Mid Term Daily Load Forecasting using ARIMA, Wavelet-ARIMA and Machine Learning

Akshita Gupta
Research Scholar
HRED, Indian Institute of Technology
Roorkee, India
26akshita@gmail.com

Arun Kumar
Professor
HRED, Indian Institute of Technology
Roorkee, India
arun.kumar@hre.iitr.ac.in

***Abstract*— The load forecasting is one of the important topics of discussion while studying the power system. The importance of load forecasting becomes even more important with the expansion of the horizon of the power system operation, may it be with the changing climatic patterns or increasing uncertain renewable penetration in the grid. This paper takes up the mid-term daily load forecasting using the three models, namely, ARIMA, Wavelet-ARIMA and Machine Learning. The first two models are time series based, while the third one is the application of the Artificial Intelligence and combines the past data along with the climatic patterns. The Wavelet decomposition is performed in order to show the effect of decomposition on the forecasting of time series and to check the performance of the different discrete wavelets. The machine learning uses averaging and boosting ensemble methods in order to combine the single regression techniques. The results show that the performance of machine leaning algorithm was found to be better than the time-series algorithms, thus providing the idea that inclusion of climatic factors is very important in the load forecasting models. Moreover, the averaging type of ensemble models were found to perform better for most of the months than the boosting type. Thus, the best forecasting accuracy is obtained by combining the climatic patterns in the study as well as using the ensemble models.**

Keywords— ARIMA, Wavelet-ARIMA, Daily Load Forecasting, Machine Learning

I. INTRODUCTION

The study of power system planning, operation and reliability is incomplete without the study of Load forecasting. Load forecasting is an essential and important part of power systems study and depending upon the type of study to be undertaken, can be classified on the basis of time horizon as short-term load forecasting (STLF), mid-term load forecasting (MTLF) and long-term load forecasting (LTLF). The STLF is important for the studies involving economic dispatch, unit commitment and the real time control of power system and they are ranged in between an hour to one week. The MTLF is generally beneficial for the utilities in order to plan the purchase of fuels enough for the proper functioning of the power system and used to calculate the changes in the tariff based upon the expenditure. The duration of this ranges from as low as one month to five years [1]. The last one known as the LTLF is typically used for planning the size and type of the generating plants, transmission lines etc., minimizing the variable as well as the fixed cost. The general time frame for this ranges from five to 20 years. The vast application of load forecasting in various stages of power system makes it an important subject [2]. And thus, an inaccurate forecasting could cause a severe impact on the reliability and cost of operation of the power system, along with the cost of production. This becomes more important in the scenario of high market competition in the deregulated electricity market, along with the changing climatic variations.

There are a number of techniques that have been applied in the field of load forecasting in the past decade. The regression models are the most commonly employed ones. The other effectively applied techniques include various forms of Artificial Intelligence (AI) [3, 4], time series models like Auto-regressive Moving Average (ARMA) and their modified form [5-8], Particle Swarm Optimization (PSO) [9, 10]. These techniques are found to be very effective. Also, the development of the hybrid techniques[11, 12], combining the different individual techniques have also been explored in the literature. But the major drawback that all the traditional time series models experience is the lack of accountability of the climatic parameters which are very important in the today's era of climate change. The parameters like temperature, humidity, precipitation plays a major role in the load forecasting [13]. Another aspect that has been explored in the literature is the decomposition [14, 15] of the load time series in the low and high frequency components employing Kalman filters, wavelet transform etc. The combination of time-series analysis and decomposition together has produced some good forecasting in certain application with reduced computational cost. The major criteria for the selection of the load forecasting technique are the horizon of the forecast, the computational time and the aim of the forecasting.

This work presents a study for the MTLF using the traditional time-series model, ARIMA, Wavelet-ARIMA model and ensemble regression models in Machine Learning, with an objective to predict the daily load for a case area. The study analyses different discrete wavelets for the decomposition of the daily time series and then applying ARIMA model for forecasting. Apart from these in order to give an outlook of how weather parameters affect the load forecasting, ensemble models have been applied in Machine Learning.

