Logistics & Sustainable Transport
Vol. 11, No. 1, February 2020, 51-76
doi: 10.2478/jlst-2020-0004

Methods and Models for Electric Load Forecasting: A Comprehensive Review

Mahmoud A. HAMMAD[1], Borut JEREB[2], Bojan Rosi[2], Dejan DRAGAN[2*]

[1] Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt
[2] University of Maribor/Faculty of Logistics, Celje, Slovenia

[Corresponding Author indicated by an asterisk *]

***Abstract* — Electric load forecasting (ELF) is a vital process in the planning of the electricity industry and plays a crucial role in electric capacity scheduling and power systems management and, therefore, it has attracted increasing academic interest. Hence, the accuracy of electric load forecasting has great importance for energy generating capacity scheduling and power system management. This paper presents a review of forecasting methods and models for electricity load. About 45 academic papers have been used for the comparison based on specified criteria such as time frame, inputs, outputs, the scale of the project, and value. The review reveals that despite the relative simplicity of all reviewed models, the regression analysis is still widely used and efficient for long-term forecasting. As for short-term predictions, machine learning or artificial intelligence-based models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Fuzzy logic are favored.**

***Key words* — Electric load forecasting; Modeling electricity loads; Methods and models of forecasting.**

I. INTRODUCTION

Electric is a clean and efficient source of energy, which plays an irreplaceable role in our daily life [1]. The significance of electricity has been increasing drastically recently [2], and therefore it has become an essential subject in research [3]. Besides, electric power is more suitable and efficient for the requirement of environment-friendly society compared with other traditional sources of energy such as natural gas, coal, and oil [1]. However, electricity as a product has different characteristics compared to material products since it cannot be stored in bulk as it should be generated as soon as it is demanded [2, 4]. Also, the demand pattern of electricity is complex due to the deregulation of electricity markets such as electricity oversupply and shortage, which could lead to inaccurate forecasting and causes significant financial loss [3, 4]. Moreover, the global electricity demand is expected to multiply with increasing population and living standards improvement. Also, economies expand as well as using high-power electrical appliances and developing technology such as smart grids, electric cars, and renewable energy production. All these factors make it difficult to manage the power system [3, 5, 6]. Hence, it is necessary to predict the needs/loads of electricity in advance and before deciding on the generation of it [7].

Electric load (EL)/demand forecasting is a vital process in the planning of the electricity industry [4] and plays a crucial role in the operation of electric power systems [3]. The electric power load forecast is highly related to the economy's development, and it is also related to national security and the daily operation of society [5]. Therefore, the accuracy of electric load forecasting has great importance for energy generating capacity scheduling and power system management [1], as these accurate forecasts lead to substantial savings in operating and maintenance costs, and correct decisions for future development [4]. Furthermore, electric power load forecasting represents the initial step in developing future generation, transmission, and distribution facilities [8]. However, the accuracy of electric load forecasting (ELF) cannot often fulfill our desired result because it is influenced by various uncertain and uncontrollable factors such as economic development, human social activities, country policies, and climate change [9].

So far, there is no precise standard for classifying the range of load forecasts [3]. However, some authors have divided load forecasting in terms of the prediction duration into three categories [4, 6, 10, 11]: short-term forecasts, medium-term forecasts, and long-term forecasts. Other researchers go for classifying load forecasting into four groups [3, 12-14]: long-term forecasts, mid-term forecasts, short-term forecasts, and very short-term forecasts, as follows:

- ***Long-term load forecasting (LTLF)****:* is for more than one year to 20 years ahead. This type of forecast is fundamental for strategic planning, construction of new generations, and develops the power supply and delivery system (generation units, transmission system, and distribution system).
- ***Medium-term load forecasting (MTLF)****:* is usually for a week up to a year, which is used for maintenance scheduling and planning fuel purchases as well as energy trading and revenue assessment for the utilities.
- ***Short-term load forecasting (STLF)****:* is for intervals ranging from one hour to a week, it is very important for day-to-day operations of a utility, schedule the generation and transmission of electricity.
- ***Ultra/very short-term load forecasting (VSTLF)****:* ranges from a few minutes to an hour ahead and is used for real-time control.

Although numerous forecasting methods and models were developed to compute an accurate load forecasting, finding an appropriate forecasting model for a specific electricity network is not an easy task, and none of them can be generalized for all demand patterns [4, 9, 15]. Based on [1], electric load forecasting models can be divided into two types:

- ***multi-factor forecasting methods***, and
- ***time series forecasting methods***.

The ***multi-factor/cross-sectional forecasting method*** focuses on the search of the causal relationships between different influencing factors and forecasting values. On the other hand, ***the time series forecasting method*** is depending more on the historical series. Accordingly, lots of researchers turn to utilize the time series forecasting method to forecast electric load to avoid the complicated and non-objective factors that might effect on establishing an accurate forecasting model using a multi-factor forecasting method. Thus, the time series forecasting method is much easier and quicker. The ***most frequently and widely used time series forecasting models*** can be divided into three subcategories [1]:

statistical models,

machine learning models, and

hybrid models.

The main contribution of this paper is believed to be enriching the existing literature about electric load forecasting, with the main emphasis to complement the explanation given in already published review papers from the field. In this context, there have been three most important papers detected about a comprehensive review regarding the methods, models, and different methodologies about the electric load forecasting [16-18]. Four review papers discuss about the computational intelligence, data mining, and machine learning approaches in the ELF field [19-22], while the two papers address a long-term ELF comprehensively [23, 24]. One paper is entirely dedicated to the short-term ELF only [25], while the work of Weron [26] focuses its attention to the statistical approaches for modeling and forecasting electricity loads and prices. Finally, the review and analysis of machine learning and regression models related to commercial buildings are presented in the paper of Yildiz and his colleagues [27].

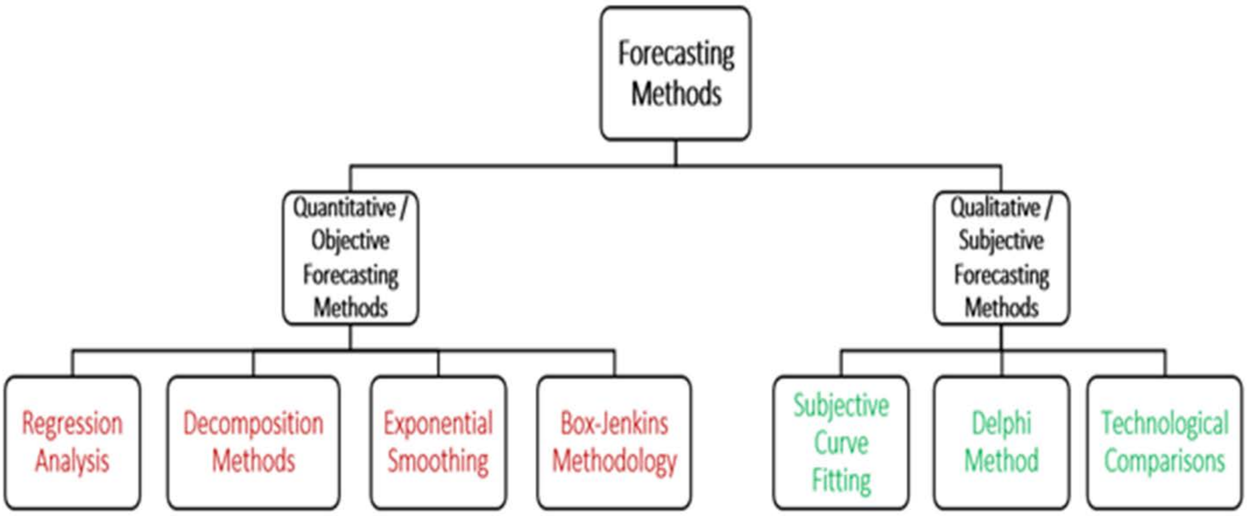
The rest of the paper is organized as follows. In the next section, the most frequently used methods and models of load forecasting in the literature are briefly explained. Section III applies the methodology used in this work. In Section IV, the research findings and results are discussed and analyzed. Finally, in Section V, the conclusions of the paper are given.

II. THE LITERATURE REVIEW

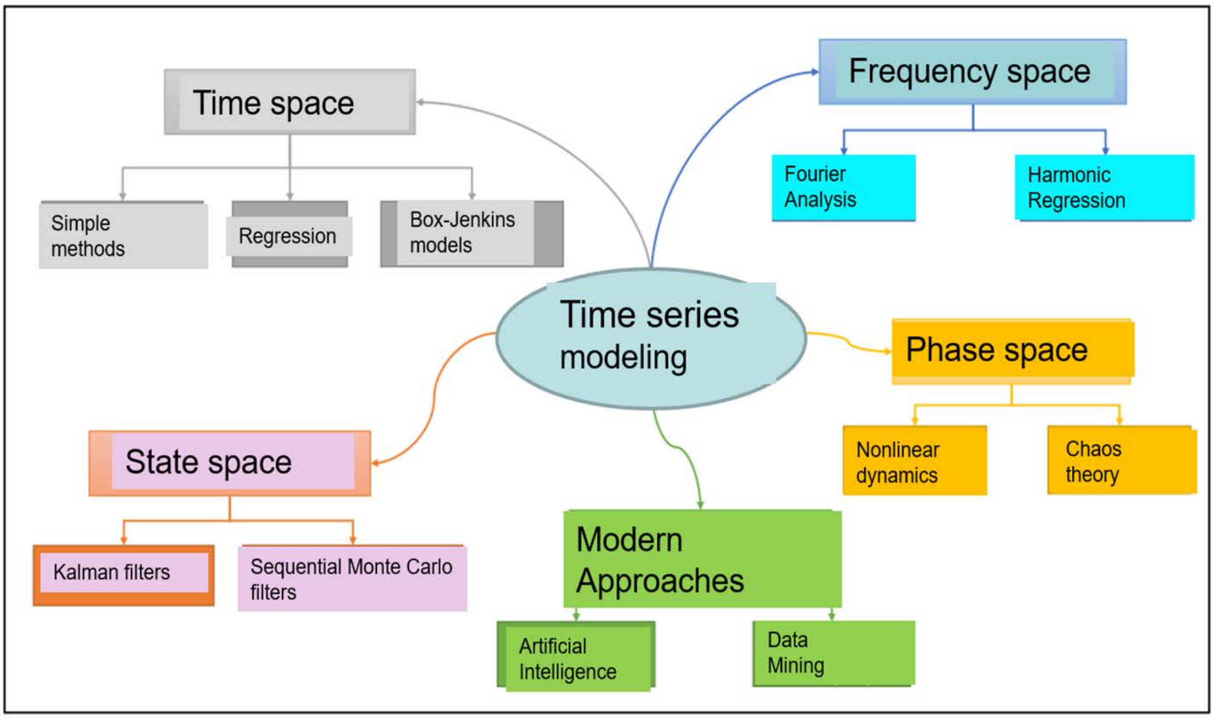
In this section, the most widely used methods and models in the area of electric load forecasting are briefly discussed by reviewing the relevant previous works. According to Feinberg [28], the majority of forecasting methods for the ELF forecasting are related with the artificial intelligence algorithms and statistical approaches. In these two spheres, the regression models, fuzzy logic, neural networks, and expert systems are particularly important. Moreover, for the medium-term and long-term forecasting, an econometric approach and so-called "end-use" approach are also the prevailing ones. On the other side, for the short-term forecasting, the following approaches are significant: neural networks, various time series and regression models, statistical learning approaches, so-called "similar day approach", fuzzy logic models, and expert systems.

A. Basic Forecasting/Prediction Methods

According to [13, 29, 30], the basic forecasting methods are classified into two basic types: ***qualitative and quantitative methods***, and selecting the appropriate type depends mainly on the data available. In qualitative/subjective forecasting methods, the future load is predicted subjectively based on using the opinions of experts; however, they are not purely guesswork, but they are developed structured approaches for obtaining good forecasts without using historical data. Hence, such methods are useful and implemented when historical data are not available or scarce. These methods include: ***subjective curve fitting, the Delphi method, and technological comparisons***. On the other hand, the quantitative/objective forecasting methods are based on mathematical and statistical formulations. They are applied when the data are available, but two conditions must be satisfied: numerical information about the past is available, and it is reasonable to assume that some aspects of the past patterns will continue in the future. The quantitative forecasting methods involve a wide range of methods, and each method has its own properties, accuracies, and costs that must be considered when choosing a specific method within specific disciplines for specific purposes. Quantitative methods include, among the others: ***regression analysis, decomposition methods, exponential smoothing, and the Box-Jenkins methodology***. Most quantitative prediction problems demand either the time series data collected at regular periods over time or cross-sectional data, which are collected at a single point in time. In order to summarize the structure/types of forecasting methods, one of the ways of classification can be composed as shown in Fig. 1 (a) [13, 29, 30]. Regarding the basic classification of the forecasting models (discussed in the next section), one of the possible ways is shown in Fig. 1 (b). An illustration of the more precise classification of the forecasting models will follow later in the text (c.f. Fig. 4).



a) One of the possible classifications of the major types of basic forecasting methods.



b) One of the possible basic classifications of the forecasting models.

