A Review and Taxonomy on Data Driven Models for Estimating Electrical Load for Power Systems

Richa Kumari[1]
M.Tech.Scholar[1]
Technocrats Institute of Technology & Science, Bhopal.[1]
rangoliricha@gmail.com[1]

Devendra Sharma[2]
Asst. Professor [2]
Technocrats Institute of Technology & Science, Bhopal.[2]
devendrasharma798@gmail.com[2]

Dr.Vikas Gupta[3]
Professor[3]
Technocrats Institute of Technology &Science, Bhopal.[3]
vikasgupta.bhopal@gmail.com[3]

Vasant Acharya[4]
Asst. Professor[4]
Technocrats Institute of Technology, Bhopal.[4]

Abstract: Due to increasing number of consumers, managing electrical load has become a serious challenge. Hence a good estimation of electrical load is necessary for smooth conduction of power systems. Previously statistical models were used for the prediction of electrical load, but with the advent of evolutionary algorithms, load forecasting using such approaches has become an active area of research. The present paper introduces the need of electrical load prediction and subsequently the use of evolutionary algorithms. Different techniques and their properties are presented here for a clear understanding of the tools being used. Finally evaluation parameters are discussed which evaluate the performance of any proposed system.

Keywords:-
Electrical Load Forecasting, evolutionary algorithms, iterations, mean absolute percentage error, mean square error.

I.Introduction

For a long time, the electricity has been the most sought after form of energy resource. Since the civilization has begun, electrical form of energy demands have become surplus. With rapid advancement and the digital revolution taking place all over the world, the electricity consumption has increased manifolds. Also rising population and urbane lifestyle has increased the electrical load. The energy in the electrical form has been the most easy and conducive form of energy. Henceforth the power distribution relies majorly on converting the other forms of energy into electrical form and then distributing it. The power grid system in an area is responsible for the distribution of the power across power lines and sub stations. This inter connection system design is very helpful as it aids in electric load sharing and is very economical to use.

The power demands keeps on varying with changing seasons. For instance, in summer, the consumption rate of power is very high due lots of cooling electrical appliances in use. In winters, the energy consumption reduces drastically. Also, sometimes the energy requirement can vary depending on various factors. So, it kind of keeps varying and fluctuating. So the power generation needs to be in sync with the energy requirement and consumption. There needs to be a balance between the energy generated and the energy consumed. Then only there will be proper utilization of the electrical load and minimal wastage. Consequently an efficient mechanism is need that would help in forecasting the electric load that would help in a stable power management system.

II. Previous Work

D. Cao et al. in [1] proposes a robust deep Gaussian processes (DGP)-based probabilistic load forecasting method using a limited number of data. Since the proposed method only requires a limited number of training samples for load forecasting, it allows us to deal with extreme scenarios that cause short-term load behavior changes. In particular, the load forecasting at the beginning of abnormal event is cast as a regression problem with limited training samples and solved by double stochastic variational inference DGP. The mobility data are also utilized to deal with the uncertainties and pattern changes and enhance the flexibility of the forecasting model. The proposed method can quantify the uncertainties of load forecasting outcomes, which would be essential under uncertain inputs.

M. Imani et al. in [2] Combination of long short-term memory (LSTM) network with support vector regression (SVR) is proposed for short-term electrical load forecasting in this paper. Three different cases are introduced and assessed for load forecasting. Although, LSTM can be lonely used for both feature extraction and forecasting, but, the features extracted by LSTM can be used as input of a SVR for forecasting. The experimental results show that using the load features and the temperature features extracted by individual LSTM networks beside the original values of load and temperature measurements of 24 recent hours provide good forecasting results.

N. J. Johannesen et al. [3] explores the use of regression tool for regional electric load forecasting by correlating lower distinctive categorical level (season and day of the week) and weather parameters. The historical electrical load datasets with meteorological parameters are available for the Sydney region and they have been used to test the regression tools. Data correlation over seasonal changes have been argued by means of improving Mean Absolute Percentage Error(MAPE). By examining the structure of various regressors they are compared for the lowest MAPE. The regressors show good MAPE for short term (30 min) prediction and Random Forest Regressor scores best in the range of 1-2 % MAPE.

