IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁙ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11131

Methods and Models for Electric Load Forecasting: A Comprehensive Review

Prof. Vishal V. Mehtre[1], Ms. Ayushi Agarwal[2]

Department of Electrical Engineering, Bharati Vidyapeeth (Deemed to be University)

College of Engineering, Pune, Maharashtra, India[1,2]

Abstract: Electric Load Forecasting (ELF) is an indispensable interaction in the preparation of the power business and assumes a significant part in electric limit booking and power frameworks the board, subsequently, it has drawn in expanding scholarly interest. Consequently, the exactness of electric burden anticipating has extraordinary significance for energy creating limit planning and power framework the board. This paper presents an audit of determining techniques and models for power load. Around 45 scholarly papers have been utilized for the correlation in view of indicated models, for example, time period, inputs, yields, the size of the venture, and worth. The audit uncovers that notwithstanding the overall straightforwardness of all evaluated models, relapse examination is still broadly utilized and effective for long haul determining. With respect to transient forecasts, AI or man-made consciousness-based models like Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Fuzzy rationale are leaning toward.

Watchwords - Electric Load Forecasting; Modeling power loads; Methods and models of anticipating.

INTRODUCTION

Electricity is a spotless and productive wellspring of energy, which assumes an indispensable part in our everyday existence. The meaning of power has been expanding radically as of late, and accordingly, it has a fundamental subject in research. Also, electric power is more reasonable and effective for the necessity of climate - well-disposed society contrasted and other conventional wellsprings of energy like gaseous petrol, coal, and oil. Be that as it may, power as an item has various attributes contrasted with material items since it can't be put away in mass as it ought to be produced when it is requested. Additionally, the interest example of power is mind-boggling because of the liberation of power markets, for example, power oversupply and deficiency, which could prompt off-base anticipating and causes critical monetary misfortune. In addition, the worldwide power request is relied upon to duplicate with expanding populace and expectations for everyday comforts improvement. Additionally, economies extend as well as utilizing high-power electrical apparatuses and creating innovation like brilliant frameworks, electric vehicles, and sustainable power creation. This multitude of variables makes it hard to deal with the power framework. Subsequently, it is important to foresee the requirements/heaps of power ahead of time and prior to settling on its age.

Electric burden (EL)/request determining is an imperative interaction in the preparation of the power business also assumes an urgent part in the activity of the electric power framework. The electric power load figure is profoundly connected with the economy's turn of events, and it is additionally connected with public safety and the everyday activity of society. In this way, the exactness of electric burden anticipating has extraordinary significance for energy creating limit planning and power framework the executives, as these precise conjectures lead to significant reserve funds in working and support costs, and right choices for the future turn of events. Besides, electric power load gauging addresses the underlying advance in creating a group of people yet to come, transmission, and dissemination offices. In any case, the exactness of electric burden determining (ELF) can't regularly satisfy our ideal outcome since it is impacted by different dubious and wild factors like the monetary turn of events, human social exercises, country approaches, and environmental change.

So far, there is no precise standard for classifying the range of load forecasts. However, some authors have divided load forecasting in terms of the prediction duration into three categories short-term forecasts, medium -term forecasts, and long-term forecasts. Other researchers go for classifying load forecasting into four groups 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 forecastis 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)**:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⁜ Vol. 11, Issue 1, January 2022
DOI: 10.17148/IJARCCE.2022.11131

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)**:

range s 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 aneasy task, and none of them can be generalized for all demand patterns. Based on,electric load forecasting models can be divided into two types:

