Meta-Analysis of Electric Load Forecasting for Power Systems

JSPM
JOURNAL OF STRATEGY & PERFORMANCE MANAGEMENT

Author:
Malik Usman Dilawar
Department of Electrical Engineering, Sir Syed CASE Institute of Technology, Islamabad

Abstract

Load forecasting is one of the challenges being faced by the power system industry. With the ever-growing consumer's demand, power generating companies are struggling to provide un-interrupted power supply to the users. Over the past few decades, introduction of smart grids and power deregulation has changed the dynamics of load forecasting. Most of the research is now focused towards short term load forecasting (STLF) involving an hour to a week's time forecasting. Various techniques are being used for accurately predicting the electric load. However, gold standards are yet to be defined mainly because of the variety, non-linearity and un-predictive form of the subject. In this study critical review of 25 publications has been carried out to find the most efficient method being used for STLF. On the basis of two parameters i.e., most wide acceptance in research community and minimum Mean Absolute Percentage Error (MAPE) results of proposed methods, it has been found that Artificial Neural Networks (ANN) and Auto Regressive Moving Average (ARMA) models perform better for forecasting the electric load forecasting than other methods.

Keywords: Electric load forecasting, Power generation, Power load, Modelling electricity loads, Long term/ Short term forecasting, Artificial intelligence-based forecasting.

Introduction

Electric load forecasting (ELF) has been a prime area of concern since the advent of electricity. The predictions about future load helps power utility companies in planning to meet the power generation with consumer's demands. ELF is also one of the significant factors for regulatory bodies, industries, trading and insurance companies (Hammad, 2020). With the technological advancement integration of smart devices in various technical fields has become a norm and power industry is not an exception. Additionally, inclination towards renewable energies, due to global warming issues, resulted into the introduction of smart equipment in power generation and grid systems. Same has resulted in availability of digitized data which on the other hand became helpful for analysis and future prediction (Weron, 2014). The consumer's demand of electricity is increasing day by day, as the world has moved towards an automated version of almost everything. Traditional power generating companies are facing challenges to meet user demands and their returns on investments are declining. A strong change in the power sector is observed during 1990's with the process of deregulation and market competition (Leung et al., 2019). On the other hand, development of smart electronics devices has gained popularity in generating power more efficiently. The volatility of electricity is adamant with the fact that it has to be provided promptly. Huge amount of electricity cannot be stored hence equating generation with the user demands is a tough task. ELF has thus emerged as a vibrant field for scientific community. An accurate load prediction enables decision making by the power operators. Power industry thus invested a lot in this field to compete in the market and to avoid burning extra fuel or running machinery to generate abundant amount of electricity.

ELF is generally categorized in long, medium- and short-term forecasting on the temporal basis. Though no standard categorization has been laid so far and all of them are interconnected in the broader perspective. Long Term Load Forecasting (LTLF) - forecasting for 3 years or more. For less than 3-year time period, it is termed as medium forecasting (MTLF). Finally, in short term (STLF), forecasting is carried out from an hour/ half-hour to a week times (Kuster et al., 2017). With the growing renewable power generation systems, introduction of smart grid systems and trend of privatizations, the short term and very short-term forecasting have gained popularity. In this study critical review will be carried out to look for the best technique being used for electric load forecasting. Rest of the paper is divided in four major parts. In the first part literature review is carried out, the next one describes the research methodology applied in this study. Findings with comparative analysis and results is explained in next part. Finally, discussion is carried out before concluding the study.

Literature Review

Calculation of load is one of the significant factors for power companies. All the operations and planning of power generation, transmission, maintenance etc. are based upon the future load value. The forecasting helps in decision making as well as reducing the risk of non-availability of power. Several conventional methods of forecasting are already in practice. Over the period of time various techniques have been researched to improve the load forecasting.

The qualitative methods forecast based upon the opinions and discussion with domain experts. These methods are employed when historical data is not available for a forthcoming event. Estimates are generally vague and can lead to a blackout. The Quantitative techniques involve Time Series Analysis and Econometric Analysis. In time series a variable of interest is defined such that its value is estimated relying on the relevant historical data. Base band model, Trend model, Linear Regression models are few examples. Econometric analysis considers the drivers such as business index, weather index etc. such that they further lead to estimate demand requirements. Recently, Artificial Intelligence (AI) has out-performed the conventional methods in the fields where non-linear and complex data is involved. The non-linear demands, transmission losses, climate factors etc. and their relationships have made the load forecasting a potential field for AI techniques. Artificial Neural Networks (ANN), Support Vector Machines (SVM), Genetic Algorithm, Fuzzy Logic, Self-Organizing Maps, Extreme Learning Machines are few other AI techniques that can also be employed in load forecasting. Various reviews and analysis done on the subject are consulted to develop a comprehensive meta-analysis approach for this study. The methods followed by the authors are depicted in the Table 1.

Table 1
Previous Reviews on ELF

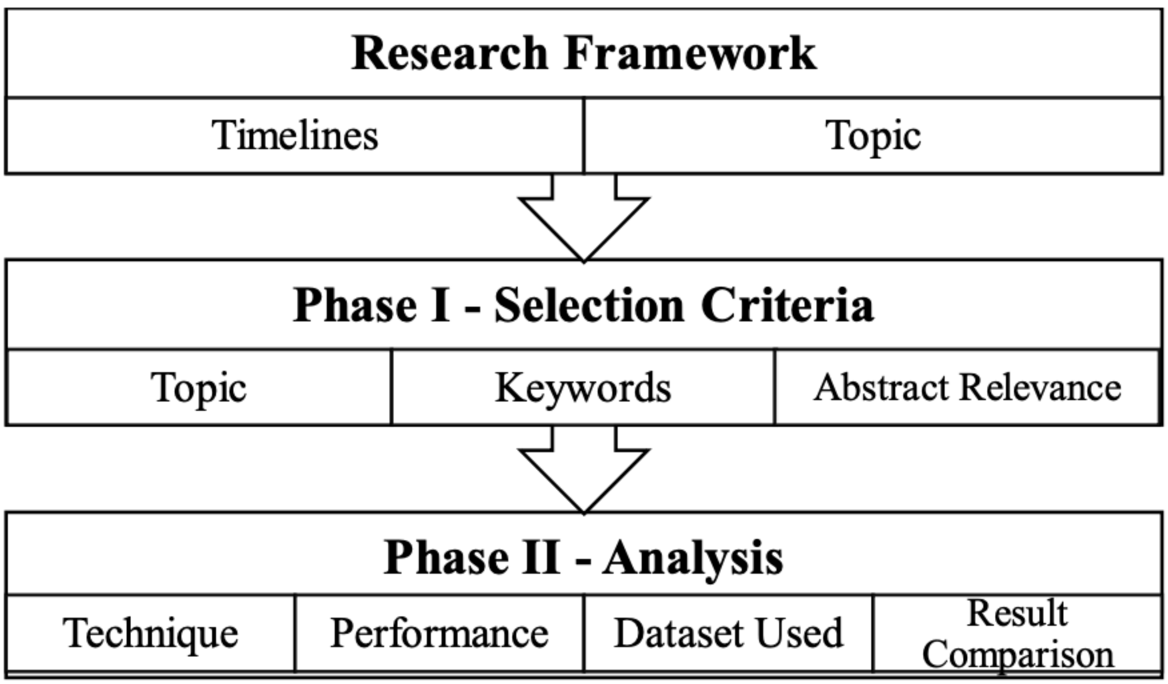
Ser	Review	Method
1.	Zhao & Tang, 2016	MAPE Percentages
2.	Zor et al., 2017	
3.	Kuster et al., 2017	Commonly used by Expert Community
4.	Hammad et al. 2020	
5.	Almalaq & Edwards, 2017	RMSE
6.	Upadhaya et a., 2019	Data and Error measured (MAPE, RMSE)

Research Approach

Research is an ongoing process where methods and theories are developed and supported by logics and proofs. Meta-analysis is a common field of almost all research disciplines. Its main objective is to combine published methods and theories of one category and compare them with that of another, in a systematic manner to reach some conclusion (Mohammed et al., 2017). It consists of five basic steps involving finding relevant studies on the subject, developing consistent criteria for comparison, recording relevant information from the study as per the criteria, analyzing information to compile them in broad contours and finally drawing conclusion basing upon these findings (Neuman, 2007). Aim of this study is to review the academic literature to explore the most efficient methods being used for STLF. Critical analysis is carried out to analyze the dynamics and performance of various methods and techniques employed.