The study compares the time-series and data-driven techniques and the effect of wavelet decomposition on the forecast. Also, the in-depth analysis of the discrete wavelets is obtained from the study which could be used in similar type of forecasting study.

II. METHODOLOGY

This section presents the brief description of the techniques employed in the study namely, ARIMA, wavelet decomposition and Machine Learning.

A. Auto-Regressive Integrated Moving Average (ARIMA) model

ARIMA model is obtained by combining autoregressive and moving average models. This model has been widely applied and tested for different types of time-series [16, 17].

Ministry of Human Resources and Development (MHRD), India

978-1-7281-7455-6/20/$31.00 ©2020 IEEE

Most of the real-world data consist of seasonal time series and thus the modelling of seasonal time series besides non-seasonal series is also required to be discussed. The seasonal time series modelling is known as multiplicative model and is defined as per Equation 1.

$$(1-\phi_1 B^x - \phi_2 B^{2x} - \ldots - \phi_P B^{Px})(1-\varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)(1-B^x)^D (1-B)^d Z_t =$$
$$(1-\Theta_1 B^x - \Theta_2 B^{2x} - \ldots - \Theta_Q B^{Qx})(1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)\varepsilon_t \quad (1)$$

where is random variable, x is periodic term, B is the difference operator given as B(Zt) = Zt-1, (1-Bx)D is Dth seasonal difference of x, (1-B)d is dth non-seasonal difference, p is the order of non-seasonal autoregressive model, q is the order of non-seasonal moving average model, P is the order of seasonal autoregressive model, Q is the order of seasonal moving average model, is the parameter of non-seasonal autoregressive model, is the parameter of non-seasonal moving average model, is the parameter of seasonal autoregressive model, and is the seasonal moving average model [7, 18]. The determination of the order of AR and MA terms in ARIMA model is very important and is performed using Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) curves. [19].

B. *Wavelet Decomposition*

The Wavelet Transform has a high resolution in both the frequency- and the time-domain. It does not only tell us which frequencies are present in a signal, but also at which time these frequencies have occurred. Wavelet decomposition helps in gathering information by breaking the signal in scaled and the shifted forms of the mother wavelet. As the decomposition is carried out, original signal tends to become lower in resolution. The two types of filters are there for the decomposition or reconstruction of the signal, namely, low-pass and high-pass filters. These both filters form a system. The original signal is passed to each filter, decomposing it into detailed coefficient (D1) and approximated coefficient (A1). The coefficient D1 is obtained using the high-pass filter and contains high-frequency component which represents the short-term period pattern. The approximated coefficient A1, on the other hand is obtained using the low-pass filter and contains a low-frequency component, representing the long-term period pattern. The level of decomposition decides whether further decomposition takes place or not. If decomposition level is two or more, then only the approximate coefficient is decomposed. The approximate coefficient A1 is decomposed in A2 and D2. The same sequence is repeated in the following levels of decomposition. The maximum decomposition can take place till the individual details consists of a single sample or pixel. This is illustrated in Fig.1 for the chirp signal.

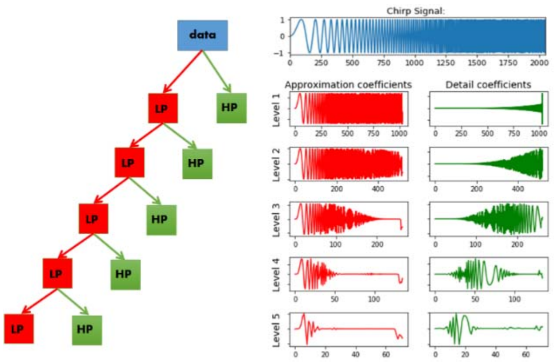

Fig.1. The approximation and detail coefficients for chirp signal

The wavelet-decomposed components can be reconstructed without any loss of data and the process is called as reconstruction of wave[20]. The reconstruction of the original signal can be achieved from the coefficients of the details and approximations. During the process of reconstruction only the coefficient vectors are required, which can be produced by down sampling the length of the signal by half.