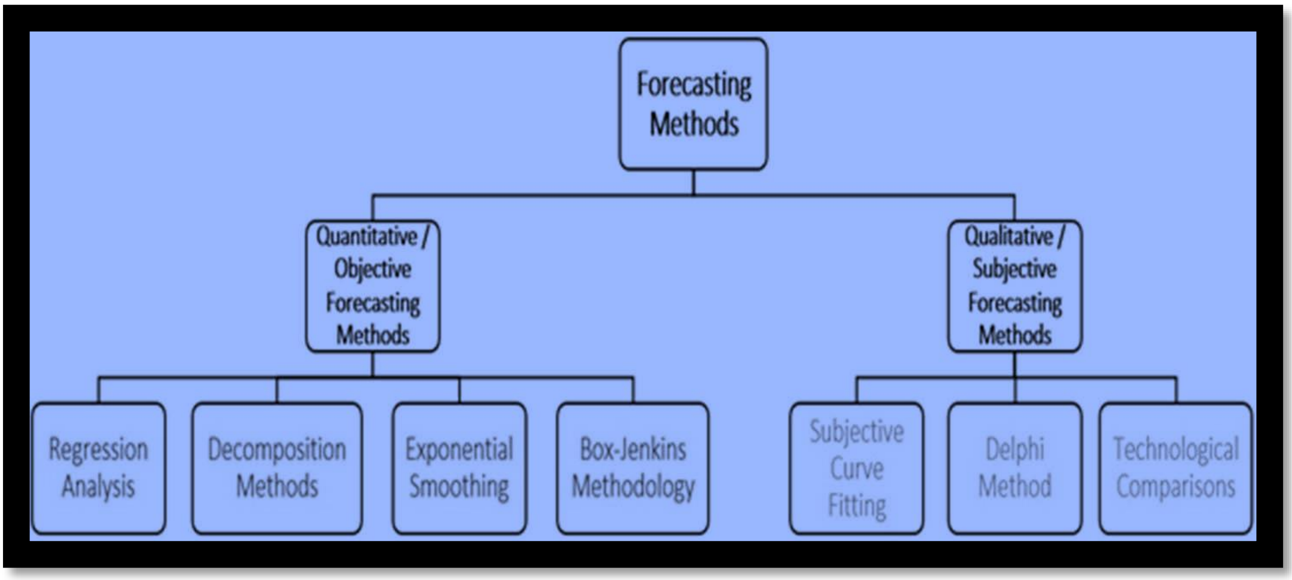
- **multi-factor forecasting methods**, and
- **time series forecasting methods**.

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, **thetime series forecasting method** is depending more on the historical series. Accordingly, lots ofresearchers 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 bedivided into three subcategories.
statistical models, machine learning models, hybrid models.

THE LITERATURE REVIEW

In this section, the most widely used methods and models in the area of electric load forecastingare briefly discussed by reviewing the relevant previous works. According to Feinberg, themajority 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 aresignificant: 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



Electric load (EL)/request determining is a fundamental cycle in the preparation of the power business what's more, assumes an urgent part in the activity of the electric power framework. The electric power load gauge is profoundly connected with the economy's turn of events, and it is likewise connected with public safety and the day-by-day activity of society. In this manner, the exactness of electric burden determining has extraordinary significance for energy creating limit booking and power framework the executives, as these precise conjectures lead to significant reserve funds in working and upkeep costs, and right choices for the future turn of events. Moreover, electric power load gauging addresses the underlying advance in creating a group of people yet to come, transmission, and dissemination offices. Notwithstanding, the precision of ELF can't regularly satisfy our ideal outcome since it is impacted by different dubious and wild factors like the financial turn of events, human social exercises, country approaches, and environmental change.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11131

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.

B.) Forecasting Models

The early electrical load forecasting models were almost entirely limited to traditional statisticalmethods, 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. This section defines anddescribes the most commonly used load forecasting models, whether traditional or modern intelligent models. It includes only some of the most commonly used clusters of the forecasting models. In the sequel, the discussion aboutthe 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 intelligenceapproaches.**

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 amathematical 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. In the sequel, the following models are going to be brieflydiscussed:

- **Box-Jenkins basic models (AR, MA, ARMA, ARIMA, ARMAX, and ARIMAX).**
- **Kalman Filtering Algorithm s in the State space**
- **Grey models.**
- **Exponential Smoothing.**

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 Auto Regressive (AR) model can be used to forecast future load values. A pth-order autoregressive, AR(p), model is defined as:

$$y_t - \sum_{i=1}^{p} \phi_i \, y_{t-i} = \varepsilon_t \,, \quad (1)$$

where: $\phi_1, \phi_2, \ldots, \phi_p$ are the unknown AR coefficient s, 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.

Moving Average (MA) Model

The moving average model that mimics the behavior of the moving average process, is a linearregression 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 a s 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:

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

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 MAprocess 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.

Autoregressive Moving Average (ARMA) Model

The Autoregressive Moving Average has been introduced in 1970 by George Box and Gwilym Jenkins. 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

IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⸎ Vol. 11, Issue 1, January 2022
DOI: 10.17148/IJARCCE.2022.11131

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

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

ARMA models have been a popular choice and extensively applied to load forecasting researchesdue to their relative simplicity and effectiveness.

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 19 76 with a purpose to include the case of non-stationarity as well. The ARIMA Box– Jenkins models have three types of parameters: th e autoregressive parameters $(\emptyset_1, \emptyset_2, \dots, \emptyset_p)$, the moving average parameters $(\theta_1, \dots, \theta_q)$, 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 usingthe 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 $(p, d, q) \times (P, D, Q)_s$, where s refers to the number of periods per season and P, D and Q are the seasonal equivalents of p, 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.