The general framework of the study comprises of two phases. In the first phase research papers and articles are searched in top most databases on internet centered upon specific keywords. In the next phase developed methods and their results are analyzed statistically. General framework of this research is shown in Figure 1.

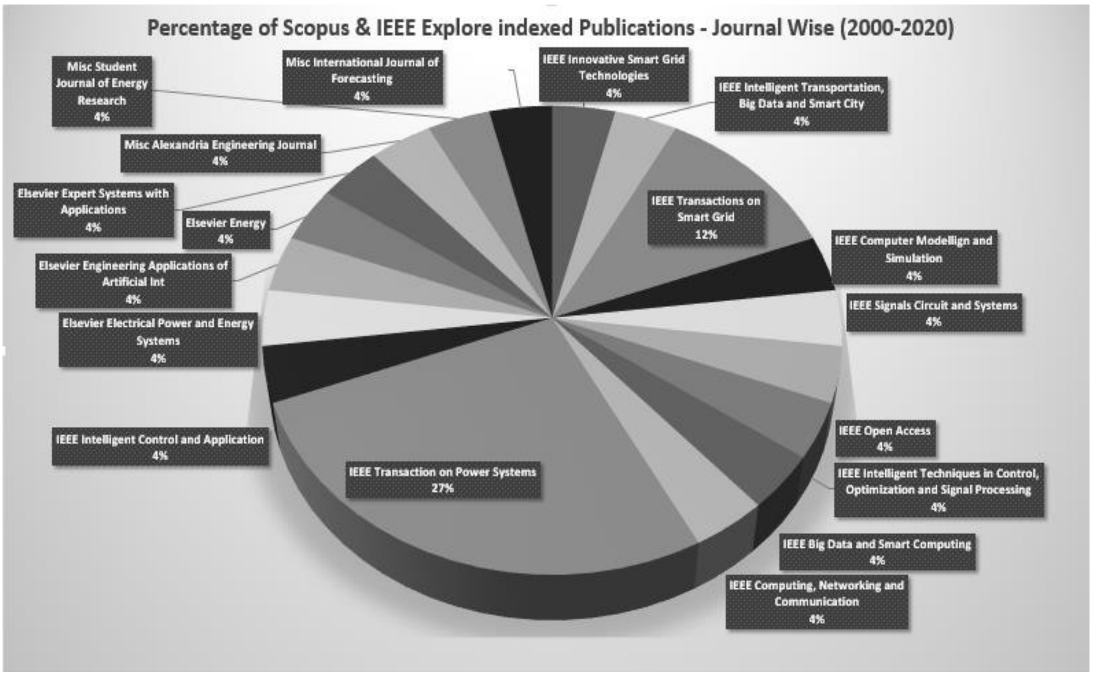
Figure 1
General Framework



Research Catalogue

In phase I, to systematically review the literature, search is mainly carried out from year 2000-2020 by using specific keywords and search engines. Scopus and IEEE Xplore are the most reliable databases in scientific community. Both these databases are used with the keywords, "electric load forecasting", "power load", "modelling electricity loads" and "long term/ Short term forecasting". 8,505 number of papers and articles came out as a result of search initially, including papers from areas of computing, power market and wind energy as well. The advanced research tool is used to narrow down the search to specific area of power engineering which then resulted in 5,009 papers. The search is further refined based upon the title of the papers to locate 1865 papers relevant to electric load forecasting and STLF. Keeping in view the time constraint, scope of the project and resources available 25 papers on STLF are selected for review and meta-analysis purposes. The top most journals that contributed towards the selected topics are found to be IEEE Transactions on Power Systems, IEEE Transactions on Power Grids, International Journal of Forecasters and International Journal of Electrical Power and Energy System as shown in the graph below.

Graph 1
Journal Wise No of Publications from 2000-2020



Analysis Approach

Finally in phase II, each publication is studied in detail for comparative analysis after selecting publications during initial phase. Owing to the variability of consumer's load demands due to various meteorological conditions, socio-economic conditions two-pronged approach is applied in this study. Firstly, a specific criterion is developed to analyze and compare the studies in detail. Since each step of the research contribute towards the final results of the studies, criteria are developed in such a sense that it covers complete research methodology. Secondly for statistical analysis the proposed methods are compared as per their Mean Absolute Percentage Error (MAPE) results. This study assumes that all the results published in the studies are correct, methods used by majority of the expert community are best and finally MAPE percentages of the studies are compared.

Criteria of Analysis

Different performance measures and results are used in different papers as per their requirements. However, meaningful meta-analysis can only be done on the basis of some criteria. This criterion needs to be selected very thoughtfully, as if it misses the relevant parameters of respective research theme then chances are high that meta-analysis may not be able to make correct assessments. Henceforth, various systematic literature reviews and studies on the subject are consulted before defining comparison criteria for this study. Criteria given in Table 2 are used to compare the papers in this study.

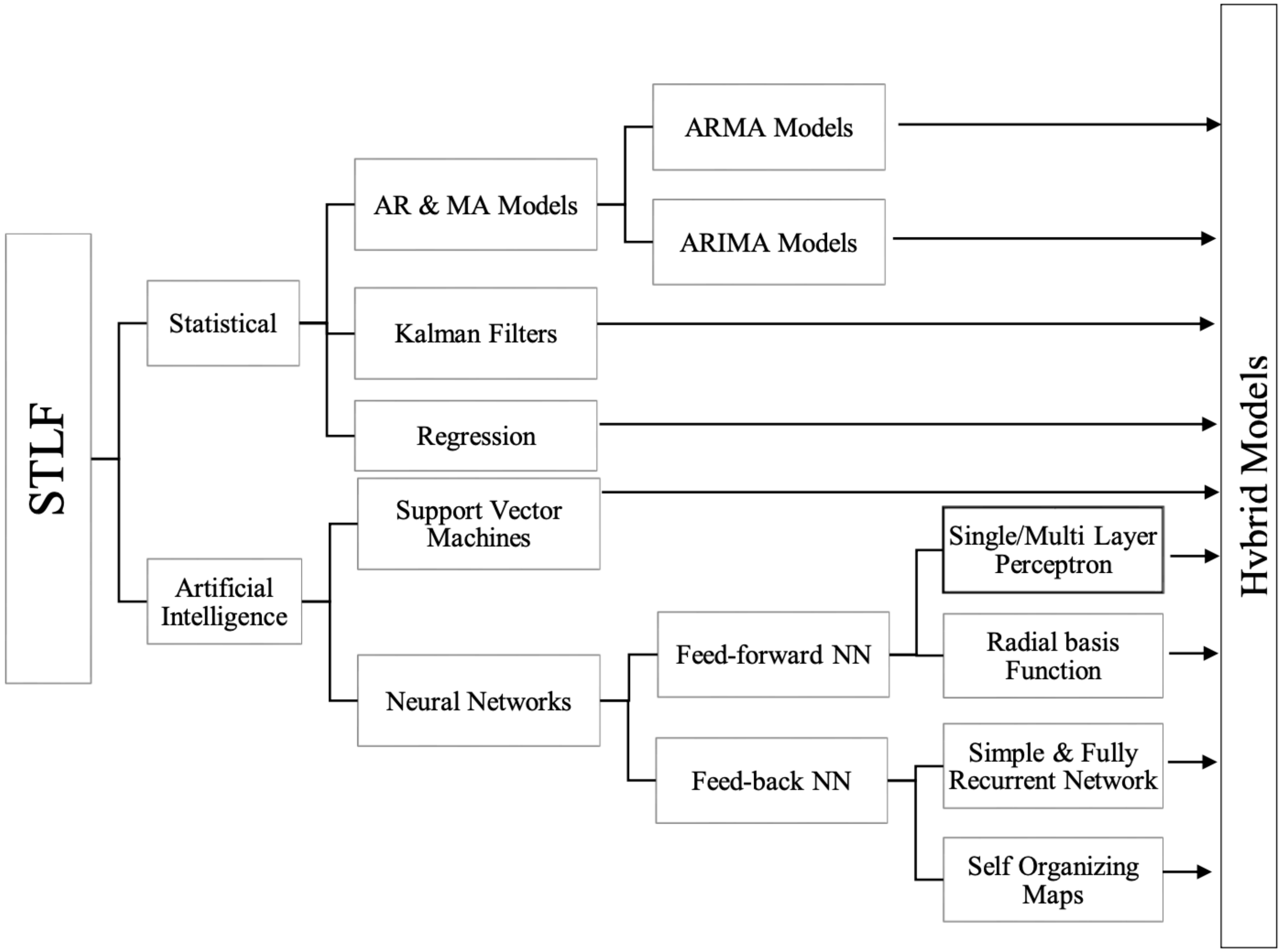
Table 2
Criteria for Analysis of various Studies

