Neural Network-based Load Forecasting and Error Implication for Short-term Horizon

Khuntia, S.R.; Rueda, J.L.; van der Meijden, M.A.M.M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Neural Network-based Load Forecasting and Error Implication for Short-term Horizon

S. R. Khuntia, *Graduate Student Member*,
J. L. Rueda, *Senior Member*
Department of Electrical Sustainable Energy
Delft University of Technology
The Netherlands
E-mail: s.r.khuntia@tudelft.nl

M. A. M. M. van der Meijden[1,2], *Member*
[1]Department of Electrical Sustainable Energy
Delft University of Technology
[2]TenneT TSO B.V.
The Netherlands

Abstract—Load forecasting is considered vital along with many other important entities required for assessing the reliability of power system. Thus, the primary concern is not to forecast load with a novel model, rather to forecast load with the highest accuracy. Short-term load forecast accuracy is often hindered due to various load impacting factors. Two of the major impacting factors are day-ahead weather forecast and subsequent variation in electricity demand that is independent of weather. To tackle the uncertainty in short-term load forecasting, this paper presents a neural network-based load forecasting technique for short-term horizon based on data corresponding to a U.S. independent system operator. With the real life data, a better understanding of forecasting error is carried out while further identifying the time periods when the load is supposedly to be over- or under-forecast.

Keywords—Error analysis; forecasting; forecast error; load forecast uncertainty; neural network; short-term load forecast

I. INTRODUCTION

Accurate load forecasting is vital to both the electric industry as well as smooth operation and planning of a utility company. Load forecasting is not something new, and it dates back to late 1960s when the first paper on load forecasting techniques was published [1]. Today, load forecasting has become an integral part of planning for more than just utilities; system operators, energy suppliers, financial institutions, and participants in the generation, transmission, and distribution of electricity have a vested interest in load forecast accuracy. Based on time-scale [2], load forecast can be broadly classified into three main categories [3]:

- Short-Term Load Forecast (STLF): The time-period of STLF lasts for few minutes, hours to one-day ahead or a week. STLF aims at economic dispatch and optimal generator unit commitment, while addressing real-time control and security assessment.

- Mid-Term Load Forecast (MTLF): The time-period of MTLF is a month to a year or two. MTLF aims at maintenance scheduling, coordination of load dispatch and price settlement so that demand and generation is balanced.

- Long-Term Load Forecast (LTLF): The time-period of LTLF is few years to 10-20 years ahead. LTLF aims at system expansion planning, i.e., generation, transmission and distribution. In some cases, it also affects the purchase new generating units.

In recent years, significance of short-term load forecast (STLF) has increased and it will continue because over- or under-estimate of load demand has a significant impact on the efficiency of operation of any electrical utility. Various operational decisions such as economic scheduling of the generating capacity, scheduling of fuel purchase and system security assessment are based on such forecasts [4]. Literature study reveals that short-term load forecasting has been extensively studied, and many load modeling and forecasting techniques have been developed [5-7]. These methodologies can be broadly classified into two categories: 1) Traditional STLF methods like time series, regression analysis and gray model, which is based on load patterns. 2) Non-traditional or modern STLF methods like fuzzy, neural networks and other intelligent load forecasting methods.

The scope of this paper is forecasting electrical load in short-term horizon using neural networks. Neural network-based load forecasting is one of the most widely used non-traditional load forecasting methods, and it is evident from literature study [8-10]. Ref. [10] can be accessed for a complete review of different neural network-based methods developed lately. In short-term horizon, the challenge for planners is not only to forecast load but also to form an accurate picture of the day-ahead load. The day-ahead load includes point forecasts of the load in each hour and also acknowledges the precision, or lack hereof, associated with the forecasted value. This work addresses the uncertainty associated with day-ahead scheduling and error implication associated with this uncertainty.