Figure 1: Classification of the major types of basic forecasting methods and forecasting models

B. Forecasting Models

The early electrical load forecasting models were almost entirely limited to traditional statistical methods, but with the progress of modern science, load forecasting technologies have been considerably developed. Recently, forecasting models based on machine learning theories are becoming more and more popular in the power load forecasting [14]. This section defines and describes the most commonly used load forecasting models, whether traditional or modern intelligent models (c.f. figure 1b)). At this place, it is worth to mention that Fig. 1b) includes only some of the most commonly used clusters of the forecasting models. In the sequel, the discussion about the forecasting models is going to be directed in the following two directions:

- ***1. Statistical models.***
- ***2. Modern models based on machine learning, data mining, and artificial intelligence approaches.***

1. Statistical Models

The statistical model is a mathematical model that embodies a set of statistical assumptions concerning the generation of sample data. A statistical model represents the data-generating process with a considerably idealized form. The statistical model is usually specified as a mathematical relationship between one or more random variables and other non-random variables. Several statistical models have been developed for making predictions and forecasting, according to some criteria of optimal fit [31]. In the sequel, the following models are going to be briefly discussed:

- ***Box-Jenkins basic models (AR, MA, ARMA, ARIMA, ARMAX, and ARIMAX) [32-34].***
- ***Kalman Filtering Algorithms in the State space [35-40].***

- ***Grey models [41-50].***
- ***Exponential Smoothing (e.g., see [51] for an ELF, [52-54] for a general explanation).***

Autoregressive (AR) Model

The main idea of autoregressive models is that the current value of the series, y_t, can be expressed as a linear combination of previous/past loads, then the AutoRegressive (AR) model can be used to forecast future load values. A pth-order autoregressive, AR(p), model is defined as:

$$y_t - \sum_{i=1}^{p} \emptyset_i \, y_{t-\mathrm{i}} = \varepsilon_t \, , \qquad (1)$$

where: $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ are the unknown AR coefficients, while ε_t is random white noise. The order of the model tells how many lagged past values are involved. Thus, the AR model can predict future behavior based on past behaviors. It is used to forecast when there is some correlation between the current values of y_t in a time series and its past values, where y_t is also disturbed with the random noise ε_t. Autoregressive models have been used for decades in many fields, such as economics, electric load forecasting, and digital signal processing [15, 29, 55].

Moving Average (MA) Model

The moving average model that mimics the behavior of the moving average process, is a linear regression model that regresses the current values against the white noise of one or more past values. I.e. Moving average model can also be treated as a model in which the time series is regarded as a moving average (unevenly weighted) of a random shock series ε_t. Thus, the moving average model of order q "MA(q)" is given by:

$$y_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \, \varepsilon_{t-\mathrm{i}} \, , \qquad (2)$$

The noise series can be approximate by the forecast errors or model's residuals when the load observations become available. There exists a "duality", i.e., invertibility principle between the MA process and the AR(∞) process, that is, the moving average model can be rewritten (inverted) into an autoregressive form (of infinite order). However, this can only be done if the MA parameters follow certain conditions, that is if the model is invertible. Otherwise, the Box-Jenkins requirements about stationarity, invertibility, and stability of the model will be violated. [15, 29, 30, 55].

Autoregressive Moving Average (ARMA) Model

The Autoregressive Moving Average has been introduced in 1970 by George Box and Gwilym Jenkins [15, 32]. The ARMA(p, q) models represent a combination of an autoregressive models AR(p) and a moving average models MA(q). In the ARMA models, the current value y_t is expressed linearly in terms of its past values and in terms of current and previous values of the noise. Mathematically an ARMA(p, q) model is written as [29, 55, 56]:

$$y_t - \sum_{i=1}^{p} \emptyset_i \, y_{t-\mathrm{i}} = \varepsilon_t + \sum_{i=1}^{q} \theta_i \, \varepsilon_{t-\mathrm{i}} \, , \qquad (3)$$

ARMA models have been a popular choice and extensively applied to load forecasting researches due to their relative simplicity and effectiveness [55, 57].

Autoregressive Integrated Moving Average (ARIMA) Model

The AR, MA, or ARMA models, discussed above, can only be used for stationary time series data. Although in practice, many time series like those related to business and socio-economic possess a non-stationary behavior. Thus from the application perspective, the ARMA model is inadequate to describe the non-stationary time series appropriately. Therefore, the ARIMA models were proposed by Box and Jenkins in 1976 with a purpose to include the case of non-stationarity as well. The ARIMA Box–Jenkins models have three types of parameters: the autoregressive parameters $(\emptyset_1, \emptyset_2, \ldots, \emptyset_p)$, the moving average parameters $(\theta_1, \ldots, \theta_q)$, and the number of differencing *d* conducted to (1- *B)*, where *B* represents a lag operator. The mathematical formulation of the ARIMA(p, d, q) model using the lag polynomials is given below:

$$\emptyset(B).\nabla^d.y_t = \theta(B).\varepsilon_t\ , \text{ i.e.,} \qquad (4)$$

$$\left[1 - \sum_{i=1}^{p} \emptyset_i\, B^i\right].[1-B]^d.y_t = \left[1 + \sum_{j=1}^{q} \theta_j\, B^j\right].\varepsilon_t \qquad (5)$$

where p represents the order of the autoregressive, q represents the moving average terms, and d is the number of differences to make the original time series stationary. The general ARIMA(p, d, q) model is a non-seasonal model, and therefore the seasonal model for it is considered as an extension of the general one which can be written as ARIMA$(p, \mathrm{d}, q)\times(P, D, Q)_s$, where s refers to the number of periods per season and P, D and Q are the seasonal equivalents of $p, \mathrm{d},$ and q. Hence, the seasonal variants of ARIMA model are known as (SARIMA) models. Another useful generalization of ARIMA models is the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model, which allows non-integer values of the differencing parameter d. The ARFIMA has useful applications in modeling time series with a long memory. Accordingly, as detected in the literature, the ARIMA models and their variants have achieved considerable success for electric load forecasting [4, 29, 55, 56, 58, 59].

ARMAX and ARIMAX Models

The ARMA and ARIMA models use only the time and load as input parameters, besides the random noise that disturbs the whole process. Because loads depend on the weather and time of the day, exogenous variables sometimes can be included to give the ARMAX and ARIMAX models [29]. In the ARMAX (autoregressive moving average with exogenous inputs) model, the current value of the time series y_t is expressed linearly in terms of its past values, in terms of current and previous values of the noise, and additionally, in terms of present and past values of the exogenous variable(s). The ARMAX$(p, q, r_1, \ldots\ldots, r_k)$ can be compactly written as:

$$\emptyset(B).y_t = \theta(\mathrm{B}).\varepsilon_t + \sum_{i=1}^{k} (\psi_0^i + \psi_1^i.\mathrm{B} + \cdots + \psi_{ri}^i.B^{ri}).v_t^i\ , \qquad \text{i.e.,}$$

$$\emptyset(B).y_t = \theta(\mathrm{B}).\varepsilon_t + \sum_{i=1}^{k} \psi^i(B).v_t^i\ , \qquad (6)$$

where the r_i's are the orders of the exogenous factors (variables) v_t^i and ψ^i(B) is a shorthand notation for ψ^i(B) = $(\psi_0^i + \psi_1^i.\mathrm{B} + \cdots + \psi_{ri}^i.B^{ri})$ with the ψ_j^i 's being the corresponding coefficients [55]. From the expression (6) follows:

$$y_t = \frac{\theta(\mathrm{B})}{\emptyset(B)}.\varepsilon_t + \sum_{i=1}^{k} \frac{\psi^i(B)}{\emptyset(B)}.v_t^i\ ,$$

$$y_t = \frac{\theta(B)}{\phi(B)} \cdot \varepsilon_t + \sum_{i=1}^{k} \tilde{\psi}^i(B)\, v_t^i \qquad (7)$$

where $\tilde{\psi}^i$'s are the adequate coefficient polynomials $\frac{\psi^i}{\phi}$ [55]. Similar expressions as for the ARMAX model can be derived for the ARIMAX model, except that the integrated (I) part must be additionally considered (via the differencing operator, similarly as for the ARIMA model) (see [32, 60]).

Kalman Filtering Algorithm in the State space

Forecasting, especially the long-term forecasting, is characterized by a high level of uncertainty due to its high dependence on socioeconomic factors; for this reason, an error level up to 10% is acceptable [29]. In this spirit, applying a Kalman filtering algorithm can significantly minimize the mean of the squared model's error. The Kalman filter (KF) is a set of mathematical equations in the state space that can provide an efficient computational (recursive) means to estimate the state of an observed process. The Kalman filter is named after Rudolph E. Kalman, who published his famous paper in 1960 describing a recursive solution to the discrete-data linear filtering problem. The Kalman filter has been hired extensively for tracking in interactive computer graphics. It has been used for motion prediction, and it is also used for multi-sensor (inertial-acoustic) fusion. Moreover, this filter is very powerful in several other aspects: it supports estimations of past, present, and future states, as well as it can do so even when the precise nature of the modeled system is unknown. A Kalman filter is also a potent tool when it comes to controlling the noisy systems, in which the electric power systems undoubtedly can be assigned [61]. According to Gaur and his colleagues [40], the main elements that affect the electric load behavior can be classified as follows:

a. Weather: This factor is the most essential extreme. It comprises humidity, wind speed, temperature, precipitation, etc. The variations in these factors straight lead to the adjustment in the habit patterns of appliances such as heaters, air conditioners, coolers, and so on.

b. Time: This factor impacts electric load at different daily periods, weekdays and weekends, holidays, and year's seasons. Here, the time-dependent electric load variation can mirror the people's lifestyle, such as their work schedules, leisure time, sleeping patterns, etc.

c. Economy: This factor is important in the deregulated market, reflecting the variable electricity price, while the load management policy has an important impact on the electric load growth or decline trend.

d. Random disturbances: The shutdown or start-up of the enormous loads such as steel mill, or wind tunnels are going to lead to the load curve impulses. The other abnormal events which are priorly known but have an uncertain effect on the load, also fall in the random disturbance category.

e. Customer factors: These factors include the consumption type (commercial, residential, agricultural, or industrial), the size of the buildings, the number of employees, as well as the number of electric utilization.

The factors mentioned above can be injected as inputs into the Kalman filter. Since it is extremely difficult to deal with the complex inputs (c., d., e.), the weather and time factors are usually included in the KF only [40]. The KF is usually looked through the eyes of discrete-time linear dynamic system. The latter has the latent state space vector $\mathbf{x}^{\mathrm{T}}(k) = \left[x_1(k) \quad x_2(k) \quad \ldots \quad x_n(k) \right]$ signifying the vector of the hidden states of the appliances, as well as the observation vector $y(k)$ demonstrating the smart meter device (SMD) readings. Here, k represents the discrete-time moments. The delayed estimator produces the state estimates $\mathbf{x}^{\mathrm{T}}(k|k-1)$ and output estimates $y(k|k-1)$ by applying the measurements only up to the *(k-1)-th* output $y(k-1)$. The KF's mechanism works in a two-step process, that is, the predictor step (PS), and the corrector step (CS). In the PS, the KF estimates the current load's state on the basis of its previous state, together with its covariance uncertainty. Once the new SMD's measurement is observed, the estimated state vector is updated by deploying a weighted average, where the higher weight is given to the estimate with a higher certainty. In such a way, the PS and CS steps are continuing to proceed recursively. Since the linear KF often cannot satisfy the rigor demands regarding the forecasting accuracy in the case of serious nonlinearities of the given problem, there have been developed several of its nonlinear variants. Thus, in order to explore the hidden nonlinearities of the problem, the Extended Kalman filter (EKF), as well as the Unscented Kalman filter (UKF), are also occasionally used.

Grey Models (GM)

Grey system theory (GST) was firstly introduced by Deng in 1982 [62]. This theory can deal with the observed systems that have partially unknown parameters, while the grey models need only a limited amount of the data to estimate the unknown system's behavior. The main task of the GST is to extract convincing governing laws of the observed system on the basis of the available data, regardless of their complexity or chaoticity [62]. The GM(1,1) grey model is one of the most frequently used grey models, which can produce forecasts of the future primitive data points. The GM(1,1) is a time series forecasting model with differential equation (DE) having the time-dependent varying coefficients. By deploying the so-called accumulated generation (AG), it becomes possible to smooth and thus lower the intensity of the uncertainty in the system. When the DE is solved, the n-step ahead predicted value of the system can be obtained. Based on using the latter, the inverse AG (IAG) can be applied to extract the predicted values of the original data [62]. According to [63], the grey forecasting models based on the grey system theory are extensively used in networks, since the models are capable of using random variations as the grey quantity, which is changed in a certain interval. The differential equation of the grey model is essential here since it provides the means to forecast the power load. When the derived model is successfully tested against adequate reliability, stability, and accuracy, it can be deployed to forecast the future load. The grey models are suitable for all three types of load forecasting, i.e., for the short-term, medium-term and long-term forecasting [63]. One of the main advantages of the GMs is that they can be developed without considering the load distribution and the changed load's trend [5]. However, their deficiency is that they are appropriate only for effective solving of the problems with the prevailing exponential growth trends [9].

Exponential Smoothing (ES)

Exponential smoothing is a pragmatic forecasting approach, whereby the prediction can be carried out from the exponentially weighted average of the past observations [54, 55, 64]. The highest weight is given to the present observation, lower weight to the immediately preceding measurement, even lower weight to the measurement before that, and so on (i.e., we are dealing with the exponential decay of the influence of past historical data) [64]. The ES models are among the most common and prevalent statistical forecasting methods because of their accuracy, simplicity, robustness, and low cost [59]. They are also crucial for the purpose of a load forecasting in the power systems. As emphasized by Peirong and his colleagues [51], the accuracy of the model importantly depends on smoothing coefficients of the EF model. This study also demonstrates how to seek the best smoothing coefficients.