Ser	Category	Description
1.	Proposed Method	Essence of this study, as we want to check which methods used for STLF are more reliable and efficient.
2.	Dataset Used	Number of samples or data used as input plays critical role in estimation.
3.	Overview of the Methodology	What methodology is used by the author Long/ Medium/ Short/ Very Short-Term Forecasting
4.	Performance Measure	How results are compared with other methods and what are proposed method's strengths and weaknesses
5.	Prediction Term	Time duration for which prediction is made.

Research Findings

During this research it is found that a number of techniques are used by researchers while estimating load forecasts. Since, no standardized model exist for the types employed, the rise of multidisciplinary collaborations in scientific community has made the types of techniques more ambiguous to categorize. However, it is found that most of the expert community has classified techniques in two main areas; statistical and artificial intelligence based as shown below.

Figure 2
Methods used for STLF



Statistical Methods

These econometrics based mathematical models are generally based on relationships between two or more variables. The relationship is multiplicative or additive in nature. These techniques mostly use the historical load series to forecast the future load (Almeshaiei & Soltan, 2011).

Auto Regressive (AR) & Moving Average (MA) Models

These are two basic models used for studying the statistical properties of a non-stationary process. Mostly researchers use their different combinations for forecasting purposes. ARMA model is actually the integration of AR and MA models (Yang & Yang, 2019). In the AR model, present value of a load series can be expressed in the form of combination of past loads (Huang & Shih, 2003). This model can be used to predict the load value based upon past values of the load having some correlation. The equation of AR model can be written as follows:

$$y_k - \sum_{i=1}^{p} \alpha_i y_{k-i} = e_k$$

Where α_i's are the unknown coefficients of AR Model, e_k is the random noise and p the order of AR model that tells us the no of past values involved in the process. The MA models are used where load value is forecasted from past values of the input random noise values. The equation can be written as:

$$y_k = e_k + \sum_{i=1}^{q} \beta_i e_{k-1}$$

Where β_i's are the unknown coefficients of the Model, $e_k$is the random noise and q the order of MA model. Written in the notation of ARMA (p, q) these models combine strengths of AR and MA model to forecast the load value. Present values of load can be expressed in the form of past values of load and current and past value of noise, as shown in the equation below:

$$y_k - \sum_{i=1}^{p} \alpha_i y_{k-1} = e_k + \sum_{i=1}^{q} \beta_i e_{k-i}$$

Huang and Shih (2003) developed an ARMA model by adding Gaussian noise to incorporate nonlinearity and then selecting a suitable model, with an order, to predict load. Parameters are estimated using gradient based methods and finally model is validated for its adequacy with the real data. The model performed well a compared to simple ARMA and ANN. Yang & Yang (2019) developed a basic ARMA model and compared it with Projection Pursuit Regression (PPR) to be better performing.

Another variant is Auto Regressive Integrated Moving Average (ARIMA), which take into account the non-linearity involved in a time series. The AR, MA and ARMA models are applicable for stationary processes only. However, when non-stationary data is involved, data has to be transformed to a stationary form. The equation of the ARIMA model is:

$$\alpha(B). \vee d \;. \; y_k = \beta(B). \; e_k$$

Where α, β are the unknown coefficients. e_k defines the noise. Amjady (2001) employed modified version of ARIMA model by incorporating temperature and operator's knowledge in to the model. The proposed model performed better than the ARMA model for predicting next year's hourly data. The ARMA and ARIMA are used successfully by Almeshaiei and Soltan (2011) to forecast the load for Kuwaiti electric network. Their approach is to mainly use segmentation and decomposition of time series into similar regions and contours to make the forecast.

Kalman Filtering Algorithm

The long-term forecasting is generally termed by a certain level of un-certainty. To cope up with this, Kalman Filters were introduced in 1960 to minimize the mean of the squared model's error. The algorithm comprises of a set of equations that gives efficient recursive means to estimate the state of an observed sequence (Gaur & Majumdar, 2016). This technique has few powerful characteristics where it can control the highly noisy systems as well as cater for small unknown variables of the system. In load forecasting the unknown variables like weather, abrupt load demands, and customer's requirements can be addressed by this algorithm. The mechanism works in two stages. In the predictor stage, the algorithm predicts load's current state based upon its previous states along-with its covariance and in the corrector stage information from metering device is collected to and estimated state vector is updated by employing the weighted average. The Kalman Filter method generally does not take into account the non-linear issues of the load forecasting. Hence its modified versions are employed as done by Gaur and Majumdar (2016). They proposed modified Kalman Filter versions Extended Kalman Filter (EKF) and Unscented Kalman Filters (UKF) to better estimate the non-linear behavior using Jacobian Matrices.

Regression Models

Regression models are widely used statistical methods in forecasting. Main gist is learning more about the relationship between dependent and independent variables of the process. Multiple regression is based on minimizing the sum of squares of difference between observed and predicted values. Fan and Hyndman (2012) used regression technique to develop a semi parametric additive model for 24-hour demand forecast. They developed 48 models, on half hourly basis, using selected historical load and temperature data. Forecast residuals and forecast errors are calculated using modified bootstrap method and finally empirical distributions are constructed around the forecast errors for load prediction.

Non-Linear Predictors

Non-linear dynamics of power industry is explored using non-linear chaotic dynamic and evolutionary strategy by many studies. Unsihuay-vila et al. (2010) used non-linear chaotic dynamic based predictor PREDICT2 for analysis of non-linear load during training stage with emphasis on the optimization of the objective function. A new Evolutionary strategy is proposed to solve the optimization problem with candidate solution vector, having a random value with standard deviation. Dong and Grumbach (2019) applied Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks to remove gradient problems in the past load data for ELF.