The Wavelet Transform is divided in two different categories: Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT). The wavelet transform W(a,b) of signal f(x), with respect to mother wavelet f(x), is described in (2).

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\psi^* \left(\frac{t-b}{a}\right) dt \quad (2)$$

where a and b are real numbers, and * denotes complex conjugate. The parameters a and b are known as scale and translation parameter and controls the spread of the wavelet, and determines its central position respectively.

The coefficient, W(a,b) represents how the original signal f(x) and the scaled/translated mother wavelet match. Thus, for all a,b associated with a particular signal, the set of all wavelet coefficients W(a,b) is the wavelet representation of the signal with respect to the mother wavelet. The CWT provides a redundant representation of the signal in the sense that the entire support of W(a,b) need not be used to recover f(t). Instead of using that approach, the mother wavelet can be scaled and translated using certain scales and positions usually based on powers of two. This is known as DWT and this scheme is more efficient as well as accurate than CWT [21]. The DWT is described in (3).

$$DWT(m,k) = \frac{1}{\sqrt{a_0^m}} \sum_n f(n)\psi\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \quad (3)$$

where a,b are same in (2) and can be expressed in m. In (3), k is an integer variable that refers to a particular point of the input signal, and n is the discrete time index.

C. *Machine Learning*

Machine learning is one of the many applications of artificial intelligence (AI) which provides the systems with an ability to learn automatically and improve from experiences without being explicitly programmed. The focus of machine learning is on developing computer programs that access data and use it to learn for themselves. The process of learning begins with observations of data, looking for the patterns in data and making better decisions in the future, based on the real time data. The primary aim is to allow the computers to learn automatically without human intervention or assistance and adjust actions accordingly [22].

The ML algorithms are generally classified into two categories: Regression and Classification. For the prediction of data, regression models are used and the basic regression model available is Multiple Linear Regression model. A Multiple Linear Regression fits a linear model with coefficients w = (w1, …, wp) in order to minimize the residual sum of squares between the observed values in the dataset, and the targets are then predicted by the linear approximation.

Other advanced and accurate techniques that are used are ensemble methods with the basic aim to combine the predictions of several base estimators that are built with given learning algorithms so as to improve generalizability as well as the robustness over a single estimator. The ensemble methods are usually of two types and are summarized as below:

1. Averaging methods: The driving principle for these methods are building of several estimators independently and then to average their predictions. For example, Bagging methods, and Forests of randomized trees, etc.

2. Boosting methods: The base estimators are built sequentially and the aim is to reduce the bias of the combined estimator. The motivation is to combine several weak models to produce a powerful ensemble. For example, AdaBoost, and Gradient Tree Boosting, etc.

III. Results

The three models have been applied on the state of Karnataka in India. The state is located in the southern part of the country as shown in Fig.2 and experiences a good potential of solar energy.

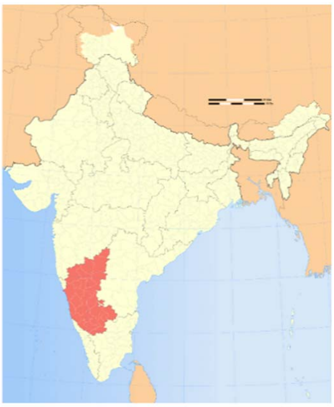

Fig.2. The index map of Karnataka

The load in the state comprises of about 20 % of domestic demand, 20% of commercial demand, 15% of agriculture demand and 40% of industrial demand. The daily data that has been used in the study have been listed below:

1. Load data for 10 years from 2008-2017.
2. For the machine learning model data for minimum temperature, maximum temperature, humidity, precipitation and cloud cover for 6 years from 2012-2017

Apart from this the Day of Month and Day of week have also been taken into account in Machine learning model. In the developed algorithm different Machine learning models have been used to predict the daily load values. The developed model is programmed in Python package and the computer used for the computations had an Intel Core i7-6700 CPU @ 3.40GHz, 3408 Mhz, 4 Core(s), 8 Logical Processor. The different models used in the study are listed in Table 1 along with the abbreviations that are used in the rest of the paper. A total of 20 models have been applied for the daily inflow prediction covering both the averaging and the boosting methods.

