actual load level in time i and y_i is the predicted value. Various references utilized different performance metrics, such as root mean square error (RMSE) or mean absolute error (MAE). However, these metrics are not comparable across different papers, as they are sometimes provided in non-normalized values. Additionally, RMSE and MAE are not consistently reported in all papers. The most prevalent metric used is MAPE, which clearly indicates the network's performance.

For one week ahead load forecasting the best error level achieved is as much as 1.34% [2]. However, the reported STLF error levels for one day ahead load are completely various and it depends on many parameters.

Table 2. ***MAPE performance comparisons of different load forecasting methods.***

Forecasting type	Percentage: MAPE
DenseNet-AO [33]	% 5.67
SOM-SVM [3]	% 3.03
ELM-GA [31]	% 2.58
CNN-LSTM [34]	% 2.19
SVM-GWO [32]	% 1.33
EWT-attention-LSTM [29]	% 1.60
Hybrid forecast engine [6]	% 1.57
RBF [7]	% 0.8184
Hybrid RBF [1]	% 0.511
The Proposed method: Signal processing + Hybrid RBF	% 0.3634

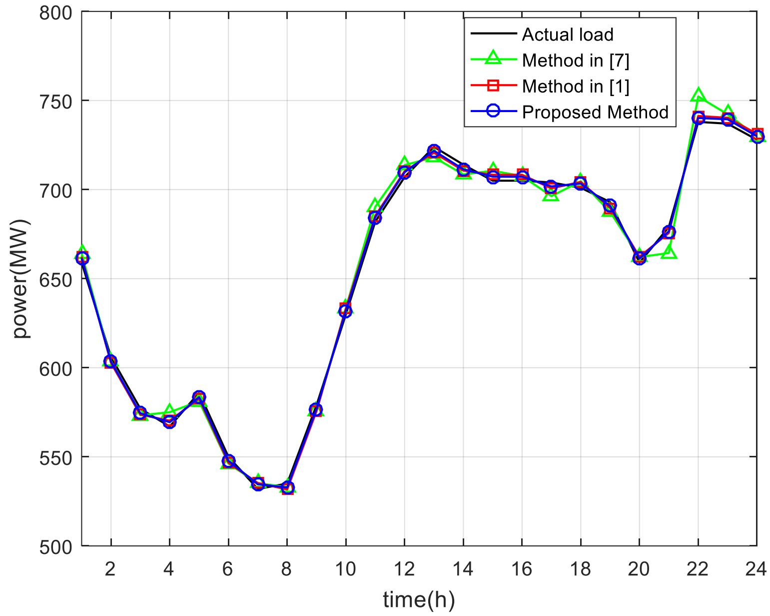

Fig. 3. Comparing the Load forecasting results for various STLF methods.

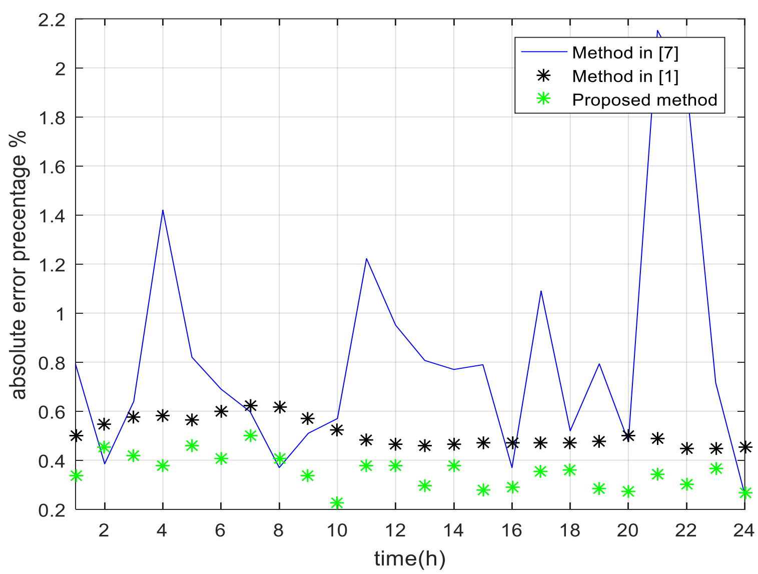

Fig. 4. Comparing STLF results based on the MAPE criteria.

To the best of our knowledge, the best error performance that is reported for this case is in [1] with the MAPE error level of 0.511%. Based on this report, we focused our attention on improving this value using the same load and weather data that are used in [1]. We refined this data using the feature extraction and selection methods and improved the results. The up-to-date methods [33-34] are mainly based on deep learning and as our approach is about data processing, feature extraction, and selection, the topics do not match. Also, the used data sets are different for different papers and the data might not be preprocessed. However, we added the results of newly presented papers to show the superiority of the proposed signal processing method. In Table II, the comparison of the proposed STLF method and the previously presented methods is given:

In Fig. 3 we can see the diagrams of actual and forecasted load using several methods, as we can see, the performance of the proposed method is better than other schemes. Fig. 4 shows a comparison between the performance of the k-means RBF network [7], the hybrid RBF of [1], and our network plus the signal processing. We can see that the mean value of error for our method is lower than that of the k-means RBF network in [7] with raw load data.