Exponential Smoothing (ES)

This forecasting technique works on the weighted average of the past observations. Highest weight is given to the present value of the load, then next lower weight to the preceding value of the present value and even lower to the observation before. Due to the simplicity and accuracy ES technique is used quite frequently for the load forecasting. The ES techniques have been divided into three further divisions. Single exponential smoothing (Brown's Method) used when there is no pattern in the given data, Double exponential smoothing (Holt's Method) when trend is observed in the data and finally Triple Exponential Smoothing (Holt Winters Method) when data reveals significant seasonal configurations. Taylor (2012) developed five ES weighted models including a Singular Value Decomposition SVD based model to reduce the data to lower dimensions with uncorrelated variables. In (Taylor & McSharry, 2007) they proved that their proposed Seasonal Holt Winters Exponential

Smoothing method out performed ARMA and PCA models. They used the models to forecast seasonal demands of European data. They added an additional index and smoothing equation for forecasting the load. Also, ARMA and PCA models are developed to compare the performance. Kavanagh (2017) calculated load forecast for Irish market using Double Seasonal Holt Winter's Exponential Smoothing with Error Correction. Seasonal parameters are initialized from the historical load data and model is proposed using the exponential smoothing algorithm. Finally, GRG nonlinear error of predicted value and actual data is calculated.

Support Vector Machines (SVM)

Presented by Vapnik in 1995, the SVM are basically classification and regression techniques. SVM mainly extract the decision rules having satisfactory generalization ability from the training data called support vectors (Zhang & Hong, 2020). Input space is mapped nonlinearly into a higher space dimension constructing an optimal hyper plane. In the training phase of the SVM, linearly constrained quadratic programming is carried out which is unique but time consuming. Fan and Chen (2006) used Self Organizing Mapping (SOM) technique to organize the input data into clusters. SVMs are then applied on each data subset to forecast the load for next day. This hybrid method proved helpful in addressing the non-stationary load time series. Pai and Hong (2005) and Zhang and Hong, (2020) applied VMD is applied to decompose input data into sub series based on certain center frequency and bandwidth. Nonlinear mapping function is used to map data in high dimension where SVR function is used to relate forecast value with input.

Artificial Intelligence (AI) Models

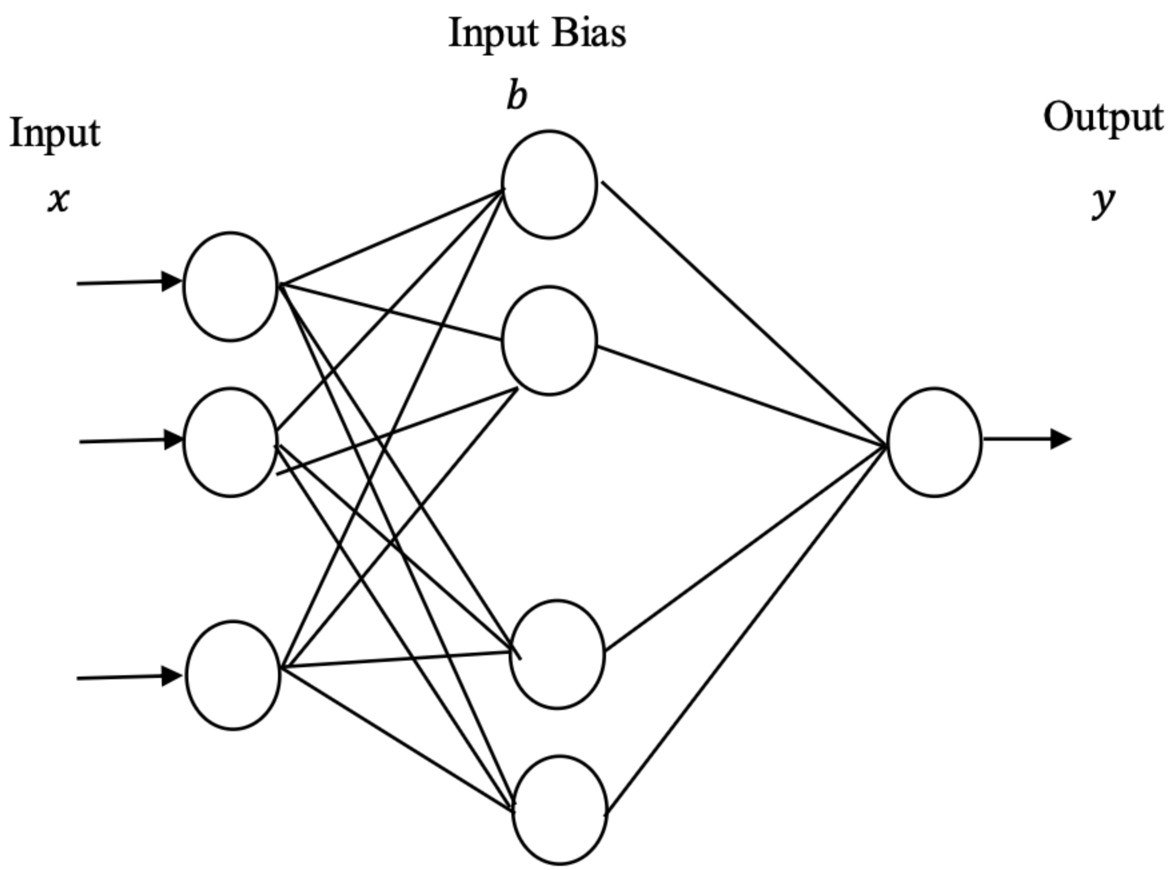
With the advent of advanced technology and high computational powers, AI systems have been developed for forecasting and estimating.

Artificial Neural Network (ANN)

Developed in 1990 by Warren McCulloch and Walter Pitts, the ANN have been applied in several areas including forecasting and classifications (Ribeiro et al., 2019). ANN is basically a non-linear circuit that can perform non-linear curve fitting. It processes information in line with the human biological systems. Inspired by the working of human brain, the NN can process the certain piece of information using its basic unit called neuron. Information received at the input node of the neuron is accumulated, processed and then further forwarded to the next neuron through output node. The ANN system is trained on the relevant historical data where it identifies the similarities and patterns of the input data. Then basing upon this prior knowledge about the data and system, the network gives generalized output. In its most simplistic form, the network consists of an input layer, a hidden layer and finally an output layer. Input I is sent to the hidden layer, along-with associated weights, which perform certain function f(x) to give an output. Based upon its topology, the ANN are generally categorized into Feed Forward (FF-NN) and Feedback or Recurrent NN

Feed Forward Neural Networks (FF-NN)

Usually preferred for forecasting and consists of various combinations of an input, hidden and an output layer. In its simplistic form Single Layer Perceptron, no hidden layer exists. The forecasts are obtained using linear combination of inputs and weight vectors, which are obtained using learning algorithm that minimized some cost function e-g MSE. With the addition of an intermediate layer, the NN takes the form of Non-Linear Multi-Layer Perceptron (MLP). Neurons are arranged in layers and connected through weight vectors with next layer. Neuron b takes the input from its predecessor neuron, if it exists, computes the weighted sum w, eliminate the bias and gives the output after applying the activation function $\boldsymbol{g}$.



Equation is given by:

$$\sum_{j=1}^{n} \beta_j\, g(w_j \,.\, x_i + b_j) = y_i$$

Where, x_i is the input, w_j is the weight, b is the neuron of hidden layers.

Feedback Neural Networks

Unlike the FF-NN, the feedback NN is dynamic in nature. Their output depends on the state of the system. Whenever a new input pattern in given the output of neurons are computed. Feedback of the neurons are modified due to the feedback system and hence the NN enters into a new state. To overcome the vanishing gradient problem of the NN, Nonlinear Autoregressive Models with Exogenous Inputs (NARX) has been developed. This three-layer FF-NN with good learning capabilities has sigmoid activation function in its hidden layer, linear activation function in output layer and delay lines for storing previous predicted values.