In general, there exist three different exponential smoothing techniques: simple, single exponential smoothing (SES) (***Brown's method***), double exponential smoothing (DES) (***Holt's method***), and triple exponential smoothing (TES) (***Holt-Winters method***). The SES model requires a little calculation, and it is used when the data pattern neither has a cyclical or seasonal variation, nor the trend in the historical data. On the other side, the DES models, which are particularly used in economics, enable the forecasted values with a trend included. Finally, the TES (Holt-Winters) models have two possible modes of computation: the additive and the multiplicative. The additive model is used if the original data are showing stable seasonal fluctuations. Conversely, the multiplicative models are used when the original data are reflecting the significant variations of the seasonal fluctuations [65].

According to the empirical evidence, the basic Holt-Winters method tends to produce over-forecasts or under-forecasts, particularly for the longer forecasting horizons [53, 54]. For this reason, in the year 1989, Gardner and McKenzie have deployed a new parameter ϕ associated with the trend component. This way, the trend is dampened to a flat line, when the future becomes more distant. Consequently, we can get the ***Holt-Winters Damped Additive model*** in the following form, including the four parameters $\alpha, \beta, \gamma, \phi$ [53, 54]:

$$
\begin{aligned}
&\mathit{Level}: && \ell(t)=\alpha\cdot\left(y(t)-s(t-m)\right)+(1-\alpha)\cdot\left[\ell(t-1)+\phi\cdot b(t-1)\right]\\
&\mathit{Growth}: && b(t)=\beta\cdot\left[\ell(t)-\ell(t-1)\right]+(1-\beta)\cdot\phi\cdot b(t-1)\\
&\mathit{Seasonality}: && s(t)=\gamma\cdot\left(y(t)-l(t-1)-\phi\cdot b(t-1)\right)+(1-\gamma)\cdot s(t-m)\\
&\mathit{Forecast}: && \hat{y}(t+h|t)=\ell(t)+\phi_h\cdot b(t)+s\left(t-m+h_m^+\right),\\
&\mathit{where}: && \phi_h=\phi+\phi^2+\ldots+\phi^h,\ h_m^+=\left[(h-1)\bmod m\right]+1,\ m-\mathit{number\ of\ seasons}|\mathit{year}\\
&h-\mathit{time\ points\ of\ the\ future\ horizon}\\
&y-\mathit{observed\ time\ series}
\end{aligned}
$$

2. Models of Artificial Intelligence, Computational Intelligence, or Machine Learning

Traditional statistical models are limited and sometimes might lead to unsatisfactory solutions. The reason is the too high number of computational possibilities leading to large solution times and the complexity of certain non-linear data patterns. Hence, machine learning and artificial intelligence-based techniques provide a promising and attractive alternative [15, 56].

Artificial Neural Network (ANN) Algorithms

The artificial neural network (ANN) approach was discovered in 1990 by Warren McCulloch and Walter Pitts as an alternative mechanism to the time series forecasting. The ANNs have been successfully applied in many different areas, especially for forecasting and classification purposes. ANNs models have been used and studied intensively as a tool to be used for electric load forecasting and gained huge popularity in the last few decades [56]. Basically, the neural network is a non-linear circuit that is capable of doing non-linear curve fitting. It represents an information processing paradigm that was inspired by the way the biological systems of humans, such as the brain, are able to process a certain piece of information. In this process, the ANNs try to recognize regularities and patterns in the input data, learn from experience, and then provide generalized results based on their known previous knowledge. An ANN is composed of several interconnected processing elements (PE), called neurons, which are changing their dynamic state response with respect to external inputs [3, 15, 56]. The simplest form of an artificial neural network containing input, hidden and output layers is shown in Fig. 2. Here, the input values to the hidden node (*I*), associated weights (w_i), hidden layer function *f(x),* and the output results (*Y*) can be seen. By changing the weights of the ANNS, the preferred output from a specific input can be achieved. Such a process, where the weights of the ANNs are adjusted, is an iterative training process [66]. The commonly employed ANN algorithms for electric load forecasting are [3, 5, 14, 16-18, 39, 41, 55, 57, 66-70]:

- ***feed-forward (FF) neural networks,***
- ***NARX (nonlinear autoregressive with exogenous inputs) neural networks,***
- ***back-propagation (BP) neural networks,***
- ***radial basis function (RBF) neural networks,***
- ***the random neural networks,***
- ***recurrent neural networks, and***

- ***self-organizing competitive neural networks.***

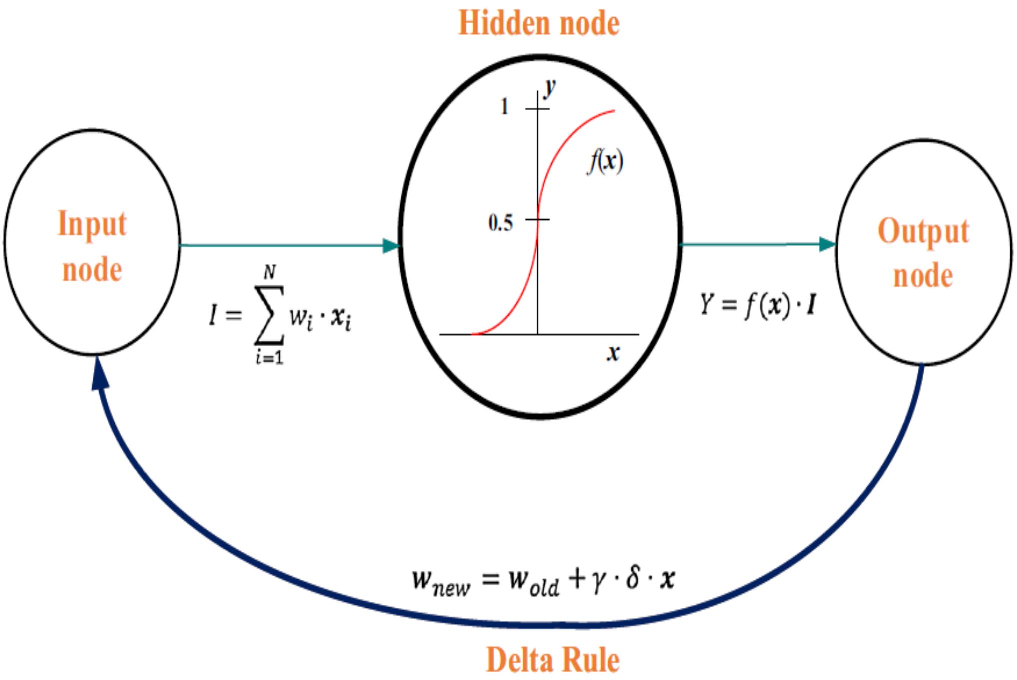


Figure 2: A simple neuron scheme in an ANN (Source: Nalcaci, Özmen, and Weber (2018)[3]).

According to [55], the proposed ANN models could be classified into the two main groups:

- ***The first group** includes those ANNs that have **only one output node**, used to forecast the next hour's load, next day's peak load, or next day's total load.*
- ***The second group** involves models that have **several output nodes** to forecast a sequence of hourly loads, typically 24 nodes, with a purpose to predict the next day's entire load profile.*

Based on [5, 9], the second group has strong robustness and strong learning ability. However, the ANN quickly falls into the local minimum because of the restriction on the generalization ability and cannot make full use of information due to the small sample size selected. Besides, the learning convergence speed is slow.

According to (Momani et al., 2015) [66], the NARX ANNs usually outperform the classical neural networks such as, for example, the FF time-delay networks. Besides, the NARX networks, together with the RBF networks, are well-known to possess some excellent properties such as their simplicity, reliability, low % error, high accuracy, and ability to create the nonlinear relationships between variables. In the aforementioned work of Momani and his colleagues, it is demonstrated that the NARX and RBF ANNs are able to provide highly accurate EL forecasts for five days ahead, while the relative error keeps its values below the reasonable level between 0.1% to 3.9% [66]. In this study, the NARX's output *y(t)* represents the predicted hourly load (in megawatts), which is a function of the inputs, including the humidity $u_1(t)$ and its past values, temperature $u_2(t)$, and its historical data, and the prior hourly load values $y(t-i), i=1,\dots,n_y$. Accordingly, the NARX ANN model can be mathematically described in the following form for this case study [66]:

$$y(t) = f_{NARX}\left(u_1(t), u_1(t-1), \dots, u_1\left(t-n_{u_1}\right), u_2(t), u_2(t-1), \dots, u_2\left(t-n_{u_2}\right), y(t-1), y(t-2), \dots, y\left(t-n_y\right)\right)$$

The Extreme Learning Machines (ELM)

The Extreme learning machines represent the special class of the FF ANNs and are appropriate for regression, classification, clustering, feature learning, and sparse approximation. They can be used for forecasting purposes as well. Huang, Zhu, and Siew proposed the extreme learning machines in 2004 [71]. They usually address a single-hidden layer FF neural network. In the ELM, the weights of hidden layer nodes are randomly selected, and a least-squares solution can analytically determine the output weights of ELM. The latter means that besides the weights that are connecting inputs to hidden nodes, the parameters of the hidden nodes need not be adjusted as well. Conversely, the hidden nodes can be randomly allocated and, afterward, never updated. The output weights of the hidden nodes are usually settled in a single step, which substantially decreases the time needed for the learning of the ANN. According to Huang and his colleagues [71], the ELM networks are capable of producing good generalization performance and can learn even thousands of times quicker than the competitive networks, which are trained using backpropagation. It is also detected from the literature that ELM models can outperform even the support vector machines, which are providing the sub-optimal solutions in both regression and classification problems.

The ELM models have also been extensively used in the field of the electric load forecasting, see for example works [72-75]. There have also been successful attempts to improve the basic ELM scheme, e.g., Ertugrul [74] has applied a novel recurrent ELM approach for the purpose of ELF forecasting, Garcia-Laencina [76] has deployed a mechanism to improve the forecasting by conducting a linear combination of multiple ELM machines and demonstrated the performance on three engineering problems, etc. The basic scheme of the ELM network is shown in Fig. 3. [74, 77].

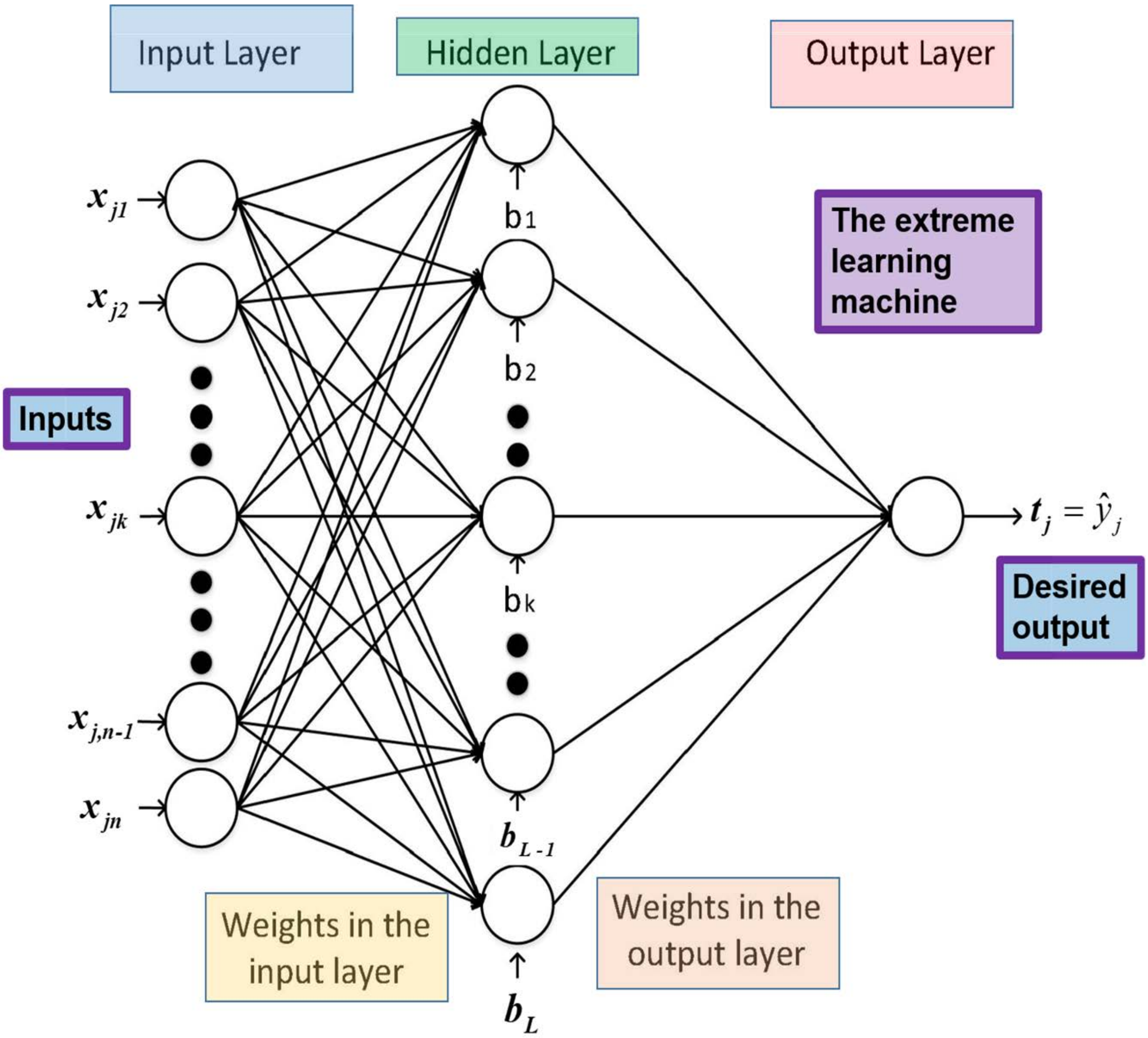


Figure 3: Diagram of the Extreme Learning Machine (Sources: Ertugrul; Albadr et al. [74, 77])