An important question is: *What is the need of error and uncertainty analysis in short-term load forecasting*? One of the vital inputs in short-term load forecast is weather data. Thus, it is evident that for day-ahead load forecast, use of weather forecast is the most viable option. There are high chances that online operation of weather forecast model introduces the associated error into load forecasting model. Chen et al. [11] performed a comprehensive study showing the significant effect of weather on load forecast. It can also impact the training of neural network as discussed by Yoo and Pimmel

[12]. For our study, weather forecast error is not considered, rather a historical database of weather information is used to train the network. In this study, statistical properties of forecast errors are studied and performance analysis of neural network-based model is performed by making comparisons to the normal distribution.

The rest of the paper is organized as follows: Section II discusses about the neural network-based load forecasting methodology. Forecast results are demonstrated in Section III. Section IV presents an in-depth learning of error analysis and uncertainty associated with forecasted results. Finally, section V concludes the paper.

II. NEURAL NETWORK-BASED SHORT-TERM LOAD FORECAST

The use of neural networks has been a widely studied electric load forecasting technique since 1990s [13]. The learning property of neural network in solving nonlinear and complex problems called for its application to forecasting problems. This is due to the fact that instead of relying on explicit rules or mathematical functions, neural networks draw a link between input and output data. Thus, in comparison to other traditional and non-traditional models, neural networks holds a good promise for the purpose of load forecasting. For our study, historical data for years 2010-2013 was used to forecast day-ahead load for 2014. The various input parameters for training data is illustrated in Fig. 1. An exception in this study is that random disturbances, consumer class, and demand side management is not considered. This is because the primary aim of this work is to analyse error and uncertainty in forecasted value.
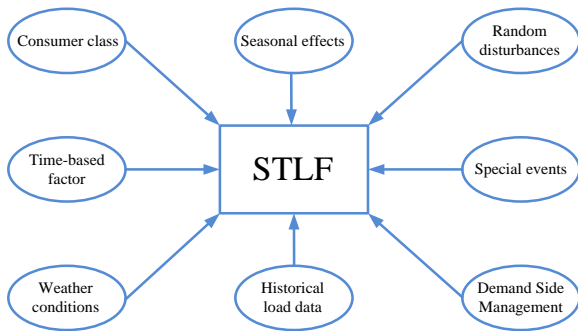

Fig. 1. Input parameters for training data [*Note*: *Consumer class, random disturbances and demand side management are excluded in this study*]

In our study, the Levenberg-Marquardt optimization technique is used as it has one of the best learning rates when compared to the other available functions in forecasting problems [14]. The proposed model is a three-layered feedforward backpropagation network, as shown in Fig. 2. The input layer comprises of 73 inputs, output layer is a 24-hour or day-ahead load forecast while choosing the hidden layer is tricky. The number of neurons in the hidden layer must be carefully chosen; too many neurons make the network overspecialized, leading to loss of generalizing capability. On the other hand, due to lack of enough hidden layer neurons, the network may find it difficult to learn the behavior of the series. For our study, the model was tested with varying number of hidden layer neurons, ranging from ten to fifteen. Fourteen neurons were finally used in our study because it offered a

better model characteristic. Similarly, the neural-network weights are adjusted based on a comparison of the output and the target, until the network output matches the target. There are no specific defined rule for optimal selection of network layers and neurons to produce the better forecast results. But to keep things simple, three-layered architecture was chosen. Networks with more than one hidden layer are generally more complex.
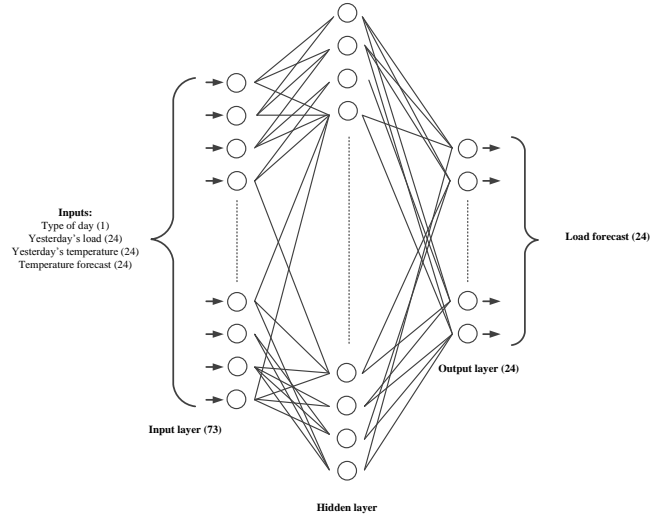