NN based STLF has been enhanced using Multi-resolution analysis (MRA) by Reis and Silva (2005). Four models are developed with different input variables among load, temperature, differenced load from first model and MRA with differenced load. The final models comprise of sub-models of first three models to decompose the load series using individual fitting. Proposed Model with load, temperature and first order differenced load as input predicted the load most accurately. Chen et al (2008) proposed a hybrid model using Wavelets with NN. The historical data is decomposed into Low and High Frequency components then NN are then applied for load forecasting. Ribeiro et al. (2019) also proposed Wavelet based NN (WNN), using previously used algorithms for generation, selection and generalization, to compare its prediction performance. However, they concluded that results of WNN are comparable with naïve methods and MLP NN on GEFCom dataset. Lv et al. (2019) improved the BP NN with introduction of GA. They used PSO to improve the convergence speed and PCA to reduce the matrix dimensionality.

Hamid and Rehman (2010) introduced Artificial Immune System with ANN. Aim is to check benefits of the robust AIS like computational strengths as its distributed, diverse with anomaly detection and self-organizing learning abilities. The performance of AIS based FF-NN has comparable results on MAPE scale with that of BP NN. However, further studies may reveal true potential of AIS in the field. Different types and scale of NN have been in use for last two decades by the researchers for load forecasting. Having received considerable success, the NN are also criticized for having too many input parameters which can lead towards data over fitting. Hippert et al. 2005 conducted a detailed review of various models of NN with traditional statistical methods. They compared large NN with linear models including Naïve forecasting, methods with one and more smoothing filters, smoothing filters with linear regression combination of smoothing filers and NN. Conclusion is that large NN can perform well because they consider more historical data and also, they have ability of interpolating high dimensional functions which improves the profile load forecasting. Michanos et al. (2003) also worked to investigate the non-linear characteristics of the power load series is identified using MLP. An attractor is then developed in phase plane to train ANN. Wang et al. (2018) proposed a set of probabilistic models as constrained quantile regression models to average and predict the future data.

While using NN Liu et al (2015) employed wavelet-based ensemble scheme. Selection of mother wavelet and decomposition level is a tricky affair. Here ensemble of wavelets is used and their output is aggregated to get the best features as output. Wavelet based ensemble networks, algorithm incorporating Levenberg–Marquardt (LM) for improved learning, Conditional Mutual Information Feature Selection (CMIFS) method is employed for feature selection and Partial Least Square Regression (PLSR) is used for forecasting purposes. In another scheme, (Charytoniuk & Chen, 2000) exploited that NN learn dynamics of load without memorizing the data for long time with accurate results. Challenges are faced while extrapolating the relationships different from those extracted from training data. Five models of three-layered FF NN are used for forecasting. Redundant hidden neurons are also eliminated by observing duplication in co-linearity with the output.

Hybrid Models

Applications of combining strengths of different models into a hybrid model have been published by many in the past. Also, for the load forecasting various methods have been combined to produce efficient methods. The probabilistic nature of power systems, make it a potential field for employing various methods to estimate the forecast for time ahead.

Results

Since there is no gold standard yet for types of forecasting and methods for prediction, reviewers take into considerations various assumptions to while comparing the publications. This study assumes that all the results published in the studies are correct, methods used by majority of the expert community are best. However, for statistical analysis MAPE percentages of all studies are also compared.

Comparative Analysis

After selecting publications for this study during initial phase, each publication is studied in detail for comparative analysis. The dynamics of topic and unpredictability of load influence the researchers to use different variables and performance standards to measure their proposed methods. In addition, heterogeneity of data due to socio-economic conditions and consumer’s profile, non-linear environmental conditions including weather, humidity for different countries add complexity to the comparison of studies. At times a simple or a particular method favor a particular situation even than the sophisticated techniques. Main objective is on the fields mentioned in the table of basis of comparison including Proposed Method, Data set used, Overview and Prediction Term.

Table 3
Comparison of Studies

Ser	Paper/ Year	Technique	Dataset/ Training & Testing	Overview
1.	Yang, & Yang, 2019	Hybrid using Auto-regression Integrated moving average (ARIMA) & projection pursuit regression (PPR)	5 min interval time series of Sichuan Electric Power Company, China with 864 observations for randomly selected data from 26-28 Sep 2016. 576 for training and 288 for testing out of the total of 864	• ARIMA is modelled estimating its order and parameters using Bayesian Information Criteria (BIC) and correlation function. • Next the PPR model is developed. Integration of both models is carried out to address the linear and non-linear dynamics of the load forecast.
2.	Almeshaiei &Soltan, 2011	Hybrid model using Autoregressive Moving Average ARMA and ARIMA.	Daily load data of Kuwaiti Electric network from 2006 – 2008.	• Data is segmented to locate the identical patterns and calculate their probability. Afterwards series is decomposed using MA for load pattern segmentation. • For load forecasting, curve fitting is conducted to identify region similarity, contours and related points.
3.	Huang Shih, 2003	ARMA including Gaussian and Non-Gaussian Processes	Hourly data of 3 months between 1998-1999 of Taipower Company, Taiwan	• Historical data is processed for Gaussian test using Bi-spectrum process. • If data is Gaussian, second order statistics are calculated otherwise MA model is applied. • Correct model identified in previous step is used to estimate parameter representation of model.
4.	Amjady, 2001	ARIMA Model integrated with operators' knowledge	Hourly load and peak load data from Iran's national grid from 1996-1998. Data from 1996-1997 is used for training, while that of 1998 for testing purposes.	• ARIMA model is proposed considering the historical data, estimating the parameters. • 16 Modified ARIMA models are used for forecasting along-with the temperature and operators' knowledge

5.	Dong & Grumbach, 2019	Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks	3-year record data of various feeders from West Canada with 1997 records 1,597 records for training and 400 out of total 1997 (80%/20% split ratio)	• Features from past data are collected based upon socio-economic and weather conditions. • Principal component Analysis (PCA) and Normalization of selected features is performed. • LSTM and GRU networks for Many to Many and One to Many configuration is developed. • These networks are better to vanish and explode gradient problems in the data.
6.	Unsihuay Vila, et al., 2010	A hybrid approach based on non-linear chaotic dynamic predictor	Hourly electricity load of year 2002 from New England, Albert and Spain. Random selection of 4 weeks, one each in 4 months of a year, for testing. Rest of the data used for training	• The time series data, set as input to the model, is divided into two segments. One segment used to predict data in second segment. • Population is initialized using candidate random variables. Next the population parameters are recombined to produce off-springs and then mutate. • This ES is used to tune the prediction parameters.
7.	Fan & Hyndman, 2012	Semi parametric additive models using Modified Bootstrap method	Half hourly demand and temperature data of Melbourne from 1997-2009 from Australian National Electricity Market. Data from 2004-2008 used for training and 2009 data for testing.	• Semi parametric model is developed to forecast demand and temperature values using their historical data. Cross validation is done to select the variables for use in models. • Forecast residuals are calculated by sequentially substituting into random forecasted model values. • Modified bootstrap method is used to obtain forecast errors.
8.	Gaur &Majumdar, 2016	Modified Non-linear Kalman Filter, Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). Weather and Wind Speed data is	Reference Energy Disaggregation Dataset (REDD) anonymously collected from Boston, US	• Standard KF is modeled using past data, temperature and wind speed data. • The model predicts the value based on past data along-with its covariance. • The output is recursively updated using the law of minimizing mean square error. • The non-linear Modified filters, EKF and UKF are applied to calculate the prediction.