Let us assume that we have *N* arbitrary unique samples $(\boldsymbol{X_i}, \boldsymbol{y_i}), \boldsymbol{X_i} = [x_{i1}, x_{i2}, ..., x_{in}]^T \in R^n, \boldsymbol{y_i} = [y_{i1}, y_{i2}, ..., y_{im}]^T \in R^m$. Mathematically, the ELM can be represented by the following equation (see Fig. 3) [71, 74, 77]:

$$\sum_{i=1}^{L} \boldsymbol{\beta_i} \cdot g_i(\boldsymbol{X_j}) = \sum_{i=1}^{L} \boldsymbol{\beta_i} \cdot g_i(\boldsymbol{W_i} \cdot \boldsymbol{X_j} + b_i) = \sum_{i=1}^{L} \boldsymbol{\beta_i} \cdot g_i(W_i \cdot X_j + b_i) \quad (8)$$

Here, $g_i(\bullet)$ are the activation functions, $\boldsymbol{W_i} = [W_{i1}, W_{i2}, ..., W_{in}]^T$ is the vector, whose weights link the hidden node and the *i-th* input nodes, while $\boldsymbol{\beta_i} = [\beta_{i1}, \beta_{i2}, ..., \beta_{im}]^T$ is the vector, whose weights link the hidden node and the *i-th* output nodes. The parametes b_i represent the ith hidden node's threshold (bias), while the product $\boldsymbol{W_i}.\boldsymbol{X_j}$ signifies the inner product between $\boldsymbol{W_i}$ and $\boldsymbol{X_j}$ [71, 74, 77]. The value *L* is the number of the neurons in the hidden layer. The learning algorithm is based on trying to obtain the zero error mean of the *N* training data, which means that the following condition must be achieved: $\sum_{j=1}^{L} \|\boldsymbol{y_j} - \boldsymbol{t_j}\| = 0, t - desired\ output$. Then it can be shown that there exist such $\boldsymbol{\beta_i}, \boldsymbol{W_i}, b_i$ that the following system of equations can be written [71, 74, 77]:

$$\begin{aligned}
&\sum_{i=1}^{L} \boldsymbol{\beta_i} \cdot g_i(W_i \cdot X_j + b_i) = \boldsymbol{t_j},\ j = 1, ..., N \\
&follows: \\
&\boldsymbol{\beta_1} \cdot g_1(W_1 \cdot X_j + b_1) + \boldsymbol{\beta_2} \cdot g_2(W_2 \cdot X_j + b_2) + ... + \boldsymbol{\beta_L} \cdot g_L(W_L \cdot X_j + b_L) = \boldsymbol{t_j},\ j = 1, ..., N \\
&follows: \\
&\boldsymbol{\beta_1} \cdot g_1(W_1 \cdot X_1 + b_1) + \boldsymbol{\beta_2} \cdot g_2(W_2 \cdot X_1 + b_2) + ... + \boldsymbol{\beta_L} \cdot g_L(W_L \cdot X_1 + b_L) = \boldsymbol{t_1} \\
&... \\
&\boldsymbol{\beta_1} \cdot g_1(W_1 \cdot X_N + b_1) + \boldsymbol{\beta_2} \cdot g_2(W_2 \cdot X_N + b_2) + ... + \boldsymbol{\beta_L} \cdot g_L(W_L \cdot X_N + b_L) = \boldsymbol{t_N}
\end{aligned} \quad (9)$$

Thus, we have the following vector-matrix system of equations [71, 74, 77]:

$$\underbrace{\begin{bmatrix} g_1(W_1 \cdot X_1 + b_1) & g_2(W_2 \cdot X_1 + b_2) & ... & g_L(W_L \cdot X_1 + b_L) \\ g_1(W_1 \cdot X_2 + b_1) & g_2(W_2 \cdot X_2 + b_2) & ... & g_L(W_L \cdot X_2 + b_L) \\ ... & ... & ... & ... \\ g_1(W_1 \cdot X_N + b_1) & g_2(W_2 \cdot X_N + b_2) & ... & g_L(W_L \cdot X_N + b_L) \end{bmatrix}}_{\mathbf{H}_{N\times L}} \cdot \underbrace{\begin{bmatrix} \boldsymbol{\beta_1}^T \\ \boldsymbol{\beta_2}^T \\ ... \\ \boldsymbol{\beta_L}^T \end{bmatrix}}_{\boldsymbol{\beta}_{L\times m}} = \underbrace{\begin{bmatrix} \boldsymbol{t_1}^T \\ \boldsymbol{t_2}^T \\ ... \\ \boldsymbol{t_L}^T \end{bmatrix}}_{\mathbf{T}_{N\times m}}$$

The vector $\boldsymbol{\beta}_{L\times m}$ can then be estimated by using the Moore-Penrose generalized inverse method as follows [71, 74, 77]:

$$\begin{aligned}
&\hat{\boldsymbol{\beta}}_{L\times m} = \mathbf{H}^+_{N\times L} \cdot \mathbf{T}_{N\times m}, \\
&\mathbf{H}^+_{N\times L} - The\ Moore - Penrose\ generalized\ inverse\ \ for\ \mathbf{H}_{N\times L}
\end{aligned}$$

Further discussion and details about the ELM networks and their more advanced variants can be found in [71-74, 76-80]. There, a more detailed explanation about efficient using of the ELM networks for the purpose of ELF prediction can be found as well.

Support Vector Machines (SVMs)

The support vector machines (SVMs) are regression and classification mechanisms [55], which was first presented by Vapnik in 1992. Later, in 1995, the soft margin classifier was proposed by Cortes and Vapnik utilizing the statistical learning theory. Initially, the SVMs were developed to deal with pattern classification problems (e.g., a face identification, optical character recognition, early medical diagnostics, the text classification, etc.). Afterward, Vapnik has extended the use of SVMs to be deployed for the regression algorithms as well (i.e., the support vector regression - SVR). Over the last two decades, the SVMs have received growing attention not only for pattern recognition and regression analysis but also for the forecasting purposes and solving of the time series prediction problems [9, 15, 55, 56]. The main objective of the SVMs is to deduct a specific decision rule with a satisfactory generalization ability by choosing some specific subset of training data, called support vectors.

In the SVM models, a nonlinear mapping of the input space into a higher dimensional feature space is deployed, and afterward, an optimally separating hyperplane is extracted. Accordingly, the complexity and quality of the SVM solutions do not directly depend on the input space. When designing an SVM model, the training process that is comparable to solving of a linearly constrained quadratic programming problem is carried out. Therefore, conversely to the other networks' training, the SVM solutions appear to be always globally optimal and unique. On the other side, the main weakness of the SVMs is that it requires an enormous amount of computations, and consequently, the time complexity of the solutions is radically increased [9].

The SVM models have also been widely used for the ELF forecasting. Fu et al. [81] have used the SVMs to predict the next day's electricity load of public buildings, while Hong [82] has combined the SVMs with the immune algorithm to forecast the electric load. Hu and his colleagues [83] have demonstrated that the SVM forecasting model, whose parameters were adjusted by a firefly based memetic algorithm, can significantly outperform the other evolutionary-based SVR models, besides some of the classical forecasting models. Furthermore, Qiang and Pu [84] have deployed the SVMs based on the particle swarm optimization to apply a short-term load forecasting. There have also been detected many other papers in the literature that are dealing with the ELF forecasting. The main objective of most of the authors was to develop efficient (nature-inspired and other) algorithms for an appropriate setting and adjusting of the SVMs' parameters. Some papers are also dedicated to the automatic model selection by means of SVMs. For example, Maldonado et al. [85] have used the ideas of SVR for feature selection purposes, which resulted in the developed mechanism for automatic lag selection. Here, the correct identification of relevant lags and seasonal patterns was demonstrated, while the entire forecasting mechanism was used to efficiently predict the electricity demand forecasting.

The details of the derivation of the SVM (and SVR) models can be found in the literature (e.g., see [81-89]). When addressing the SVM models, the structured risk minimization risk principle is considered instead of finding the minimum empirical errors [81, 83]. If we have a training dataset $\{(x_i, y_i)\} \in R^n \times R,\ x_i - input,\ y_i - output\ (target), i = 1,...,n$, the objective of the SVMs is to generate the decision function $y_{des} = f_{des}(x)$ ($\langle \bullet \rangle$ is the inner product) by means of minimization of a so-called regularized risk function *R* (*b* is the bias) [81, 83] :

$$y_{des} = f_{des}(x) = \langle w, \phi(x) \rangle + b,$$
$$w - the\ weight\ vector\ (controls\ the\ model's\ smoothness)$$
$$\phi(x) - the\ high-dimensional\ \ feature\ space$$
$$R = R_{emp} + \frac{1}{2}\cdot\|w\|^2 = \frac{C}{n}\cdot\sum_{i=1}^{n} L(y_i, f_{des}(x)) + \frac{1}{2}\cdot\|w\|^2 \qquad (10)$$
$$where:$$
$$L(y_i, f_{des}(x)) = \begin{cases} |y - f_{des}(x)| - \varepsilon & if\ |y - f_{des}(x)| \geq \varepsilon \\ 0, & otherwise \end{cases}$$

Here, R_{emp} is the empirical risk, while the function L is $\varepsilon - insensitive$ Vapnik's loss function, and ε is named the "tube size" [81, 83]. The second term in R, i.e., $\frac{1}{2}\cdot\|w\|^2$ is the so-called regularization part included as a measure of the flatness (complexity) of the function. The main purpose of the regularized constant C is to find the appropriate trade-off between the regularization term and empirical risk. The constants C and ε are the user-defined parameters and can significantly influence on the behavior of the entire SVM model. The equations in (10) can be transformed into the suitable form of the following minimization problem [83]:

$$minimize \quad R = \frac{C}{n}\cdot\sum_{i=1}^{n} L(y_i, f_{des}(x)) + \frac{1}{2}\cdot\|w\|^2 = C\cdot\sum_{i=1}^{n}(\xi_i + \xi_i^*) + \frac{1}{2}\cdot\|w\|^2$$
$$subjected\ to:$$
$$y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i$$
$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^*$$
$$\xi_i, \xi_i^* \geq 0\ (slack\ variables), \qquad i = 1, \ldots, n$$

Concerning the Wolfe's dual theorem and the saddle-point condition, the dual optimization problem can also be written in the following form [83]:

$$\underset{\alpha,\alpha^*}{maximize} \quad -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)\cdot(\alpha_j - \alpha_j^*)\cdot\langle \phi(x_i), \phi(x_j) \rangle - \varepsilon\cdot\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i\cdot(\alpha_i - \alpha_i^*)$$
$$subjected\ to:$$
$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0, \qquad \alpha_i, \alpha_i^* \in [0, C] \Rightarrow nonnegative\ Lagrange\ multipliers \qquad (11)$$
$$with$$
$$w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\cdot\phi(x_i)$$

Based on obtained $w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\cdot\phi(x_i)$ from (11) and the decision function $y_{des} = f_{des}(x)$ from (10), the later takes the following form [83]:

$$y_{des} = f_{des}(x) = \langle w, \phi(x) \rangle + b = \left\langle \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\cdot\phi(x_i), \phi(x) \right\rangle + b = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\cdot K(x_i, x_j) + b \qquad (12)$$

In equation (12), $K(x_i, x_j)$ represents a kernel function. Its value corresponds to the inner product of the two vectors x_i, x_j in the feature space $\phi(x_i), \phi(x_j)$. From this follows: $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$. It turns out that any function satisfying the Mercer's condition might have been used as the kernel function (e.g., the sigmoid kernel, radial basis function kernel, polynomial kernel, etc.) [83].

Fuzzy Logic

In many research fields, there have been a countless number of books (see for example books [26, 90-109]) and papers already published about the fuzzy logic (FL), the mechanisms behind this logic, and the basic principles. Fuzzy logic is a generalization of the conventional Boolean theory, but instead of getting a value of 0 or 1 for input, it has associated with it specific qualitative ranges. In other words, for instance, a temperature may be low, medium, or high; however, using fuzzy logic allows outputs to be deduced from noisy or fuzzy inputs and without a need to specify a precise mapping of inputs to outputs [55]. The fuzzy methods are very useful for handling uncertainties and are essential for the knowledge acquisition of human experts. A membership function can be represented for every fuzzy set, where a function for any fuzzy set, or a membership function, exhibits a specific continuous curve that is changing from 0 to 1 or vice versa, while a corresponding transition's region represents a fuzzy boundary of the term [29]. Fuzzy theory is often combined with the other methods to achieve good prediction results [5]. In short, fuzzy logic can be used in such cases [29] :

- The mathematical model does not exist, or it exists but is too difficult to encode.
- The mathematical model is too complex to be evaluated fast enough for real-time operation.
- The mathematical model includes too much memory on the designated chip architecture.
- The expert is available who can specify the rules underlying the system behavior and the fuzzy sets that represent the characteristics of each variable.
- The system has uncertainties in either its inputs or definition.
- The systems are: too complicated, too non-linear, or with too much uncertainty to implement using the traditional techniques.