Liangzhi Li et al. in [4] successfully transform the numerical prediction problem into an image processing task, and, based on that, utilize the state-of-the-art deep learning methods, which have been widely used in computer image area, to perform the electrical load forecasting. A novel deep learning based short-term forecasting (DLSF) method is proposed in the paper. Our method can perform accurate clustering on the input data using a deep Convolutional Neural Network (CNN) model. And ultimately, another neural network with three hiddenlayers is used to predict the electric load, considering various external influencing factors, e.g. temperature, humidity, wind speed, etc. Experimental results demonstrate that the proposed DLSF method performs well in both accuracy and efficiency.

A Fard, Haider Sawet & Faroogh Mohammadnia [5] in 2016 have developed a comparative study of various evolutionary short-term forecasting based on ANN. In this algorithm authors have utilized real data of load of Iran province. In this study author has generated a hybrid combination of evolutionary algorithms and ANN to forecast load. In this study, the most optimum results are obtained using modified Honey Bee Optimization(MHBMO) – ANN combination having load forecast with about 1.8%.

Kishan Bhushan Sahay, Suneet Sahu, Pragya Singh [6] in 2016 prepared a model to STLF for Toronto Canada using ANN. In ANN, they have utilized Back Propagation algorithm for training Neural Network. Also in BPNN three different algorithms LM, SCG and BR are used and separate results are generated based on above all three. On comparison MAPE of all three algorithmsauthors has concluded that out of all three back propagation algorithms LM and BR showing almost same results hence should not be used in forecasting load for short term.

Sharad Kumar, Shashank Mishra and Shashank Gupta [7] in 2016 developed two forecasting models ANN based and multiple regression based. Input data both weather and load data are sampled every 30 minutes. Author utilized data from 1st June 2015 to 15th June 2015 for training and testing. So, a total of 672 samples are available to develop model. Both models are separately trained and tested and their results are compared based on the amount of relevance the forecasted load is with actual value. In present study author found that ANN is better performing model than regression hence it should be preferred for forecasting short term load.

Victor Mayrink and Henrique S. Hippert [8] in 2016 have generated a hybrid model combining a classical and a machine learning algorithm for the forecasting of short term load. Two-time series are utilized in this study by author are load and temperature of Rio De Jenerio of Brasil dated between 1 Jan 1996 to 28 Feb 1997. Data utilized are hourly data. Two models utilized are exponential smoothing and gradient boosting. In this study for forecasting residuals of all previous steps in all iterations are refined using a base learning model. Author has proposed that the hybrid model has shown a significant improvement in results than the classical approach of exponential learning.

Ni Ding, Clémentine Benoit, Guillaume Foggia, Yvon Bésanger and Frédéric Wurtz [9] in 2015 have developed different models of machine learning for the forecasting of load. The authors in this study have utilized real data of French distribution system. The results obtained have proven that machine learning is far more accurate than time series forecasting. Author has utilized statistical methods for the purpose of variable selection. Also, the hidden layer neurons are taken as random parameters for all models since no exact method is available for deciding its exact value. In this paper, a significant improvement is MAPE is proved from naïve model by ANN for same data.

Penghua Li, Yinguo Li, Qingyu Xiong, Yi Chai and Yi Zhang [10] in 2014 have developed a hybrid quantized Elman Neural Network for the purpose of STLF. In this study author discussed the requirement of highly accurate forecasting model. According to author for every 1 % increase in forecasting error will lead to 10 million dollars increase in operating cost, which is sufficient enough motivation for forecasting. In this paper, algorithm proposed is unlike typical Elman Back propagation algorithm in which we have fixed context layer weights. Here author has now extended to hidden layer weights also thus proving to give more accurate results by understanding relation in time series in a much accurate way.