		accumulated from the website.		
9.	Taylor &McSharry, 2007	ES with Holt Winters, ARIMA and PCA	30 Weeks hourly/ half hourly data of 6/ 4 European countries from Apr – Oct 2005. First 20 weeks of each data is used for training and last 10 weeks for testing.	• Seasonal Holt Winters Exponential Smoothing is applied to forecast two seasonal demands. • Additional seasonal index and extra smoothing equation is added for double seasonal method. • The initial level and seasonal values are estimated by averaging the observations and minimizing squared sum of errors. • ARMA and PCA models are also developed to compare the performance
10.	Kavanagh, 2017	Double Seasonal Holt-Winters Exponential Smoothing with error correction	Half hourly data of 15 months from an Irish supply company from Jan 2013-March 2014	The seasonal parameters defined in the form of Days, Weeks and Seasons are initialized from the historical load data. The model is proposed using the exponential smoothing algorithm. Finally, GRG nonlinear error of predicted value and actual data is calculated.
11.	Taylor, 2012	Five exponentially weighted methods incl new Singular Value Decomposition SVD based ES	Half hourly observation from 2007-2009, first 2 years used for training and last year for testing	• SVD based approach is used to reduce the data to lower dimensions with uncorrelated variables. • Modified Holt Winter ES (HWT) • Discounted weight regression (DWR)
12.	Fan & Chen 2006	Hybrid approach to combine SOM with SVM	Hourly data of one year from 2003-2004 of New York City, US	• In the first stage SOM is used to group the training data with similar properties. • SVM network of 24 machines is then applied with regression and risk minimization principle to forecast next day's load.
13.	Pai& Hong 2005	Simulated Annealing with SVM	Taiwanese load data from 1045-2003, 40 years training set from 1945-1984, 10 yrs. validation 1985-1994, 9 yrs. testing 1995-2003	• The past data is normalized using the simulated annealing algorithms. • Then SVMs are applied for load forecasting. • The proposed model is compared with ARIMA and Regression NN.

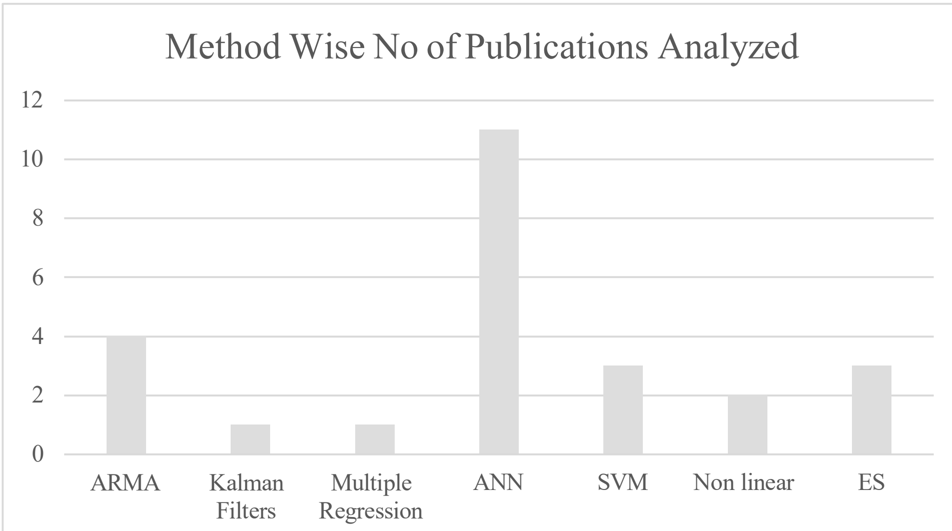
14.	Zhang & Hong, 2020	Hybrid model using Variational Mode Decomposition Self Recurrent Support Vector Regression Cuckoo Bird Cuckoo Search (VMD-SR-SVRCBCS)	Half hour load data from National Electricity Market, Queensland, Australia and hourly load data from New York Independent System, USA. Both datasets distributed into 3 x parts for training, validation and testing phase	• VMD is applied to decompose input data into sub series based on certain center frequency and bandwidth. • Nonlinear mapping function is used to map data in high dimension where SVR function is used to relate forecast value with input.
15.	Michanos, 2003	ANN based on Multilayer Perceptron	Daily peak demand of one year for 1995, 9 months data is used for training and 2 months for testing	• Time series is extended to confirm its chaotic character using correlation dimension and Lyapunov Spectrum. • A state space of differential equation is created for the time series taking into account all its variables. • Then model based on correlation dimension and state space of the data is developed.
16.	Lv et al., 2019	Neural Network optimized using Particle Swarm Optimization (PSO) and Principal component Analysis (PCA)	1 year data of a Power Grid Corporation Previous one year data for training	• PSO is used to initiate the model from initial weights and thresholds. • PCA is used to reduce the input dimension as per the set threshold with GA optimization of the model. • Load is forecasted for next 24 hours.
17.	Hamid & Rahman, 2010	Feed Forward-Neural Network (FF-NN) trained by the Artificial Immune System (AIS)	Day, time, temperature and 720 samples each from historical load data of Kuala Lumpur, Malaysia and the other from North Carolina, US 65% of the data is used for training and 35% for testing	• AIS based algorithm is developed with initial weights selected randomly between 0 and 1. • FF-NN on MLP architecture is proposed where input parameters are multiplied with weights. • Regression is performed to correlate the predicted values with the past load series.

18.	Dong et al. 2017	Convolutional Neural Network (CNN) with K Means clustering is employed.	1.4 million records of electricity data from 2012-2014 containing hourly load data from power industry. 1,003,716 samples from 2012 - 2013 are used for training and 469300 samples for testing.	• Raw data is pre-processed, converted into two subsets, training and testing, based upon selected feature analysis using K Means Clustering. • CNN is trained on one subset and then validated on testing subset.
19.	Din & Marnerides , 2017	Feed Forward Deep Neural Network (FF-DNN) and Recurrent Deep Neural Network (R-DNN)	Hourly data of NEW England, USA from 2007-2012 comprising 52600 records. 43824 samples are used for training while rest are used for testing phase.	• The data is analyzed in time and frequency domain to model it comprehensively. • In next stage, Rectifier Activation Function (ReLU) is used to model FF-DNN and R-DNN. • Separate results are computed considering only Time Domain and Time & Frequency Domain features.
20.	Chen et al. 2018	Modified Deep Residual Network adopting ensemble strategy	North American utility data set with hourly data from 1985-1992. 2-year data from 1991-1992 is used as test data, rest of the data is used for training. To check the generalization ISO-NE data is used.	• Two level basic structure is formed for forecasting 24 hours data with Scaled Exponential Linear Units (SELU) activation function. • Output is fed into Deep Residual Network (ResNet) constructed from a stack of three residual blocks. • Modifications is made, ResNetPlus, by employing number of side residual blocks and averaging output of each main residual block with these side blocks to improve error back propagation of the network. • Next ensemble strategy is used to improve generalization capability of the network.
21.	Reis & Silva, 2005	NN with Wavelet decomposition	Hourly load data of North America from 1988-1992	• Features are extracted into Low and High Frequency components using Multi Resolution Analysis. Input variables are selected by applying correlation functions. • 4 models of NN have been developed based upon MLP.