However, we have to avoid using fuzzy logic for systems in which conventional control equations and models are already optimal or entirely adequate [29]. In the field of electric load forecasting, books [26, 96, 102, 103] are dedicated to the modeling and prediction with a particular emphasis on computational and artificial intelligence approaches, including the Fuzzy logic models. Besides, Ali and his colleagues have deployed the fuzzy logic models to the short-term ELF [110], and long-term load forecasting [111]. The ELF forecasting was also addressed in works [112-115]. Further, Jamaaluddin et al. [116] were dealing with a very short-term load forecasting of a peak load time by using the fuzzy logic, while Jagbir and Singh [117] have conducted an FL-based short-term load forecasting model for a 220 kV transmission line. Laouafi et al. [118] have developed an adaptive neuro-fuzzy inference system-based approach for daily load curve prediction, while Yao et al. [114] have applied a short-term load forecasting by interval Type-2 FL system.

Wavelet Neural Networks (WNNs)

Wavelet theory is a mathematical theory that has been proposed by Grossman and Morlet in the 1980s. Later, Zhang has suggested the wavelet neural network (WNN) in 1992 in order to take advantage of both the wavelet functions and the widely used neural network. The WNNs are advocated as an alternative to the feedforward neural networks for approximating the arbitrary non-linear functions by means of a wavelet transform theory. A WNN takes the wavelet space in the spirit of feature space for pattern recognition, and it can recognize a feature extraction of the signal by calculating the internal product of the wavelets base and the signal vector. Hence, the network can efficiently learn the input and output characteristics of the system without too much prior information.

The signal in WNN is transmitted forward and the error is proliferated backward, by which a more accurate predictive value of the signal is achieved. The WNNs possess rough capability and are robust for approximating the non-linear functions [14]. The comprehensive review of the ELF forecasting by means of wavelet networks can be found in Patel et al. [119].

Genetic Algorithms (GA)

In the 1950s and 1960s, there were more computer scientists, who had independently studied evolutionary systems with the idea that evolution could be used as an optimization tool for solving engineering problems. The idea in all these systems was to develop a population of possible solutions to a given problem with the help of operators, which mimic a genetic variation and natural selection [120]. John Holland was the first who has introduced genetic algorithms and developed his idea in a book entitled "Adaptation in natural and artificial systems". David Goldberg finally popularized the GAs in 1989. Genetic algorithms have become one of the most popular and used techniques of evolutionary computation [121]. GAs represent a plethora of optimization and search techniques based on the principle of genetics and natural selection. They allow the population to be composed of more individuals exposed to specific selection rules that maximize the success of the solution ('fitness') or minimize the cost function. Holland presented GA as a heuristic (metaheuristic) method based on the principle of 'survival of the best'. This has made GAs a handy tool in resolving severe optimization problems [121]. To summarize, genetic algorithm represents a type of optimization algorithm to find the optimal solution(s) to a given computational problem that maximizes or minimizes a criterion function. It represents an important branch of the field of evolutionary computation [122]. In the last two decades, the GAs have been widely used and successfully applied to various types of optimization problems [123]. The basic components of genetic algorithms are [122, 124]:

1) A fitness function for optimization; the function that the algorithm is trying to optimize, and it is one of the most essential parts of the algorithm.

2) A population of chromosomes; chromosome refers to values that represent a candidate solution to the problem we are trying to solve. A genetic algorithm starts with a randomly chosen set of chromosomes, which serves as the first generation or initial population. Then each chromosome in the population is evaluated by the fitness function to test how well it solves the problem.

3) A selection of which chromosomes will reproduce; based on a probability distribution defined by the user.

4) A crossover to produce the next generation of chromosomes; the crossover operator resembles the biological crossing over and recombination of chromosomes in cell meiosis.

5) A random mutation of chromosomes in the new generation that randomly flips individual bits in the new chromosomes.

In the field of electric load forecasting or any kind of forecasting, the GAs have been widely used. Namely, they are frequently well-suited with nonlinear systems and they conduct a particular optimization based on the natural selection of the optimal solutions found from a wide range of the forecasting model candidates' populations [125]. Such kind of GA-based optimization is usually deployed during the model selection procedure when the most appropriate parameters of the forecasting model must be found. For instance, in work [125], the GAs have been applied to find the optimal *p, d,* and *q* parameters of the ARIMA model. Regarding the ELF forecasting, Aquino et al. [126] have employed the GAs to develop a neural-network-based load forecasting model. Gupta and Sarangi [127] have used the GA-based back-propagation method for efficient ELF forecasting. Khan et al. [128] have reported about very short-term load forecasting using Cartesian Genetic Programming evolved Recurrent Neural Networks. Moreover, the GA-based ELF forecasting is discussed in many other works, e.g., see [129, 130].

Expert Systems

Based on [55], the expert system is a computer program that has the ability to explain, understand, and expand its knowledge base as new information becomes available. Expert systems combine rules and procedures used by human experts. An expert's knowledge must be convenient for codification into software rules. In particular, the experts must be able to explain their decision process to programmers. This knowledge is later codified as facts and IF-THEN statements. It constitutes a set of relationships between the changes in the system load and changes in the exogenous factors that affect the load. Over time, some of the rules do not change, while others may have to be continuously updated.

3. Hybrid Methods and the comprehensive classification of forecasting models

Hybrid or combination models and methods can obtain better forecasting performance than the single model by integrating the advantages of different single forecasting models and therefore are widely used in many forecasting areas. In this sense, there are numerous available forecasting methods, optimization algorithms, and data processing techniques for developing different hybrid models [1, 9]. Thus, recent studies have transferred their primary research focus on the development of effective hybrid models with the hope of improving prediction performance [14]. Accordingly, it is reasonable to find modern hybrid methods, which are presented in order to cover the new technological frameworks [131]. However, there are no confessed references on how to select different methods to build a hybrid model [9]. All the widely used electrical load forecasting models that have been aforementioned and discussed in the previous two sub-sections are summarized as shown in Fig. 4.

Logistics & Sustainable Transport
Vol. 11, No. 1, February 2020, 51-76
doi: 10.2478/jlst-2020-0004

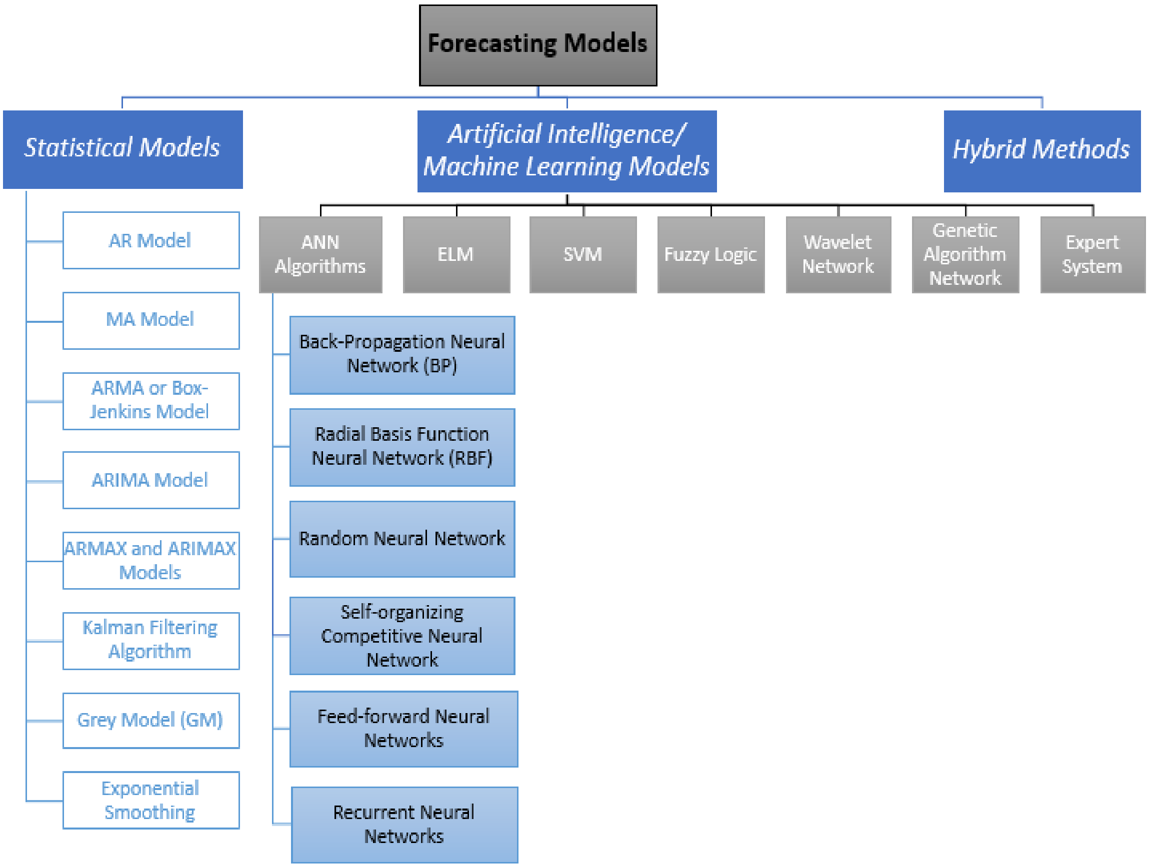


Figure 4: A comprehensive classification of forecasting models

Source: Adapted by the authors

III. THE BRIEF SYSTEMIC REVIEW OF THE ELECTRIC LOAD FORECASTING LITERATURE

The study from hereafter is based on the review of academic research aimed at electricity load forecasting. Therefore, this section describes the systematic process used for the review. This study has been applied for the conventional review without the systematic review as an initial step for discovering the general ideas of electric load forecasting models. Thus, there was no urgent need to follow the strict systematic review criteria and protocols. However, well-defined steps in order to select accurate sources and publications were followed. Firstly, the following keywords had been used: “electric and energy”, “models and methods”, and “load forecasting” in English only. Searching based on these keywords has addressed the online database Web of Science (WOS), the most trusted citation databases in the world.

The initial search results have given N = 276 scientific papers. Secondly, the research area had been narrowed down in order to find the M = 145 most relevant studies by excluding the presence of the aforementioned keywords in the main text and leave the latter only in the context of their referring in the title and abstract. Subsequently, a quick overview of the found works has enabled the identification of some irrelevant studies and papers associated with specific keywords like: “electricity pricing”, “electric vehicles”, or “wind power”. After their removal, in the end, K = 52 works have been identified and had constituted the basis of this review study. Among the selected 52 studies, 30 are articles, 15 are conference papers (together L = 45 papers), four are books, and three are theses, as shown in Fig. 5. All of them have been reviewed in depth.

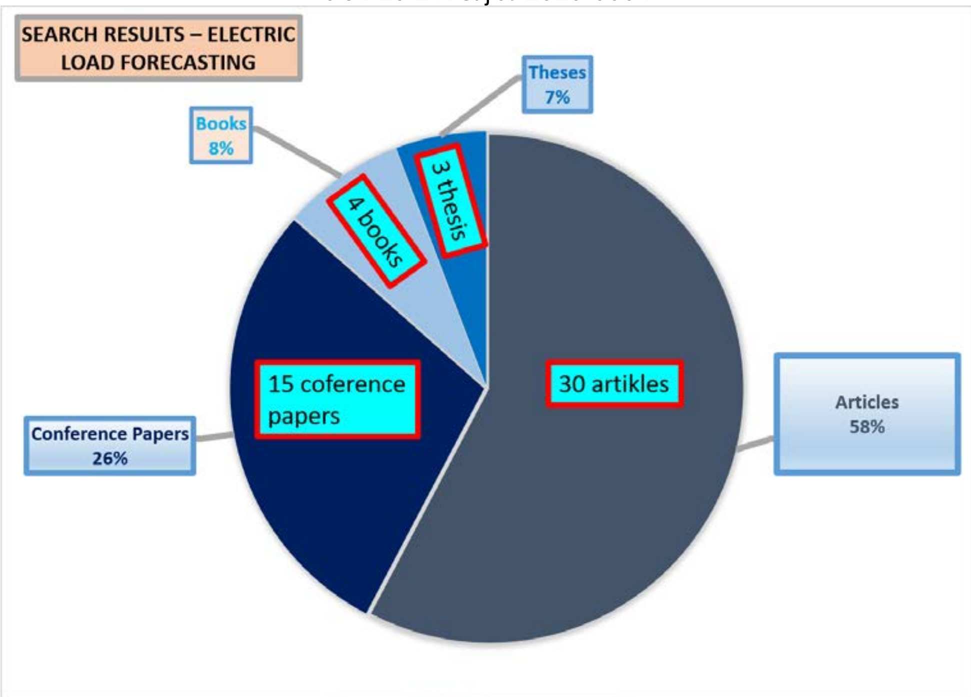


Figure 5: Search results by percentages for a review study (for K = 45 papers)

IV. RESULTS OF THE REVIEW STUDY AND DISCUSSION

From the L=45 studied papers, about 15 different models have been frequently identified. The first observation is that some models can be categorized under the same label such as: AR, MA, ARMA, ARIMA, and SARIMA models can be seen as a part of time series analysis (and forecasting) models. Therefore, they can be gathered into the label "Time Series models" [15]. Besides the huge number of papers presenting a load forecasting modeling in different areas of research, there are several general books and scientific theses available and published about modeling and forecasting electricity loads. Some of them only introduce the basic concepts of load forecasting modeling, while others cover all the advanced methodologies and modeling issues, let be by using statistical or artificial intelligence approaches. Table 1 introduces some of the typical books and theses, which dominate the field of load forecasting modeling.