III. Evolutionary Algorithms

Evolutionary algorithms try to mimic the human attributes of thinking which are:

1) Parallel data processing
2) Self-Organization
3) Learning from experiences

Some of the commonly used techniques are discussed below:

1) Statistical Regression: These techniques are based on the time series approach based on the fitting problem that accurately fits the data set at hand. The approach generally uses the auto-regressive models and means statistical measures. They can be further classified as:

a) Linear
b) Non-Linear

Mathematically:
Let the time series data set be expressed as:

$Y = \{Y1, Y2 \ldots \ldots \ldots \ldots \ldots \ldots . . Yt)$

Here,
Y represents the data set
t represents the number of samples
Let the lags in the data be expressed as the consecutive differences.
The first lag is given by:

$\Delta Y_1 = Y_{t-1}$

Similarly, the j^{th} lag is given by:

$\Delta Y_j = Y_{t-j}$

2) Correlation based fitting of time series data: The correlation based approaches try to fit the data based on the correlation among the individual lags. Mathematically it can be given by:

$A_t = corr(Y_t, Y_{t-1})$

Here,
Corr represents the auto-correlation (which is also called the serial correlation)
Y_t is the t^{th} lagged value
Y_{t-1} is the $(t-1)^{st}$ lagged value
The mathematical expression for the correlation is given by

$corr(Y_t, Y_{t-1}) = \frac{conv(Y_t, Y_{t-1})}{\sqrt{varY_t, varY_{t-1}}}$

Here,
Conv represents convolution given by:

$conv\{x(t), h(t)\} = \int_{t=1}^{\infty} x(\vartheta)h(t-\vartheta)d\vartheta$

Here,
ϑ is a dummy shifting variable for the entire span of the time series data
t represents time
Y_t is the t^{th} lagged value
Y_{t-1} is the $(t-1)^{st}$ lagged value
X is function 1
H is function 2
Var represents the variance given by:

$var(X) = X_i - E(X)$

Here,
Xi is the random variable sample
E represents the expectation or mean of the random variable X

3) Finite Distribution Lag Model (FDL): This model tries to design a finite distribution model comprising of lags fitted to some distribution such as the normal or lognormal distributions. Mathematically:

$Y_t = \alpha_t + \delta_1 z_1 + \cdots \ldots \ldots \ldots \ldots . \delta_t z_t + \mu_t$

Here,
Y_t is the time series data set
α_t is a time dependent variable
δ_1 is a time-varying co-efficient
z is the variable (time variable)
t is the time index
μ_t is the time dependent combination-coefficient

4) Artificial Neural Networks (ANN): In this approach, the time series data is fed to a neural network resembling the working of the human based brain architecture with a self-organizing memory technique.

The approach uses the ANN and works by training and testing the datasets required for the same. The general rule of the thumb is that 70% of the data is used for training and 30% is used for testing. The neural network can work on the fundamental properties or attributes of the human brain i.e. parallel structure and adaptive self-organizing learning ability. Mathematically, the neural network is governed by the following expression:

$$Y = \sum_{i=1}^{n} X_i . W_i + \theta_i$$

Here,
X_i represents the parallel data streams
W_i represents the weights
θ represents the bias or decision logic
The second point is critically important owing to the fact that the data in time series problems such as sales forecasting may follow a highly non-correlative pattern and pattern recognition in such a data set can be difficult. Mathematically:

$$x = f(t)$$

Here,
x is the function
t is the time variable.
The relation f is often difficult to find being highly random in nature.
The neural network tries to find the relation f given the data set (D) for a functional dependence of x(t).
The data is fed to the neural network as training data and then the neural network is tested on the grounds of future data prediction. The actual outputs (targets) are then compared with the predicted data (output) to find the errors in prediction. Such a training-testing rule is associated for neural network. The conceptual mathematical architecture for neural networks is shown in the figure below where the input data is x and fed to the neural network.

IV. Evaluation Parameters

Since errors can be both negative and positive in polarity, therefore its immaterial to consider errors with signs which may lead to cancellation and hence inaccurate evaluation of errors. Therefore we consider mean square error and mean absolute percentage errors for evaluation. Mean Square Error is defined as:

$$MSE = [\sum_{i=1}^{n} (X - X')^2]/n$$

Mean Absolute Percentage Error is defined as:

$$MAPE = [\sum_{i=1}^{n} (X - X')/X']/n \times 100\%$$

Here, X is the predicted value and X' is the actual value and n is the number of samples.