				• The model with inputs of load, temperature and first order differenced performed the best among other NN.
22.	Chen et al., 2008	Wavelet Neural Networks	Hourly load data of NEW England from 2003-2005 as training and year 2006 data used for testing.	• Wavelets are used to decompose the load into Low and High Frequency components • MLP based NN are then applied for load forecasting.
23.	Hippert et al., 2005	Large NN and regression methods	Hourly data from 1996-1997 of a city of Brazil. Data is split for training, testing and validating phase.	• Various models are developed based upon Naïve forecasting, methods with one and more smoothing filters, smoothing filters with linear regression combination of smoothing filers and NN and large NN
24.	Charytoniuk & Chen, 2000	NN	Load data of previous 1 hour is used to predict next 20 minutes load for a power company in US.	• They used relative load curves of past data instead of load increments to improve the forecasting accuracy as is done in traditional NN models. • Input variables are selected basing upon their string statistical correlation with outputs. • Supervised training is carried out for the proposed NN using the previous load data and minimizing the error function.
25.	Li et al., 2015	Hybrid NN based on Wavelets	Hourly load data from ISO England for year 2009-2010.	• ELM-LM algorithm is developed by randomly initializing the weights and biases to estimate the output weights. • Wavelet transform is used to employ frequency components along-with temporal dimensions of the past load series. • PLSR is used to combine the forecasts of different wavelets. • Hourly load data is fed into 24 FF-NN with the detailed extraction of frequency components using wavelet transforms.

ANN stands out to be the most used methods for ELF, after going through all the publications during this study as shown in the graph below. Eleven out of 25 publications used ANN in its one form or the other. ARMA based models are next most frequently used method for ELF by the researchers.

Graph 2
No of Publication Studied as per Methods



Statistical Analysis

For performance measurement, researchers have used various forecasting standards like including Mean Absolute Percentage Error (MAPE), Root Mean Square Error etc. However, MAPE is used more frequently in the statistical studies. The MAPE difference of the proposed and other methods are calculated. Then the smallest value is located to identify the best method. It is revealed that ARMA models and ANN gave the lowest MAPE values. A stable difference criterion has also been defined by setting the value of alpha from 0.01 to 5. This means that difference values less than 0.01 and greater than 5 are ignored in this study. Difference value is calculated by taking absolute difference between proposed method and other methods with which it is compared. The mean MAPE for ANN and ARMA Models is 0.799 and 0.8446, respectively.

Table 4
Comparison of MAPE & Standard Deviation

#	Cat	Ref	Proposed Method	Bench Mark Method	% MAPE	Mean - MAPE
1.	ARMA / ARIMA	Yang, & Yang, 2019	Hybrid using ARIMA & Projection Pursuit Regression (PPR)	ARIMA PPR	0.634 0.403	**0.8446**
2.		Almeshaiei & Soltan, 2011	Hybrid model using ARMA and ARIMA.	Real data	0.5	

3.		Huang & Shih, 2003	ARMA including Gaussian and Non Gaussian Processes	ARMA ANN	0.05 0.58	
4.		Amjady, 2001	ARIMA Model integrated with operators' knowledge	ARIMA ANN Operators	1.24 1.27 2.08	
5.	Non-Linear	Dong & Grumbach, 2019	Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks	FNN Modified FNN	5.1 2.79	5.05
6.		Unsihuay Vila, et al., 2010	A hybrid approach based on non-linear chaotic dynamic predictor	ANN ARIMA	5 4.5	
7.	Regression	Fan & Hyndman, 2012	Semi parametric additive models using Modified Bootstrap method - Regression	ANN Hybrid	0.85 0.4	0.85
8.	Kalman Filter	Gaur & Majumdar, 2016	Modified Non-linear Kalman Filter - Kalman	EKF	0	0
9.	Exponential Smoothing (ES)	Taylor & McSharry, 2007	ES with Holt Winters.	ARMA PCA AR	0.059 0.05 0.086	0.69
10.		Kavanagh, 2017	Double Seasonal Holt-Winters Exponential Smoothing with error correction	Naïve w/o EC	1.66 5.3	
11.		Taylor, 2012	Five exponentially weighted methods incl new Singular Value Decomposition SVD based ES	ANN HWT with SM NEW SVD Weather based	0.02 0.01 0.01 0.016	

12.	Vector Machines (SVM)	Fan & Chen 2006	Hybrid approach to combine SOM with SVM	ISO SVM	1.15 0.65	2.42
13.		Pai & Hong 2005	Simulated Annealing with SVM	ARIMA GRNN	8.55 3.42	
14.		Zhang & Hong, 2020	Hybrid model using VMD-SR-SVRCBCS	ARIMA GRNN BPNN SVR	7 5.3 5.1 3.8	
15.	Neural Networks (NN)	Michanos, 2003	ANN based on Multilayer Perceptron	Others	0.4%	**0.799**
16.		Lv et al., 2019	NN optimized using PSO and with PC Analysis	No PC Reduction	1.1	
17.		Hamid & Rahman, 2010	FF-NN trained by the Artificial Immune System (AIS)	Data 1 – AIS Data 2 – AIS	0.473 1.347	
18.		Dong et al. 2017	CNN with K Means clustering is employed.	LR SVR SVR & K Means NN NN & K-Means	25 8.97 0.89 0.163 0.115	
19.		Din & Marnerides, 2017	FF-DNN and Recurrent Deep Neural Network (R-DNN)	Time Frequency	12 0.01	
20.		Chen et al. 2018	Modified Deep Residual Network adopting Ensemble Strategy	Temperature-1 Temperature-2 Temperature-3	0.02 0.05 0.11	
21.		Reis & Silva, 2005	4 models based on NN with Wavelet decomposition with different inputs	Model-1 Model-2 Model-3	0.17 0.42 0.94	
22.		Chen et al., 2008	Wavelet NN	NN w/o weather NN & weather Similar day	0.188 0.07 0.22	
23.		Hippert et al., 2005	Large NN and regression methods	Smoothing (small NN) Large NN	0.1 1	
24.		Charytoniuk & Chen, 2000	NN	Forecaster-1 Forecaster-2 Forecaster-3	0.59 0.39 0.23	

25.		Li et al., 2015	Hybrid NN based on Wavelets	Abductive MLR RBFNN Random forest	0.79 0.83 0.56 0.92	

Discussion

Load forecasting has become a topic of significance in past few decades. Various techniques have been used by researchers to identify the best performing methods. However, the non-linear dynamics of the topic implies that not one method can be classified as the best. Availability of historical load data is the prime factor in forecasting, however heterogeneity in this data itself challenges the analysis. The data is dispersed in different patterns with different power companies. Somewhere it is calculated on hourly basis, whereas at the other places it is recorded on seasonal basis.

Most of the statistical methods employ past load series and weather information for prediction. These past load data are used as input to Regression techniques along-with weather and its functional relationship. Same is then solved regressively to reduce the square error of the prediction. Exponential smoothing models are developed by linear combination of time series and other variables. Kalman filtering use filtering techniques to reduce the noise in data to predict future load.

The ANN techniques have performed quite well for ELF. When combined with Wavelet decomposition forecasting is improved further as it employs frequency component of data series as well. However, their main concern is data fitment. The NN employ layers of neurons and a large no of parameters which raise the concern of over parameterization in performing the task. Mostly, large NN perform better in forecasting results but the theory behind this remains a black box.

Meteorological conditions also risk the load forecasting. Although in today's digital world, previous data as well as future weather forecasts are also available. Still the unpredictability of the weather, humidity conditions play significant role in the load forecasting. Then the socio-economic conditions of the consumers dictate the variability of load demands. One cannot take into account the functions, gatherings or other related activities at a specific place.

Another important concern is about the transmission network dynamics. Equipment failures and accidents make the power unavailable at one region thus causing variation in demand at another generating region.