The discrete wavelet decomposition is applied to the time series month-wise and 106 wavelets have been used for decomposing time-series for each month and then forecasting the decomposed components using the ARIMA model. The Root Mean Square Error (RMSE) of each wavelet-ARIMA model relative to ARIMA model is shown in Fig.3. The figure shows that for most of the wavelet-ARIMA combination the relative RMSE values are below 1.5. The 4 wavelets were found to be comparable to ARIMA model, namely, bior 1.1, bior 1.3, coif 8 and rbio1.1.

TABLE I. Different Models of Regression

S.No.	Model	Abbreviation
Single Regression		
1.	Multiple Linear Regression	MLR
2.	Gradient Boost Regression	GBR
3.	Random Forest Regression	RFR
4.	Extra-Tree Regression	ETR
5.	Ada-Boost Regression	ABR
Bagging Regression		
6.	MLR	BRLR
7.	GBR	BRGBR
8.	RFR	BRRFR
9.	ETR	BRETR
10.	ABR	BRABR
Voting Regression		
11.	MLR, GBR, RFR	VMGR
12.	MLR, GBR, ETR	VMGE
13.	MLR, GBR, ABR	VMGA
14.	MLR, RFR, ETR	VMRE
15.	MLR, RFR, ABR	VMRA
16.	MLR, ETR, ABR	VMEA
17.	GBR, RFR, ETR	VGRE
18.	GBR, RFR, ABR	VGRA
19.	GBR, ETR, ABR	VGEA
20.	RFR, ETR, ABR	VREA

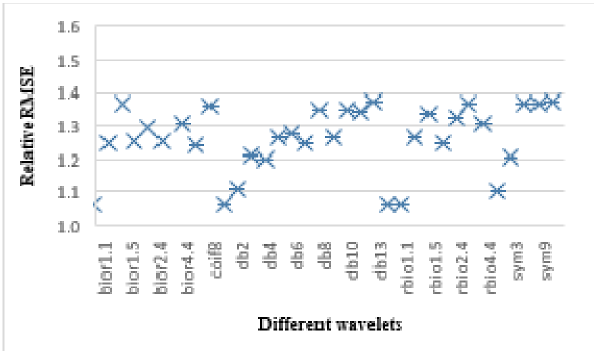

Fig.3. Relative RMSE of Wavelet-ARIMA model to ARIMA model

The ARIMA model is applied month-wise for daily prediction, with no monthly seasonality. The values of p, q, d for different month is shown in Table II.

TABLE II. Month-wise ARIMA model parameters

S.No.	Month	p	q	d
1	Jan	8	1	1
2	Feb	5	4	1
3	Mar	9	1	1
4	Apr	5	5	1
5	May	1	6	1
6	Jun	4	3	1
7	Jul	3	3	1
8	Aug	8	2	1
9	Sep	3	3	1
10	Oct	2	6	1
11	Nov	9	1	1
12	Dec	5	3	1

The developed algorithm combines the month-wise daily forecast for all the three models by combining the best wavelets for each month in case of wavelet-ARIMA model and best regression technique in case of Machine learning technique. The best Wavelet and machine learning technique for different months can be visualized from Table III. It can be observed that for most of the months the best wavelet-ARIMA model was found corresponding to bior1.1 wavelet. In case of Machine Learning the averaging methods were found to perform better for most of the months.

TABLE III. BEST WAVELET AND ML TECHNIQUES

S.No.	Month	Best wavelet relative to ARIMA	Best ML Technique
1	Jan	db2	MLR
2	Feb	bior1.3	MLR
3	Mar	sym17	BRGBR
4	Apr	bior1.1	VMGE
5	May	bior1.1	VREA
6	Jun	sym2	BRRFR
7	Jul	coif8	MLR
8	Aug	bior2.2	GBR
9	Sep	coif15	VGRE
10	Oct	bior1.1	VGRE
11	Nov	bior1.1	ABR
12	Dec	bior1.1	BRGBR

The comparison of the three techniques for the daily load forecasting can be seen from Fig. 3. It can be seen that the Machine Learning and the Wavelet- ARIMA formed by combining the best techniques listed in Table 2, closely follows the actual time series for a given year as compared to the forecasting using ARIMA model.