With these results, we can determine how closely we can predict the future load values. As we can see, the proposed method presents the closest forecasting results to the actual load values. In order to manage costs for a charging station or parking for electric vehicles, we need accurate knowledge of the maximum and minimum load demand times in the network. In smart power distribution networks, energy costs are highest during peak hours and lowest during minimum usage hours. Therefore, starting to charge cars at the minimum consumption time and discharging cars, if possible, at the peak time and selling load to the network can greatly reduce the costs of electric car charging stations and parking lots. In modern smart networks, the cost of purchasing electric load from customers also varies according to the times of load selling to the network, and the maximum profit for discharge is at maximum consumption times. Increasing the accuracy of load forecasting can help station and parking lot owners precisely adjust the charging and discharging times of cars. This setting should also be based on the time of presence of cars in parking lots and stations.

As can be inferred from the forecasted results in Fig. 3, the minimum load hours of the network in Urmia city are between 6 and 8 in the morning, which is the best time to charge cars, and the maximum load hours are between 22 and 24 PM, which is the best time to discharge cars in the station or parking lot. Next, we present our results based on the MAPE criteria:

As we can see in Fig 4, the lowest MAPE error values are achievable by using the hybrid “K-means, RLS” learning scheme of this paper. Based on these findings we conclude that the added signal processing part of the load forecasting scheme can improve the load forecasting results considerably.

5. CONCLUSION

The aim of investigating the new STLF methods is to propose a precise and easy-to-use strategy. For the EV charging stations, the STLF is vital because of cost management. For this reason, in this paper, we proposed an improved hybrid RBF network with upgraded feature extraction and selection techniques for the STLF. The proposed method is both fast and accurate while it is less computationally demanding. Our proposed STLF system works based on WT of load data collected from the previous 3 months. For reducing the redundant data from input matrices, the Gram–Schmidt GS feature selection is used. The reduction of redundant data can speed up the learning process and improve the generalization capability of the prediction scheme. The collected features are then fed to the RBF network in order to predict the power usage one day ahead. The learning algorithm that is used for the RBF network is the hybrid "k-means, RLS" algorithm. We showed that the proposed RBF network works more accurately or at least equally to the previously presented SVM predictor. Also, the computational complexity of the RBF network is much less than SVM and consequently, the time consumption of the presented system is far less than recently proposed methods. Our main purpose in this paper was to compare the RBF network and SVM in the short-term load forecasting task and to show that RBF is more efficient than SVM. Based on the STLF results, we can conclude that signal-processing methods can greatly improve load forecasting for pre-processed and standard datasets. The proposed method can serve as a path opener for other researchers to incorporate signal processing in various prediction and forecasting applications. In future works, we will work on the deep neural network schemes in order to further amplify the precision of the STLF.

REFERENCES

[1] E. Mostafapour, M. Panahi, M. Farsadi, “A hybrid K-means, VSS-LMS learning method for RBF network in short-term load forecasting,” *IEEE 9th international conference on electrical and electronics engineering (ELECO)*, pp. 961-965, Bursa, Turkey, February, 2015.

[2] A. Abdoos, M. Hemmati, A. A. Abdoos, “Short term load forecasting using a hybrid intelligent method,” *Knowledge-based systems*, vol. 76, pp. 139-147, 2015.

[3] S. Fan, L. Chen, “Short-Term Load Forecasting Based on an Adaptive Hybrid Method,” *IEEE Trans. Power Syst.* Vol. 21, no. 1, *pp.* 392-401, 2006.

[4] D. Liu, L. Zeng, C. Li, K. Ma, Y. Chen, Y. Cao, “A Distributed Short-Term Load Forecasting Method Based on Local Weather Information,” *IEEE Systems Journal*, vol. 12, no.1, pp. 208–215, 2018.

[5] R. Jiao, T. Zhang, Y. Jiang, H. He, “Short-Term Non-residential Load Forecasting based on Multiple Sequences LSTM Recurrent Neural Network,” *IEEE Access*, vol. 6, pp. 59438-59448, 2018.

[6] N. Amjady, F. Keynia, “A New Neural Network Approach to Short Term Load Forecasting of Electrical Power Systems,” *Energies Journal*, vol. 4, pp. 488–503, *March 2011.*

[7] N. Dongxiao J. Ling, T. Jie, “Improved RBF network applied to short-term load forecasting,” IEEE 2nd International Conference on Software Engineering and Service Science, *Beijing, July 2011*.

[8] Q. Liu, Y. Shen, L. Wu, J. Li, L. Zhuang, S. Wang, “A hybrid FCW-EMD and KF-BA-SVM based model for short term load forecasting,” *CSEE journal of power and energy systems*, vol. 4, No. 2, pp. 226-237 June 2018.