Conclusion:

It can be concluded form the above discussions that the power management system is the most dependable source of power. This system needs to be very robust and effective in managing the generated power and supplying it properly to the required units. A good forecasting method for electrical loads will ensure that the power management stations utilize the supply of electricity to its optimal capacity. The balance of supply and demand of the electricity shall be met properly and optimally with this concept. This paper presents the various recent approaches and their salient points.

References

[1] D. Cao et al., "Robust Deep Gaussian Process-based Probabilistic Electrical Load Forecasting against Anomalous Events", in IEEE Transactions on Industrial Informatics, 2022, vol. 18, no. 2, pp. 1142-1153.
[2] M. Imani, "Long Short-Term Memory Network and Support Vector Regression for Electrical Load Forecasting," 2019 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET), 2019, pp. 1-6.
[3] N. J. Johannesen, M. Kolhe and M. Goodwin, "Comparison of Regression Tools for Regional Electric Load Forecasting," 2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech), 2018, pp. 1-6.
[4] L. Li, K. Ota and M. Dong, "Everything is Image: CNN-based Short-Term Electrical Load Forecasting for Smart Grid," 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC), 2017, pp. 344-351
[5] Abdollah KavousiFard, Haidar Samet, Foroogh Mohammadnia, "Evolutionary Short-Term Load Forecasting Based on ANN: A Comparison with the

Most Well-known Algorithms", 24th Iranian Conference on Electrical Engineering (ICEE), IEEE, 2016.
[6] Victor Mayrink, Henrique S. Hippert, "A Hybrid Method Using Exponential Smoothing and Gradient Boosting for Electrical Short-TermLoad Forecasting", 2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI), IEEE, 2016.
[7] Ni Ding, Clémentine Benoit, Guillaume Foggia, Yvon Bésanger, and Frédéric Wurtz, "Neural Network-Based Model Design for Short-Term Load Forecast in Distribution Systems", IEEE Transactions on Power Systems, IEEE, Volume: 31, Issue: 1, pp 72 – 81, 2015.
[8]Penghua Li, Yinguo Li, QingyuXiong, Yi Chai, Yi Zhang, "Application of a hybrid quantized Elman neural network in short-term load forecasting" ,International Journal of Electrical Power & Energy Systems, Elsevier,Volume 55, pp 749-759, 2014.
[9]Hao Quan, Dipti Srinivasan, Abbas Khosravi ,"Uncertainty handling using neural network-based prediction intervals for electrical load forecasting", Energy, Elsevier,Volume 73, pp 916-925, 2014.
[10] Ajay Gupta and Pradeepta K. Sarangi, "Electrical Load Forecasting Using Genetic Algorithm Based Back- Propagation Method", ARPN Journal of Engineering and Applied Sciences, Vol. 7, No. 8, 2012.
[11]Hao Quan, Dipti Srinivasan, Abbas Khosravi,"Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals" , IEEE Transactions on Neural Networks and Learning Systems,IEEE, Volume: 25, Issue: 2, pp 303 – 315, 2013.
[12]Sharad Kumar, Shashank Mishra, Shashank Gupta, "Short Term Load Forecasting Using ANN and Multiple Linear Regression", Second International Conference on Computational Intelligence & Communication Technology, IEEE, 2016.
[13]Kishan Bhushan Sahay, SuneetSahu, Pragya Singh, "Short-Term Load Forecasting of Toronto Canada by Using Different ANN Algorithms", IEEE 6th International Conference on Power Systems (ICPS), 2016.
[14]H.L. Willis. Spatial Electric Load Forecasting. Marcel Dekker, New York, 1996.
[11] Renuka Achanta, "Long Term Electric Load Forecasting using Neural Networks and Support Vector Machines",International Journal of Computer Science and Technology, Vol. 3, Issue 1, pp 266-269, 2012.
[15] Anand Mohan, "Mid Term Electrical Load Forecasting for State of Himachal Pradesh Using Different Weather Conditions via ANN", Model International Journal of Research in Management, Science & Technology, Vol. 1, No. 2, pp 60-63, 2013.
[16] S.S. Somam, H. Zareipour, O. Malik, and P. Mandal, "A Review of Wind Power and Wind Speed Forecasting Methods with Different Time Horizons". North American Power Symposium, Arlington, pp. 26-28, 2010.