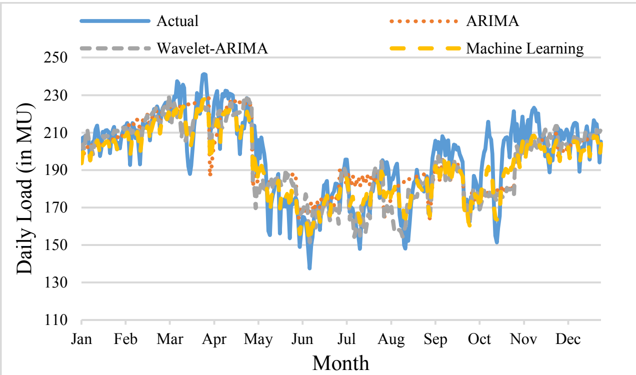

Fig. 3. Daily load forecast from the three models

The above findings can also be supported by the Relative percentage RMSE (RRMSE) values for the three techniques. The RRMSE value of 3.51 % for Machine Learning was found to be lowest out of three as compared to ARIMA and Wavelet-ARIMA with the value of 5.72% and 4.36 % respectively.

Apart from this the computational time is also very important apart from the accuracy. The ARIMA model takes the least time compared to Wavelet-ARIMA and ML model. The Wavelet-ARIMA model takes the maximum time for all the wavelets but once the best wavelet is identified it only needs to be used. The ML takes longer time for the single iteration as well as the storage requirement is higher on account of the larger number of predictors.

IV. CONCLUSION

The load forecasting has always been the subject of great importance for power system planning and operation. The importance becomes even more important in today's era of climate change and increasing penetration of renewable energy in the grid. This study takes up the daily load forecasting problem using the time-series modelling and the Artificial Intelligence method. The time-series modelling has been performed using ARIMA model and to investigate the effect of time-series decomposition on the forecasting values, wavelet decomposition has been applied. The wavelet decomposition uses different discrete wavelets to decompose the time-series, followed by ARIMA model for forecasting the daily load values. The results show that the Wavelet-ARIMA models were found to perform better than the ARIMA model. The best wavelet out of 108 wavelets for maximum number of months was bior1.1. The decomposition decreases the forecasting error of ARIMA as well as the follows the actual trend better.

The change in the climatic conditions in the recent years has made it an important factor to be included in the load forecasting studies. In this study the load forecasting using regression techniques in Machine learning has been done taking the climatic parameters, namely temperature, humidity, precipitation and cloud cover as the predictors. The ensemble models combining the different regression techniques have also been applied in this study. The averaging type techniques were found performing better than the boosting type. Further, the forecasting done using the climatic parameters gives better results as compared to the time-series analysis, which only takes into account the inherent pattern in the time-series. The forecasting done using Machine Learning is also able to follow the actual load profile variations better as compared to the time-series analysis. Thus, the study shows that the time-series decomposition increases the accuracy of forecasting. The inclusion of climatic factors makes the forecasting much better but more data intensive.

ACKNOWLEDGMENT

The first author would extend a sincere gratitude to the Ministry of Human Resources and Development (MHRD) for providing the financial and technical support in order to carry out the research work at the Indian Institute of Technology (IIT), Roorkee.

REFERENCES

[1] E. A. Feinberg and D. Genethliou, "Load Forecasting," in Applied Mathematics for Restructured Electric Power Systems. Power Electronics and Power Systems, C. J.H., W. F.F., and M. J., Eds., ed Boston, MA: Springer, 2005, pp. 266-285.

[2] W. Sun and Y. Bai, "Short-term load forecasting based on wavelet transform and BP neural network," presented at the Second International Conference on Mechanic Automation and Control Engineering, China, 2011.

[3] S. R and H. A. Abdulqader, "Load Forecating for Power System Planning and Operation using Artificial Neural Network at Al Batinah region oman," Journal of Engineering Science and Technology, vol. 7, pp. 498-504, 2012.