Table 1: Some of available books and theses dominating at the field of load forecasting modeling

	Author	Title
Books	Rafal Weron [55]	**Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach**
	Soliman and Al-Kandari [29]	**Electrical Load Forecasting: Modeling and Model Construction**
	Hong et al. [20]	**Short-Term Load Forecasting by Artificial Intelligent Technologies**
	Adhikari and Agrawal [56]	**An Introductory Study on Time Series Modeling and Forecasting (the ELF partially covered)**
	Hyndman and Athanasopoulos [30]	**Forecasting: Principles and Practice (the ELF partially covered)**
Theses	Manish Singla [132]	**Electrical Load Forecasting Using Neural Networks**
	Jaime Buitrago [133]	**Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Multivariable Inputs**
	Mahmoud Shepero [134]	**Modeling And Forecasting The Load in The Future Electricity Grid: Spatial Electric Vehicle Load Modeling and Residential Load Forecasting**

It is vital also to understand the general trend of researchers' interest in search and investigate the time-progress of developed models of electricity load forecasting over time in order to improve the existing results and applications. Thus, Fig. 6 shows the rates and number of published papers, in terms of time coverage from 2003 to 2019 (considering L=45 papers), whereby we can notice that the trend of attention increases and the highest number of publications was recorded in 2017. The results reflect the importance of studying such a topic recently, in addition, the linear-forecast predicts a positive growing in this trend in the future.

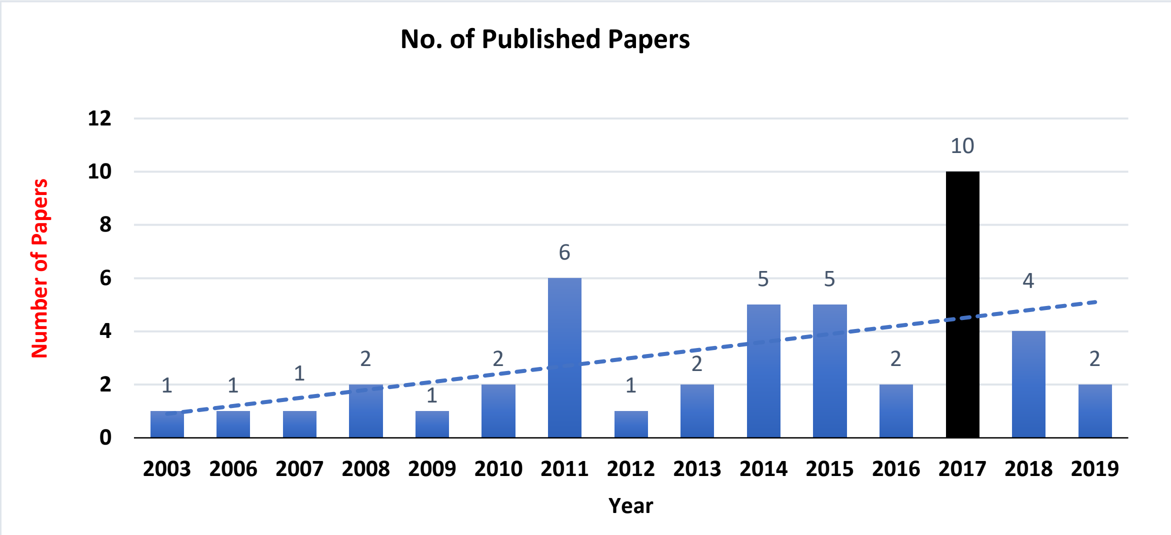


Figure 6: The Publication Pattern in the Field of Electric Load Forecasting Modeling (for K = 45 papers)

Fifteen different forecasting models have been identified and classified in this study. Indeed, it is interesting to study their distribution through all the M = 145 most relevant references in order to have a clear vision for the current trend in the forecasting models used. The distribution of the different analyzed forecasting models is illustrated in Fig. 7. A clear orientation is observed in the use of forecasting models. The ANN is the most widely used present in 27 papers, followed by the regression model present in 19 papers. Then the fuzzy logic and the SVM came after that in 15 papers. Moreover, the ARMA and ARIMA models have been hired in 13 and 12 papers respectively, and the rest of the models were used but by limited proportions. The relatively high quantity of regression, ANN, fuzzy logic and SVM models can be explained by their popularity in the research community. This observation strengthens their status as leading models in the field.

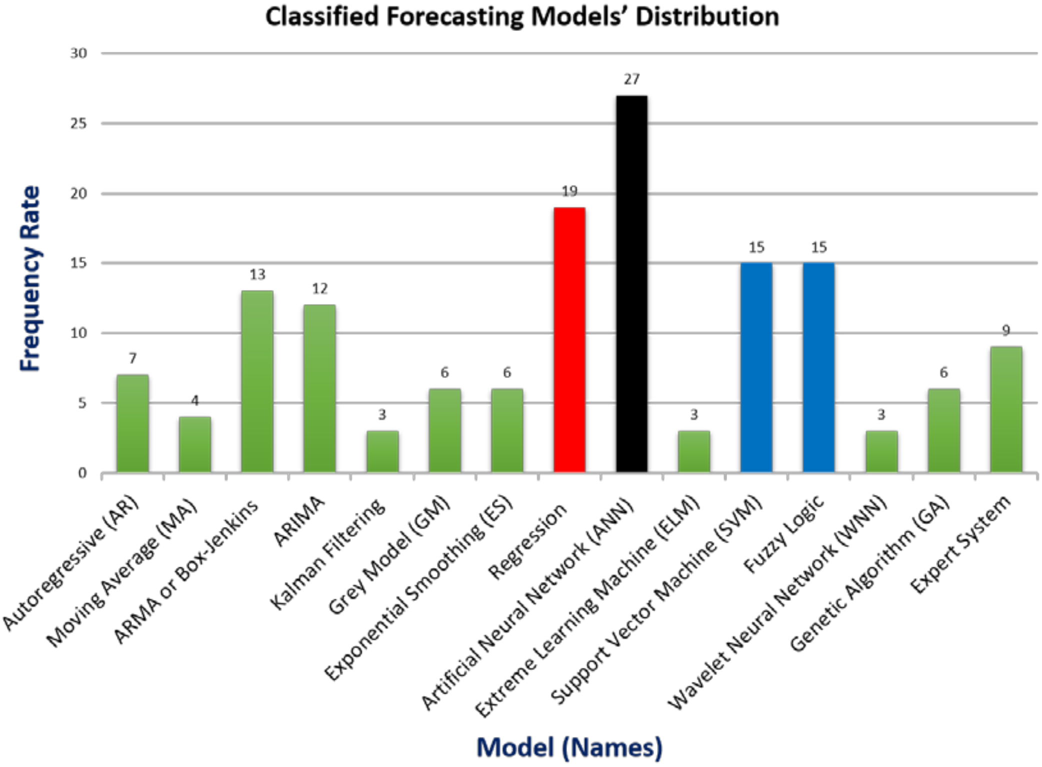


Figure 7: Distribution of Classified Forecasting Models (for M = 145 papers)

Another important characteristic is the horizon or time-term considered for the prediction. In fact, the time frame and solution that are chosen for any prediction will highly influence the results and the choice of a model over another. The timeframe of prediction is classified into four categories: very short term, short term, mid-term and long term. The forecasting horizon distribution through the different reviewed papers is shown in Table 2. The distribution in percentage is based on the number of papers in which the forecasting time frame is relevant or emphasized. The results reveal that **short-term and long-term predictions** have contributed to the highest percentage within the reviewed papers by **44.4% and 22.2%** respectively. In contrast, very short-term and mid-term predictions are not highly represented within the cases.

Table 2: The forecasting horizon distribution through the reviewed papers (for K = 45 papers)

Time Frame	Number of Papers (Journal & Conference)	Distribution Percentage
Very Short-Term	1	2.22%
Short-Term	**20**	**44.44%**
Mid-Term	5	11.11%
Long-Term	**10**	**22.22%**
None	9	20%
Total	**45**	---

In terms of geographical coverage, China is the most addressed and studied country through 4 studies followed by the USA and Turkey, then Australia and UAE, as shown in Fig. 8.

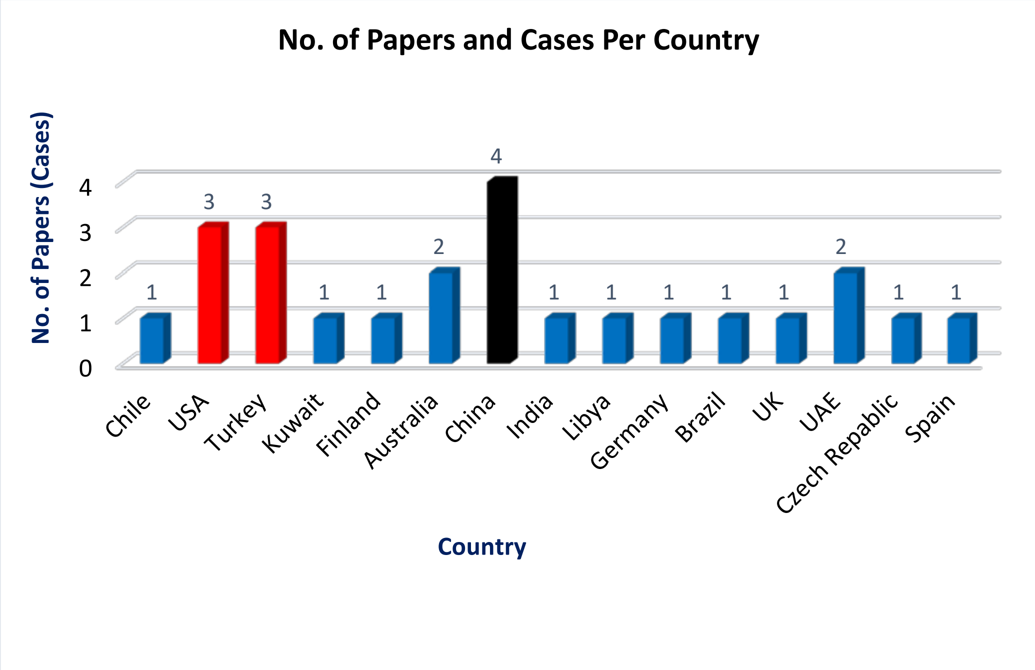


Figure 8: Coverage by Area for the Studied Countries (for K = 45 papers)

V. CONCLUSIONS AND FURTHER RESEARCH

In this paper, over 15 different forecasting models distributed into 45 most relevant scientific papers about the electric load forecasting have been more in-depth investigated. Several criteria have been checked and examined such as the scale of the project, the prediction horizon time frame, time resolution, inputs, outputs, data pre-processing, etc. The study also analyzed some patterns in the use of these models. Some of them are more appropriate and preferred for electric load forecasts such as regression analysis based models and artificial neural networks (ANN), which are the most utilized models in electricity predictions. In this scope, the artificial neural networks (ANN) models are mainly employed for short-term predictions where electricity and power consumption patterns are more complicated. On the other side, the regression models are still widely applied and efficient for long-term forecasting where periodicity and changes are less significant. Additionally, the fuzzy logic and support vector machine (SVM) models are present in a significant proportion of papers showing increasing attention thereof. Conversely, the statistical models (The Box-Jenkins models' family in particular) are not so dominant anymore as have been in the past, but their share still cannot be neglectable. Despite all these detected studies, the research gate is still wide open to apply and adapt a lot of novel combined models for electricity and power prediction. Moreover, certain magnified attention to study very short-term and mid-term load forecasting should have been additionally dedicated to fulfilling the detected gap in the field.

REFERENCES

[1] Y. Lin, H. Luo, D. Wang, H. Guo, and K. Zhu, "An Ensemble Model Based on Machine Learning Methods and Data Preprocessing for Short-Term Electric Load Forecasting," *Energies*, vol. 10, no. 1186, 2017.

[2] M. D. Reddy, "Load Forecasting using Linear Regression Analysis in Time series model for RGUKT, R.K. Valley Campus HT Feeder," *International Journal of Engineering Research & Technology (IJERT)*, vol. 6, no. 5, 2017.

[3] G. Nalcaci, A. Özmen, and G. W. Weber, "Long-term Load Forecasting: Models Based on MARS, ANN and LR methods," *Central European Journal of Operations Research (CEJOR), Springer-Verlag GmbH Germany*, vol. 27, no. 2019, pp. 1033–1049, 2018.

[4] E. Almeshaiei and H. Soltan, "A Methodology for Electric Power Load Forecasting," *Alexandria Engineering Journal*, vol. 50, no. 2011, pp. 137–144, 2011.