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. 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 filteris also a potent tool when it com es to controlling the noisy systems, in which the electric power systems undoubtedly can be assigned. According to Gaur and his colleagues, the mainelements 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 variation s 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 schedule s, 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 priorlyknown 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 utilizations.

The factors mentioned above can be injected as inputs into the Kalman filter. Since it is extremely difficult to deal with the complex

DOI: 10.17148/IJARCCE.2022.11131

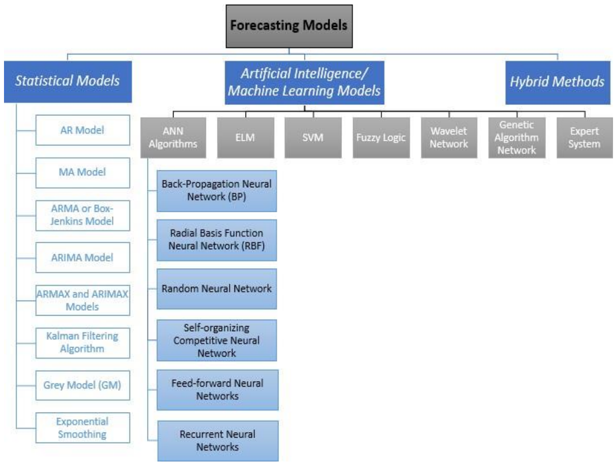
inputs (c., d., e.), the weather and time factors are usually included in the KF only. The KF is usually looked through the eyes of discrete -time linear dynamic system of the hidden states of the appliances, as well as the observation vector
The KF's mechanism works in a two-step
process, that is, the predictor step (PS), and the corrector step (CS). In the PS, the KF estimates thecurrent 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.

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 in telligence -based techniques provide a promising and attractive alternative.

The Extreme Learning Machines (ELM)

The Extreme learning machines represent the special class of the FF ANNs and are appropriatefor 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. 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. Theoutput 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, the ELM networks are capable of producing good generalization performance and can learn eventhousands of times quicker than the competitive networks, which are trained usingbackpropagation. It is also detected from the literature that ELM models can outperform even thesupport 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. There have also been successful attempts to improve the basic ELMscheme, e.g., Ertugrul has applied a novel recurrent ELM approach for the purpose of ELF forecasting, Garcia-Laencina has deployed a mechanism to improve the forecasting byconducting a linear combination of multiple ELM machines and demonstrated the performance on three engineering problems, etc.



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 studyhas been applied for the conventional review without the systema tic review as an initial step for discovering the general idea s 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

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⋇ Vol. 11, Issue 1, January 2022

DOI: 10.17148/IJARCCE.2022.11131

energy", "models and methods", and "load forecasting" in English only.Searching based on these keywords has addressed the online database Web of Science (WOS), themost 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 associate d 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 constitute d the basis of this review study. Among the selected 52 studies, 30 are article s, 15 are conference papers (together L = 45 papers), four are books, and three are theses.. All of them have been reviewed in depth.

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 thedifferent reviewed papers is shown in Table 2. The distribution in percentage is based on the number of paper s 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 4studies followed by the USA and Turkey, then Australia and UAE

CONCLUSIONS

In this paper, north of 15 different estimating models disseminated into 4 5 most applicable logical papers about the ELF have been more inside and out exploring. A few measures have been checked and inspected, for example, the size of the venture, the prediction skyline time span, time goal, inputs, yields, information pre-handling, and so on the concentrate additionally investigated a few examples in the utilization of these models. Some of them are more suitable and liked for electric burden conjectures, for example, relapse examination-based models and fake neural organizations (ANN), which are the most used models in electricity forecasts. To this degree, the counterfeit neural organizations (ANN) models are for the most part utilized for short-term expectations where power and power utilization designs are more convoluted. On the opposite side, the relapse models are still generally applied and proficient for long-term determining where periodicity and changes are less significant. Additionally, the fluffy rationale and backing vector machine (SVM) models are available in a huge extent of papers showing expanding consideration thereof. Conversely, the measurable models (The Box - Jenkins models' family specifically) are not so predominant any longer as have been before, yet their portion actually can't be neglectable. In spite of every one of these detected studies, the research entryway is still totally open to applying and adjusting a lot of novels joined models for power and power prediction. Moreover, certain amplified considerations regarding concentrating on the extremely present moment and mid-term load determining ought to have been also devoted to satisfying the recognized hole in the field.