Conclusion

Meta-Analysis is carried out by studying 25 publications on ELF modeling proposed by researchers and compared with various other forecasting methods. The criterion for comparison is selected including technique employed, data set used, overall methodology, performance and MAPE measurement. The comparative results show that various non-linear factors play significant role in ELF, importantly weather conditions. Few methods are preferred because of their fast computation power and linear relationship among variables.

ANN and ARMA are found to be the best performing methods. ANN are mostly used when changes are occurring at faster pace like frequent change in weather or environmental conditions. However, with larger NN the issues of data over fitment need to be taken into consideration. The ARMA models are attractive due to ease in their practical interpretation. They are usually criticized for their limitation to deal with non-linearity behavior of processes.

References

Abdel-Aal, R. E. (2004). Short-term hourly load forecasting using abductive networks. IEEE Transactions on Power Systems, 19(1), 164-173.

Almalaq, A., & Edwards, G. (2017, December). A review of deep learning methods applied on load forecasting. In 2017 16th IEEE international conference on machine learning and applications (ICMLA) (pp. 511-516). IEEE.

Almeshaiei, E., &Soltan, H. (2011). A methodology for electric power load forecasting. Alexandria Engineering Journal, 50(2), 137-144.

Amjady, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability. IEEE Transactions on power systems, 16(3), 498-505.

Chen, K., Chen, K., Wang, Q., He, Z., Hu, J., & He, J. (2018). Short-term load forecasting with deep residual networks. IEEE Transactions on Smart Grid, 10(4), 3943-3952.

Charytoniuk, W., & Chen, M. S. (2000). Very short-term load forecasting using artificial neural networks. IEEE transactions on Power Systems, 15(1), 263-268.

Chen, Y., Luh, P. B., & Rourke, S. J. (2008, June). Short-term load forecasting: Similar day based wavelet neural networks. In 2008 7th World Congress on Intelligent Control and Automation (pp. 3353-3358). IEEE.

Din, G. M. U., &Marnerides, A. K. (2017, January). Short term power load forecasting using deep neural networks. In 2017 International Conference on Computing, Networking and Communications (ICNC) (pp. 594-598). IEEE.

Dong, M., & Grumbach, L. (2019). A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction. IEEE Transactions on Smart Grid, 11(1), 470-482.

Dong, X., Qian, L., & Huang, L. (2017, February). Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach. In 2017 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 119-125). IEEE.

Fan, S., & Chen, L. (2006). Short-term load forecasting based on an adaptive hybrid method. IEEE Transactions on Power Systems, 21(1), 392-401.

Fan, S., & Hyndman, R. J. (2011). Short-term load forecasting based on a semi-parametric additive model. IEEE Transactions on Power Systems, 27(1), 134-141.

Gaur, M., &Majumdar, A. (2017). One-Day-Ahead Load Forecasting using nonlinear Kalman filtering algorithms.

Hamid, M. A., & Rahman, T. A. (2010, March). Short term load forecasting using an artificial neural network trained by artificial immune system learning algorithm. In 2010 12th International Conference on Computer Modelling and Simulation (pp. 408-413). IEEE.

Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. (2020). Methods and Models for Electric Load Forecasting: A Comprehensive Review. Logistics & Sustainable Transport, 11(1), 51-76.

Hippert, H. S., Bunn, D. W., & Souza, R. C. (2005). Large neural networks for electricity load forecasting: Are they overfitted?. International Journal of forecasting, 21(3), 425-434.

Huang, S. J., & Shih, K. R. (2003). Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. IEEE Transactions on power systems, 18(2), 673-679.

Kavanagh, K. (2017). Short Term Demand Forecasting for the Integrated Electricity Market. Student Journal of Energy Research, 2(1), 1.

Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: A critical systematic review. Sustainable cities and society, 35, 257-270.

Kumar, M. S., Subathra, B., & Srinivasan, S. (2017, March). Ensemble wavelet learners for demand forecasting in energy grids. In 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) (pp. 1-6). IEEE.

Kermanshahi, B., &Iwamiya, H. (2002). Up to year 2020 load forecasting using neural nets. International Journal of Electrical Power & Energy Systems, 24(9), 789-797.

Li, S., Wang, P., & Goel, L. (2015). A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. IEEE Transactions on power systems, 31(3), 1788-1798.

Lindberg, K. B., Seljom, P., Madsen, H., Fischer, D., & Korpås, M. (2019). Long-term electricity load forecasting: Current and future trends. Utilities Policy, 58, 102-119.

Lv, Y. C., Xu, X., Xu, R. L., &Ren, H. (2019, January). Research on Short-Term Load Forecasting Approach for Smart Grid. In 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS) (pp. 602-605). IEEE.

Michanos, S. P., Tsakoumis, A. C., Fessas, P., Vladov, S. S., &Mladenov, V. M. (2003, July). Short-term load forecasting using a chaotic time series. In Signals, Circuits and Systems, 2003. SCS 2003. International Symposium on (Vol. 2, pp. 437-440). IEEE.

Mohammed, Z. C., Humberto, A. R. J., Kiplel., M. C. & Chepkoech, M. (2017, Dec). A Meta-Analysis on Sustainable Supply Chain Management: An Analytical approach. European Journal of Logistics, Purchasing and Supply Chain Management

Neuman, L. W. (2007). Social research methods, 7/E. Pearson Education India.

Niu, D., Wang, Y., & Wu, D. D. (2010). Power load forecasting using support vector machine and ant colony optimization. Expert Systems with Applications, 37(3), 2531-2539.

Pappas, S. S., Ekonomou, L., Karamousantas, D. C., Chatzarakis, G. E., Katsikas, S. K., &Liatsis,

P. (2008). Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. Energy, 33(9), 1353-1360.

Ribeiro, G. T., Mariani, V. C., & dos Santos Coelho, L. (2019). Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting. Engineering Applications of Artificial Intelligence, 82, 272-281.

Reis, A. R., & Da Silva, A. A. (2005). Feature extraction via multiresolution analysis for short term load forecasting. IEEE Transactions on power systems, 20(1), 189-198.

Taylor, J. W., & McSharry, P. E. (2007). Short-term load forecasting methods: An evaluation based on European data. IEEE Transactions on Power Systems, 22(4), 2213-2219.

Taylor, J. W. (2011). Short-term load forecasting with exponentially weighted methods. IEEE Transactions on Power Systems, 27(1), 458-464.

Unsihuay-Vila, C., De Souza, A. Z., Marangon-Lima, J. W., & Balestrassi, P. P. (2010). Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. International journal of electrical power & energy systems, 32(2), 108-116.

Upadhaya, D., Thakur, R., & Singh, N. K. (2019, October). A systematic review on the methods of short term load forecasting. In 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC) (pp. 6-11). IEEE.

Wang, Y., Zhang, N., Tan, Y., Hong, T., Kirschen, D. S., & Kang, C. (2018). Combining probabilistic load forecasts. IEEE Transactions on Smart Grid, 10(4), 3664-3674.

Yang, L., & Yang, H. (2019, May). A Combined ARIMA-PPR Model for Short-Term Load Forecasting. In 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia) (pp. 3363-3367). IEEE.

Zhao, H., & Tang, Z. (2016, June). The review of demand side management and load forecasting in smart grid. In 2016 12th World Congress on Intelligent Control and Automation (WCICA) (pp. 625-629). IEEE.

Zhang, Z., Hong, W. C., & Li, J. (2020). Electric load forecasting by hybrid self-recurrent support vector regression model with variational mode decomposition and improved cuckoo search algorithm. IEEE Access, 8, 14642-14658.

Zor, K., Timur, O., & Teke, A. (2017, June). A state-of-the-art review of artificial intelligence techniques for short-term electric load forecasting. In 2017 6th International Youth Conference on Energy (IYCE) (pp. 1-7). IEEE.