[5] J. Zhang, "Research on Power Load Forecasting Based on the Improved Elman Neural Network," *The Italian Association of Chemical Engineering (AIDIC)*, vol. 51, no. 2016, pp. 589-594, 2016.
[6] M. Y. Khamaira, A. S. Krzma, and A. M. Alnass, "Long Term Peak Load Forecasting for the Libyan Network," in *Conference for Engineering Sciences and Technology (CEST)*, 2018, vol. 1, pp. 185-193: AIJR Publisher.
[7] O. Demirel, A. Kakilli, and M. Tektas, "Electric Energy Load Forecasting Using ANFIS and ARMA Methods," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 25, no. 3, pp. 601-610, 2010.
[8] Q. Wang, Y. Wang, and L. Zhang, "Research on Post-Evaluation Model and Method of Electric Powe Demand Forecasting," presented at the *Chinese Control and Decision Conference (CCDC)*, China, 2008.
[9] X. Zhanga, J. Wanga, and K. Zhang, "Short-Term Electric Load Forecasting Based on Singular Spectrum Analysis and Support Vector Machine Optimized by Cuckoo Search Algorithm," *Electric Power Systems Research*, vol. 146, no. 2017, pp. 270–285, 2017.
[10] Ü. B. Filik, Ö. N. Gerek, and M. Kurban, "A Novel Modeling Approach for Hourly Forecasting of Long-Term Electric Energy Demand," *Energy Conversion and Management*, vol. 52, no. 2011, pp. 199–211, 2011.
[11] R. Gordillo-Orquera, L. M. Lopez-Ramos, S. Muñoz-Romero, P. Iglesias-Casarrubios, D. Arcos-Avilés, A. G. Marques, and J. L. Rojo-Álvarez, "Analyzing and Forecasting Electrical Load Consumption in Healthcare Buildings," *Energies*, vol. 11, no. 493, 2018.
[12] L. Friedrich and A. Afshari, "Short-Term Forecasting of the Abu Dhabi Electricity Load Using Multiple Weather Variables," presented at the *7th International Conference on Applied Energy (ICAE)*, 2015.
[13] N. Abu-Shikhah and F. Elkarmi, "Medium-Term Electric Load Forecasting Using Singular Value Decomposition," *Energy Conversion and Management*, vol. 36, no. 7, pp. 4259-4271, 2011.
[14] R. Wanga, J. Wangb, and Y. Xu, "A Novel Combined Model Based on Bybrid Optimization Algorithm for Electrical Load Forecasting," *Applied Soft Computing Journal*, vol. 82, no. 2019, p. 105548, 2019.
[15] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical Load Forecasting Models: A Critical Systematic Review," in *Sustainable Cities and Society*, ed, 2017.
[16] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, 2002.
[17] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustainable Cities and Society*, vol. 35, pp. 257–270, 2017, Art. no. Pii: s2210670717305899.
[18] A. K. Singh, Ibraheem, S. Khatoon, M. Muazzam, and D. K. Chaturvedi, "Load forecasting techniques and methodologies: A review," in *2012 2nd International Conference on Power, Control and Embedded Systems*, 2012, pp. 1–10.
[19] S. Fallah, M. Ganjkhani, S. Shamshirband, and K.-w. Chau, "Computational Intelligence on Short-Term Load Forecasting: A Methodological Overview," *Energies*, vol. 12, no. 3, p. 393, 2019, Art. no. PII: en12030393.
[20] W. C. Hong, M. W. Li, and G. F. Fan, *Short-Term Load Forecasting by Artificial Intelligent Technologies*. MDPI AG, 2019.
[21] I. A. b. W. A. Razak, S. b. Majid, M. S. b. M. Aras, and A. b. Ahmad, "Electricity Load Forecasting Using Data Mining Technique," ed: IntechOpen, 2012.
[22] F. Su, Y. Xu, and X. Tang, "Short-and mid-term load forecasting using machine learning models," in *2017 China International Electrical and Energy Conference (CIEEC)*, 2017, pp. 406–411.
[23] S. HemaChandra, V. Harish, C. R. Kumar, and V. Nagarjuna, "A Review of Long Term Electrical Load Forecasting Methods," (in English), *Artificial Intelligent Systems and Machine Learning*, vol. 4, no. 10, pp. 566–569, 2012.
[24] S. K. Panda, S. N. Mohanty, and A. K. Jagadev, "Long Term Electrical Load Forecasting: An Empirical Study across Techniques and Domains," *Indian Journal of Science and Technology*, vol. 10, no. 26, pp. 1–16, 2017.
[25] M. Jacob, C. Neves, and D. Vukadinović Greetham, "Short Term Load Forecasting," in *FORECASTING AND ASSESSING RISK OF INDIVIDUAL ELECTRICITY PEAKS*, vol. 33, M. N. C. V. G. D. Jacob, Ed. (Mathematics of Planet Earth, [S.l.]: SPRINGER NATURE, 2019, pp. 15–37.
[26] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley, 2007.
[27] B. Yildiz, J. I. Bilbao, and A. B. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 1104–1122, 2017, Art. no. Pii: s1364032117302265.
[28] E. A. Feinberg and D. Genethliou, "Load Forecasting," in *Applied Mathematics for Restructured Electric Power Systems.Optimization, Control, and Computational Intelligence*, J. H. Chow, F. F. Wu, and J. A. Momoh, Eds. Dordrecht: Springer-Verlag New York Inc, 2006, pp. 269-285.
[29] S. A.-h. Soliman and A. M. Al-Kandari, *Electrical Load Forecasting: Modeling and Model Construction*, 1st ed. Butterworth–Heineman, 2010.
[30] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. OTexts: Melbourne, Australia, 2018.
[31] R.M.Dawes, "Clinical versus Actuarial Prediction," *International Encyclopedia of the Social & Behavioral Sciences*, pp. 2048-2051, 2001. Elsevier Ltd.
[32] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: Forecasting and control / George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel*, 4th ed. ed. (Wiley series in probability and statistics). Oxford: Wiley, 2008.
[33] P. S. P. Cowpertwait and A. V. Metcalfe, *Introductory Time Series with R*. Springer New York, 2009.
[34] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control*. Wiley, 2013.
[35] Casals, Jose, Garcia-Hiernaux, Alfredo, Jerez, Miguel, Sotoca, Sonia, Trindade, and A. Alexandre, "State-Space Methods for Time Series Analysis: Theory, Applications and Software."
[36] J. J. F. Commandeur and S. J. Koopman, *An introduction to state space time series analysis* (Practical econometrics). Oxford; New York: Oxford University Press, 2007, pp. xiv, 174.
[37] J. Durbin and S. J. Koopman, *Time series analysis by state space methods*, 2nd ed. ed. (Oxford statistical science series, no. 38). Oxford: Oxford University Press, 2012, pp. xxi, 346.
[38] L. Huang, Y. Yang, H. Zhao, X. Wang, and H. Zheng, "Time series modeling and filtering method of electric power load stochastic noise," *Protection and Control of Modern Power Systems*, vol. 2, no. 1, p. 7, 2017, Art. no. Pii: 59.
[39] S. Markoulakis, "Short-term load forecasting based on the Kalman filter and the neural-fuzzy network (ANFIS)."
[40] M. Gaur and A. Majumdar, "One-Day-Ahead Load Forecasting using nonlinear Kalman filtering algorithms," 2016.

[41] J. Cheng, W. Xiong, and L. Ai, "Electric Load Forecasting Based on Improved Grey Neural Network," in *Recent advances in computer science and information engineering*, vol. 124, Z. Qian, Ed. (Lecture Notes in Electrical Engineering, no. 124-129) Heidelberg: Springer, 2012, pp. 651–655.
[42] C. Herui, B. Tao, and L. Yanzi, "Short-term Power Load Forecasting Based on Gray Theory," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 11, 2013.
[43] M. Jin, X. Zhou, Z. M. Zhang, and M. M. Tentzeris, "Short-term power load forecasting using grey correlation contest modeling," *Expert Systems with Applications*, vol. 39, no. 1, pp. 773–779, 2012, Art. no. Pii: s0957417411010347.
[44] J. A. S. Kelso, P. Érdi, K. J. Friston, H. Haken, J. Kacprzyk, J. Kurths, L. E. Reichl, P. Schuster, F. Schweitzer, D. Sornette, S. Liu, and Y. Lin, *Grey Systems* (no. 68). Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, p. 401.
[45] S. Liu and Y. Lin, *Grey systems: Theory and applications / Sifeng Liu and Yi Lin* (Understanding complex systems). Berlin: Springer Verlag, 2010.
[46] S. Liu, Y. Yang, and J. Forrest, *Grey data analysis: Methods, models and applications / Sifeng Liu, Yingjie Yang, Jeffrey Forrest* (Computational risk management). Singapore: Springer, 2017.
[47] Y. Lu, Y. Teng, and H. Wang, "Load Prediction in Power System with Grey Theory and its Diagnosis of Stabilization," *Electric Power Components and Systems*, vol. 47, no. 6-7, pp. 619–628, 2019.
[48] J. Mi, L. Fan, X. Duan, and Y. Qiu, "Short-Term Power Load Forecasting Method Based on Improved Exponential Smoothing Grey Model," *Mathematical Problems in Engineering*, vol. 2018, no. 1, pp. 1–11, 2018, Art. no. Pii: 3894723.
[49] T. Ozcan, T. Küçükdeniz, and F. H. Sezgin, "Comparative Analysis of Statistical, Machine Learning, and Grey Methods for Short-Term Electricity Load Forecasting," in *Nature inspired computing*, vol. 1, I. R. Management Association, Ed. Hershey PA: IGI Global, 2017, pp. 1161–1183.
[50] H. Zhao and S. Guo, "An optimized grey model for annual power load forecasting," *Energy*, vol. 107, pp. 272–286, 2016, Art. no. Pii: s0360544216304066.
[51] P. Ji, D. Xiong, P. Wang, and J. Chen, "A Study on Exponential Smoothing Model for Load Forecasting," presented at the *Asia-Pacific Power and Energy Engineering Conference (APPEEC), 2012.27-29 March 2012, Shanghai, China ; proceedings*, Piscataway, NJ, 2012. Available: http://ieeexplore.ieee.org/document/6307555/
[52] D. Dragan, T. Kramberger, and M. Intihar, "A comparison of Methods for Forecasting the Container Throughput in North Adriatic Ports," presented at the *IAME 2014*, Norfolk, 2014.
[53] D. Dragan, A. Keshavarzsaleh, T. Kramberger, B. Jereb, and M. Rosi, "Forecasting US Tourists' inflow to Slovenia by modified Holt-Winters Damped model: A case in the Tourism industry logistics and supply chains," *Logistics & Sustainable Transport*, vol. 10, no. 1, pp. 11–30, 2019.
[54] R. J. Hyndman, *Forecasting with exponential smoothing*. Berlin; London: Springer, 2008.
[55] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. England: John Wiley & Sons Ltd, 2006.
[56] R. Adhikari and R. K. Agrawal, *An Introductory Study on Time Series Modeling and Forecasting*. LAP Lambert Academic Publishing, 2013.
[57] J. G.Jetcheva, MostafaMajidpour, and Wei-PengChen, "Neural Network Model Ensembles for Building Level Electricity Load Forecasts," *Energy andBuildings*, vol. 84, no. 2014, pp. 214–223, 2014.
[58] M. Sarhani and A. E. Afia, "Electric Load Forecasting Using Hybrid Machine Learning Approach Incorporating Feature Selection," in *International Conference on Big Data Cloud and Applications*, Tetuan, Morocco, 2015.
[59] X. Wang, K. Smith-Miles, and R. Hyndman, "Rule Induction for Forecasting Method Selection: Meta-Learning the Characteristics of Univariate Time Series," *Neurocomputing*, vol. 72, no. 10-12, pp. 2581-2594, 2009.
[60] M. Intihar, T. Kramberger, and D. Dragan, "Container Throughput Forecasting Using Dynamic Factor Analysis and ARIMAX Model," *PROMET - Traffic&Transportation*, vol. 29, no. 5, pp. 529–542, 2017.
[61] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," University of North Carolina, Chapel Hill2004, vol. TR 95-041.
[62] E. Kayacan, B. Ulutas, and O. Kaynak, "Grey system theory-based models in time series prediction," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1784-1789, 2010.
[63] Y. Feng, "Study on Medium and Long Term Power Load Forecasting Based on Combination Forecasting Model," *Chemical Engineering Transactions*, vol. 51, no. 2015, pp. 859-864, 2015.
[64] E. Ostertagová and O. Ostertag, "The Simple Exponential Smoothing Model," presented at the *Modelling of Mechanical and Mechatronic Systems 2011: The 4th International conference*, Faculty of Mechanical engineering, Technical university of Košice, 2011.
[65] A. Chusyairi, R. N. S. Pelsri, and Bagio, "The Use of Exponential Smoothing Method to Predict Missing Service E-Report," presented at the *Information Systems and Electrical Engineering (ICITISEE): 2nd International Conferences on Information Technology*, 2017.
[66] M. A. Momani, W. H. Alrousan, and A. T. Alqudah, "Short-Term Load Forecasting Based on NARX and Radial Basis Neural Networks Approaches for the Jordanian Power Grid " *Jordan Journal of Electrical Engineering*, vol. 2, no. 1, pp. 81-93, 2016.
[67] L. C. M. d. Andrade, M. Oleskovicz, A. Q. Santos, D. V. Coury, and R. A. S. Fernandes, "Very short-term load forecasting based on NARX recurrent neural networks," in *2014 IEEE PES general meeting*Piscataway, NJ: IEEE, 2014, pp. 1–5.
[68] J. Buitrago and S. Asfour, "Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Vector Inputs," *Energies*, vol. 10, no. 1, p. 40, 2017, Art. no. PII: en10010040.
[69] W. X. Jiatang Cheng and L. Ai, "LNEE 124 - Electric Load Forecasting Based on Improved Grey Neural Network."
[70] H. Li, Y. Zhu, J. Hu, and Z. Li, "A localized NARX Neural Network model for Short-term load forecasting based upon Self-Organizing Mapping," in *2017 IEEE 3rd International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017 - ECCE Asia)*, I. I. F. E. E. Conference, Ed. [Piscataway, NJ]: IEEE, 2017, pp. 749–754.
[71] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489-501, 2006.
[72] Y. Chen, M. Kloft, Y. Yang, C. Li, and L. Li, "Mixed kernel based extreme learning machine for electric load forecasting," *Neurocomputing*, vol. 312, pp. 90-106, 2018.