[9] C.M. Lee, C.N. Ko, Short-term load forecasting using lifting scheme and ARIMA models, Expert Syst. Appl. 38 (2011) 5902–5911.

[10] H.-Z. Li, S. Guo, C-J. Li, J-Q. Sun, A

hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm, Knowl.-Based Syst. 37 (2013) 378–387.

[11] A. Demirorena, G. Ceylana, Middle Anatolian region short-term load forecasting using artificial neural networks, Electr. Power Compon. Syst. 34 (2006) 707–724.

[12] M. Lopez, S. Valeroa, C. Senabrea, J. Apariciob, A. Gabaldonc, Application of SOM neural networks to short-term load forecasting: the Spanish electricity market case study, Electr. Power Syst. Res. 91 (2012) 18–27.

[13] B.J. Chen, M.W. Chang, C.J. Lin, Load forecasting using support vector machines: a study on EUNITE competition 2001, IEEE Trans. Power Syst. 19 (2004) 1821–1830.

[14] C.N. Ko, C.M. Lee, Short-term load forecasting using SVR (Support Vector Regression)-based radial basis function neural network with dual extended Kalman filter, Energy 49 (2013) 413–422.

[15] H. Wu, X. Chang, Power load forecasting with least square support vector machines and chaos theory, in: Proceeding IEEE PES General Meeting, July 2009, pp. 1–6.

[16] Q. Wu, Power load forecasts based on hybrid PSO with Gaussian and adaptive mutation and Wm-SVM, Expert Syst. Appl. 37 (2010) 194–201.

[17] S. Tomonobu, T. Yoshinori, T. Hitoshi, U. Katsumi, Next day load curve forecasting using wavelet analysis with neural network, Electr. Power Compon. Syst. 30 (2002) 1167–1178.

[18] Z.A. Bashir, M.E. El-Hawary, applying wavelet to short term load forecasting using PSO based neural networks, IEEE Trans. Power Syst. 24 (2009) 20–27.

[19] Y. Chen, P.B. Luh, Short term load forecasting: similar day-based wavelet neural networks, IEEE Trans. Power Syst. 25 (2010) 322–330.

[20] R. Hooshmand, H. Amooshahi, M. Parastegari, A hybrid intelligent algorithm based short-term load forecasting approach, Int. J. Electr. Power Energy Syst. 45 (2013) 313–324.

[21] Q. Wu, Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts, Expert Syst. Appl. 38 (2011) 379–385.

[22] M. Ulagammai, P. Venkatesh, P.S. Kannan, N.P. Padhy, Application of bacterial foraging technique trained artificial and wavelet neural networks in load forecasting, Neurocomputing 70 (2007) 2659–2667.

[23] B.K. Chauhan, M. Hanmandlu, Load forecasting using wavelet fuzzy neural network, Int. J. Knowl.-Based Intell. Eng. Syst. 14 (2010) 57–71.

[24] T. Hossein, S. J. Plathottam, R. K. Angamuthu, P. Ranganathan, "Short-term load forecasting using deep neural networks (DNN)," North American power symposium, 17-19 Sept. 2017.

[25] [Simon Haykin, *Neural networks and learning machines*. Prentice Hall, pp. 230-263, 2008.

[26] Load data is obtained from website:

http://sccis- rep.igmc.ir/RptParam.aspx and the regional Power management office of the West Azarbayjan province.

[27] Weather data is obtained from website: www.weather.com.

[28] X. Zhang, "Short-term load forecasting for electric bus charging stations based on fuzzy clustering and least squares support vector machine optimized by wolf pack algorithm," Energies, 11(6), p.1449, 2018.

[29] Peng, Lu, Lin Wang, De Xia, and Qinglu Gao. "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory." *Energy* 238 (2022): 121756.

[30] Pugliese, Raffaele, Stefano Regondi, and Riccardo Marini. "Machine learning-based approach: Global trends, research directions, and regulatory standpoints." *Data Science and Management* 4 (2021): 19-29.

[31] Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., & Glowacz, A. (2020). Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. Energies, 13(11), 2907.

[32] Ayub, N., Irfan, M., Awais, M., Ali, U., Ali, T., Hamdi, M., ... & Muhammad, F. (2020). Big data analytics for short and medium-term electricity load forecasting using an AI techniques ensembler. Energies, 13(19), 5193.

[33] Irfan, M., Raza, A., Althobiani, F., Ayub, N., Idrees, M., Ali, Z., ... & Alqhtani, S. (2022). Week Ahead Electricity Power and Price Forecasting Using Improved DenseNet-121 Method. Computers, Materials & Continua, 72(3).

[34] Rafi, Shafiul Hasan, Shohana Rahman Deeba, and Eklas Hossain. "A short-term load forecasting method using integrated CNN and LSTM network." *IEEE Access* 9 (2021): 32436-32448.