[4] T. Senjyu, H. Takara, K. Uezato, and Toshihisa Funabashi, "One-Hour-Ahead Load Forecasting Using Neural Network," IEEE Transactions on Power Systems, vol. 17, pp. 113-118, 2011.

[5] K. Goswami, A. Ganguly, and A. K. Sil, "Day Ahead Forecasting and Peak Load Management using Multivariate Auto Regression Technique," in IEEE Applied Signal Processing Conference (ASPCON), 2018, pp. 279-282.

[6] S. Maldonado, A. Gonzalez, and S. Crone, "Automatic time series analysis for electric load forecasting via support vector regression," Applied Soft Computing Journal, vol. 83, 2019.

[7] M. Valipour, M. E. Banihabib, and S. M. R. Behbahani, "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir," Journal of Hydrology, vol. 476, pp. 433-441, 2013.

[8] P. Li, J. Zhang, C. Li, B. Zhou, Y. Zhang, M. Zhu, et al., "Dynamic Similar Sub-Series Selection Method for Time Series Forecasting," IEEE Access, vol. 6, pp. 32532-32542, 2018.

[9] C. M. Huang, C. J. Huang, and M. L. Wang, "A Particle Swarm Optimization to Identifying the ARMAX Model for Short-Term Load Forecasting," IEEE Transactions on Power Systems, vol. 20, pp. 1126-1133, 2005.

[10] Y. K. Semero, J. Zhang, and D. Zheng, "EMD–PSO–ANFIS-based hybrid approach for short-term load forecasting in microgrids," IET Generation, Transmission & Distribution, vol. 14, pp. 470-475, 2020.

[11] M. Hanmandlu and B. K. Chauhan, "Load Forecasting Using Hybrid Models," IEEE Transactions on Power Systems, vol. 26, pp. 20-29, 2011.

[12] Z. Zhang, W.-C. Hong, and J. Li, "Electric Load Forecasting by Hybrid Self-Recurrent Support Vector Regression Model With Variational Mode Decomposition and Improved Cuckoo Search Algorithm," IEEE Access, vol. 8, pp. 14642-14658, 2020.

[13] B. Li, J. Zhang, Y. He, and Y. Wang, "Short-Term Load-Forecasting Method Based on Wavelet Decomposition With Second-Order Gray Neural Network Model Combined With ADF Test," IEEE Access, vol. 5, pp. 16324-16331, 2017.

[14] L. Qiuyu, C. Qiuna, L. Sijie, Y. Yun, Y. Binjie, W. Yang, et al., "Short-term Load Forecasting Based on Load Decomposition and Numerical Weather Forecast," presented at the IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 2017.

[15] L. Alfieri and P. De Falco, "Wavelet-Based Decompositions in Probabilistic Load Forecasting," IEEE Transactions on Smart Grid, vol. 11, pp. 1367-1376, 2020.

[16] G. E. P. Box and G. Jenkins, Time Series Analysis, Forecasting and Control: Holden-Day, Inc., 1990.

[17] J. D. Salas, J. W. Delleur, and V. Yevjevich, Applied Modeling of Hydrologic Time Series: Water Resources Publications, 1985.

[18] M. Karamouz and S. Araghinejad, Advance Hydrology: Amirkabir University of Technology Press., 2012.

[19] J. D. Cryer and K. Chan, Time Series Analysis: With Applications in R. New York: Springer-Verlag New York, 2008.

[20] B.-g. Koo and J. P. Heung-seok Lee, "A Study on Short-term Electric Load Forecasting using Wavelet transform," in 5th IEEE PES Innovative Smart Grid Technologies Europe, Istanbul, October 12-15.

[21] A. S. Pandey, D. Singh, and S. K. Sinha, "Intelligent Hybrid Wavelet Models for Short-Term Load Forecasting," IEEE Transactions on Power Systems, vol. 25, pp. 1266-1273, 2010.

[22] O. Theobald, Machine learning for absolute beginners : a plain English introduction, 2018.