[73] S. K. Dash and D. Patel, "Short-term electric load forecasting using Extreme Learning Machine - a case study of Indian power market," presented at the *2015 IEEE Power, Communication and Information Technology Conference (PCITC.15-17 October, 2015, Siksha 'O' Anusandhan University, Bhubaneswar, India : PCITC-2015 proceedings*, [Piscataway, NJ], 2015. Available: http://ieeexplore.ieee.org/document/7438135/
[74] Ö. F. Ertugrul, "Forecasting electricity load by a novel recurrent extreme learning machines approach," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 429-435, 2016.
[75] C. H. Weng, W. Ting, L. Xueyong, and R. Weerasinghe, "Research on short-term electric load forecasting based on extreme learning machine," *E3S Web of Conferences*, vol. 53, p. 02009, 2018.
[76] P. J. Garcia-Laencina, "Improving Predictions Using Linear Combination Of Multiple Extreme Learning Machines," *Information Technology And Control*, vol. 42, no. 1, 2013.
[77] M. A. A. Albadr and S. Tiun, "Extreme Learning Machine: A Review " *International Journal of Applied Engineering Research*, vol. 12, no. 14, pp. 4610-4623, 2017.
[78] W. Ting and L. Xueyong, "Research on short-term electric load forecasting based on extreme learning machine," *E3S Web of Conferences*, vol. 53, p. 02009, 2018.
[79] Y. Wei, H. Huang, B. Chen, B. Zheng, and Y. Wang, "Application of Extreme Learning Machine for Predicting Chlorophyll-a Concentration Inartificial Upwelling Processes," *Mathematical Problems in Engineering*, vol. 2019, pp. 1-11, 2019.
[80] Z. Yang, T. Zhang, J. Lu, Y. Su, D. Zhang, and Y. Duan, "Extreme learning machines for regression based on V-matrix method," *Cogn Neurodyn*, vol. 11, no. 5, pp. 453-465, Oct 2017.
[81] Y. Fu, Z. Li, H. Zhang, and P. Xu, "Using Support Vector Machine to Predict Next Day Electricity Load of Public Buildings with Sub-metering Devices," *Procedia Engineering*, vol. 121, pp. 1016-1022, 2015.
[82] W.-C. Hong, "Electric load forecasting by support vector model," *Applied Mathematical Modelling*, vol. 33, no. 5, pp. 2444-2454, 2009.
[83] Z. Hu, Y. Bao, and T. Xiong, "Electricity load forecasting using support vector regression with memetic algorithms," *ScientificWorldJournal*, vol. 2013, p. 292575, 2013.
[84] S. Qiang and Y. Pu, "Short-term power load forecasting based on support vector machine and particle swarm optimization," *Journal of Algorithms & Computational Technology*, vol. 13, p. 174830181879706, 2018.
[85] S. Maldonado, A. González, and S. Crone, "Automatic time series analysis for electric load forecasting via support vector regression," *Applied Soft Computing*, vol. 83, p. 105616, 2019.
[86] N. Cristianini and J. Shawe-Taylor, *An introduction to Support Vector Machines*. New York: Cambridge University Press, 2000.
[87] N. Deng, Y. Tian, and C. Zhang, *Support vector machines* (Chapman & Hall/CRC data mining and knowledge discovery series). Boca Raton: CRC Press, 2013.
[88] I. Steinwart and A. Christmann, *Support vector machines* (Information science and statistics). New York: Springer, 2008.
[89] J. A. K. Suykens, *Least squares support vector machines*. New Jersey; London: World Scientific H1 - British Library H2 - DSC m03/18809, 2002.
[90] A. Abraham and S. Das, *Computational Intelligence in Power Engineering*. Springer Berlin Heidelberg, 2010.
[91] P. L. Anderson, *Business Economics and Finance with MATLAB, GIS, and Simulation Models*. CRC Press, 2004.
[92] M. F. Azeem, *Fuzzy Inference System: Theory and Applications*. IntechOpen, 2012.
[93] J. H. Chow, F. F. Wu, and J. A. Momoh, *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*. Springer US, 2004.
[94] S. K. Halgamuge and L. Wang, *Computational Intelligence for Modelling and Prediction*. Springer Berlin Heidelberg, 2005.
[95] R. Jensen and Q. Shen, *Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches*. Wiley, 2008.
[96] S. Kalogirou, *Artificial Intelligence in Energy and Renewable Energy Systems*. Nova Science Publishers, 2007.
[97] A. Konar and D. Bhattacharya, *Time-Series Prediction and Applications: A Machine Intelligence Approach*. Springer International Publishing, 2017.
[98] E. Ogliari and S. Leva, *Computational Intelligence in Photovoltaic Systems*. Mdpi AG, 2019.
[99] A. K. Palit and D. Popovic, *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*. Springer London, 2006.
[100] M. Paulescu, E. Paulescu, P. Gravila, and V. Badescu, *Weather Modeling and Forecasting of PV Systems Operation*. Springer London, 2012.
[101] W. Pedrycz and S. M. Chen, *Time Series Analysis, Modeling and Applications: A Computational Intelligence Perspective*. Springer Berlin Heidelberg, 2012.
[102] S. A. Soliman and A. M. Al-Kandari, *Electrical Load Forecasting: Modeling and Model Construction*. Elsevier Science, 2010.
[103] Y. H. Song, *Modern Optimisation Techniques in Power Systems*. Springer Netherlands, 1999.
[104] M. Sudha, *Applied Computational Intelligence*. Educreation Publishing, 2019.
[105] G. Tayfur, *Soft Computing in Water Resources Engineering: Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms*. WIT Press, 2014.
[106] K. E. Voges and N. Pope, *Business Applications and Computational Intelligence*. Idea Group Publishing, 2006.
[107] J. Wang, *Business Intelligence in Economic Forecasting: Technologies and Techniques: Technologies and Techniques*. Information Science Reference, 2010.
[108] L. Wang, C. Singh, and A. Kusiak, *Wind Power Systems: Applications of Computational Intelligence*. Springer Berlin Heidelberg, 2010.
[109] P. P. Wang, *Computational Intelligence in Economics and Finance*. Springer Berlin Heidelberg, 2013.
[110] A. T. Ali, E. B. Tayeb, and Z. M. Shamseldin, "Short term electrical load forecasting using fuzzy logic," *International Journal Of Advancement In Engineering Technology, Management and Applied Science (IJAETMAS)*, vol. 3, 2016.
[111] D. Ali, M. Yohanna, M. I. Puwu, and B. M. Garkida, "Long-term load forecast modelling using a fuzzy logic approach," *Pacific Science Review A: Natural Science and Engineering*, vol. 18, no. 2, pp. 123-127, 2016.
[112] M. Faysal, M. J. Islam, M. M. Murad, M. I. Islam, and M. R. Amin, "Electrical Load Forecasting Using Fuzzy System," *Journal of Computer and Communications*, vol. 07, no. 09, pp. 27-37, 2019.

[113] M. K. Singla and S. Hans, "Load Forecasting using Fuzzy Logic Tool Box," *Global Research and Development Journal for Engineering*, vol. 38, pp. 12-19, 2018.
[114] L. Yao, Y.-l. Jiang, and J. Xiao, "Short-Term Power Load Forecasting by Interval Type-2 Fuzzy Logic System," in *Information Computing and Applications*, Berlin, Heidelberg, 2011, pp. 575-582: Springer Berlin Heidelberg.
[115] Z. Ismail and R. Mansor, "Fuzzy Logic Approach for Forecasting Half-hourly Electricity Load Demand," *Fuzzy Logic Approach for Forecasting Half-hourly Electricity Load Demand*, 09/14 2011.
[116] J. Jamaaluddin, D. Hadidjaja, I. Sulistiyowati, E. A. Suprayitno, I. Anshory, and S. Syahrorini, "Very short term load forecasting peak load time using fuzzy logic," *IOP Conference Series: Materials Science and Engineering*, vol. 403, p. 012070, 2018.
[117] J. Kaur and Y. S. Brar, "Short term load forecasting using fuzzy logic of 220KV transmission line," *Int. J. Eng. Res. Technol*, vol. 3, no. 2278, p. e0181, 2014.
[118] A. Laouafi, M. Mordjaoui, and T. E. Boukelia, "An adaptive neuro-fuzzy inference system-based approach for daily load curve prediction," *Journal of Energy Systems*, pp. 115-126, 2018.
[119] S. K. Patel and S. Sharma, "A Review of very Short-Term Load Forecasting (STLF) using Wavelet Neural Networks," *International Journal of Science, Engineering and Technology Research* vol. 4, no. 2, 2015.
[120] M. Mitchell, *An introduction to genetic algorithms*, 7. print ed. (A Bradford book). Cambridge, Mass., 2001, p. 209.
[121] R. L. Haupt and S. E. Haupt, *Practical genetic algorithms*, 2nd ed. ed. Hoboken, N.J.; Chichester: Wiley-Interscience, 2004.
[122] J. Carr, "An Introduction to Genetic Algorithms," 2014.
[123] C.-C. Hsu, C.-H. Wu, S.-C. Chen, and K.-L. Peng, "Dynamically Optimizing Parameters in Support Vector Regression: An Application of Electricity Load Forecasting," presented at the *System Sciences, the 39th Annual Hawaii International Conference*, 2006.
[124] D. Beasley, D. R. Bull, and R. R. Martin, "An Overview of Genetic Algorithms : Part 1, Fundamentals," *University Computing*, vol. 15, no. 2, pp. 56-69, 1993.
[125] Y. K. Al-Douri, H. Al-Chalabi, and J. Lundberg, "Time Series Forecasting using Genetic Algorithm," in *The Twelfth International Conference on Advanced Engineering Computing and Applications in Sciences*, 2018.
[126] R. R. B. de Aquino, O. N. Neto, M. M. S. Lira, A. A. Ferreira, and K. F. Santos, "Using Genetic Algorithm to Develop a Neural-Network-Based Load Forecasting," in *Artificial Neural Networks – ICANN 2007*, Berlin, Heidelberg, 2007, pp. 738-747: Springer Berlin Heidelberg.
[127] A. Gupta and P. K. Sarangi, "Electrical load forecasting using genetic algorithm based back-propagation method," *ARPN Journal of Engineering and Applied Sciences*, vol. 7, no. 8, pp. 1017-1020, 2012.
[128] G. M. Khan, F. Zafari, and S. A. Mahmud, "Very Short Term Load Forecasting Using Cartesian Genetic Programming Evolved Recurrent Neural Networks (CGPRNN)," in *2013 12th International Conference on Machine Learning and Applications*, 2013, vol. 2, pp. 152-155.
[129] F. Li and X. Zhao, "The application of genetic algorithm in power short-term load forecasting," in *2012 International Conference on Image, Vision and Computing (ICIVC 2012)*, 2012.
[130] J. I. Silva-Ortegaa, B. Cervantes-Bolivarb, I. A. Isaac-Millanc, Y. Cardenas-Escorciab, and G. Valencia-Ochoad, "Demand energy forecasting using genetic algorithm to guarantee safety on electrical transportation system," *CHEMICAL ENGINEERING*, vol. 67, 2018.
[131] H. Verdejo, A. Awerkin, C. Becker, and G. Olguin, "Statistic Linear Parametric Techniques for Residential Electric Energy Demand Forecasting: A Review and An Implementation to Chile," *Renewable and Sustainable Energy Reviews*, vol. 74, no. 2017, pp. 512–521, 2017.
[132] M. Singla, "Load Forecasting Using Artificial Neural Network," Thapar Institute, 2018.
[133] J. Buitrago, "Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Multivariable Inputs," *Open Access Dissertations*, 2017.
[134] M. Shepero, "Modeling and forecasting the load in the future electricity grid : Spatial electric vehicle load modeling and residential load forecasting," ed: Uppsala universitet, 2018.

AUTHORS

A. Mahmoud A. Hammad, M.Sc., is Assistant Lecturer at the Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt.

B. Borut Jereb, Ph.D., is Full Professor and the Vice-Dean at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: borut.jereb@um.si).

C. Bojan Rosi, Ph.D., is Full Professor and the Dean at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: bojan.rosi@um.si).

D. Dejan Dragan, Ph.D., is an Associate Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: dejan.dragan@um.si).

Manuscript received by 23 December 2019.

Published as submitted by the author(s).