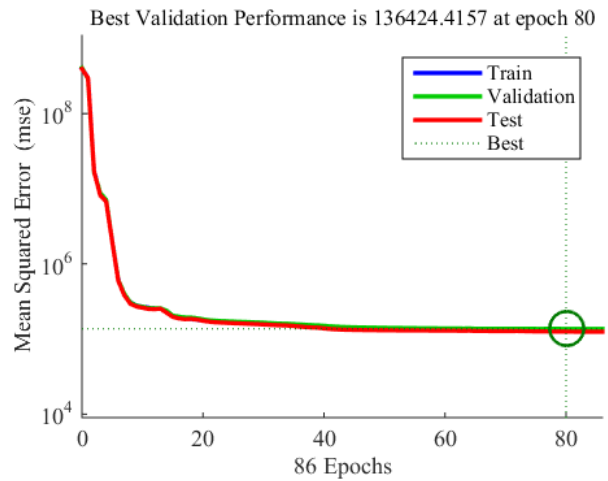Fig. 2. Neural network architecture for short-term load forecast

Fig. 3. Epoch training of neural network

Another important feature is training of network. Number of epochs is vital for network training. Generally, the error reduces after more epochs of training, but might start to increase on the validation data set as the network starts overfitting the training data. For our study, best validation performance was observed at epoch 80, as seen in Fig. 3. It can be deduced that the best performance is taken from the epoch with the lowest validation error, and it attained in less number of epochs.
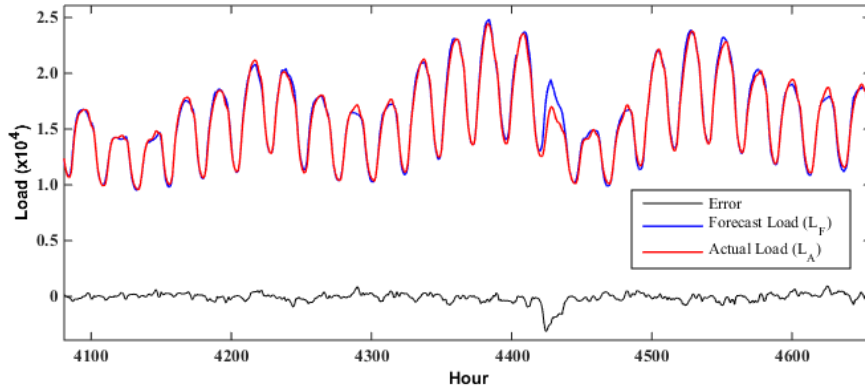
Fig. 4. Forecasted load (L_F), actual load (L_A) and error for date range (19/06/2014 – 14/07/2014)

III. FORECAST RESULTS

To assess the effectiveness of our proposed forecast methodology, the Neural Network toolbox in MATLAB [15] is used. The neural network data set is divided into two sets: first data set is used for training purpose of network and second data set is used for testing of forecasting results. First data set comprises of four years (2010-2013) 24-hourly load and weather data, that is used to train the network called training data set. Second data set comprises of targeted year (2014) 24-hourly load data, is considered as test data set. The forecasted results are compared with test data set, as shown in Fig. 4. The solid red curve shows the hourly average demand over the date range (19/06/2014 – 14/07/2014), while the blue curve shows the day-ahead load forecast for the same time period. A more detailed load profile showing ~~load~~ hourly load forecast for the same date range is illustrated in Fig. 5.

Fig. 5. Hourly load forecast for date range (19/06/2014 – 14/07/2014)

In order to model the statistical uncertainty information, large volumes of historical and real-time data are needed. As illustrated in Fig. 6, a sliding window is used for acquiring continuous statistical information on system load. The time-frame can be tuned accordingly depending on forecast need. The forecast resolution is the time interval between two consecutive data records. The time horizon is the length of the look-ahead interval, and the forecast update interval is the time

interval for updating the forecast. The structure is supported by a table of data requirements for STLF as shown in Table I.
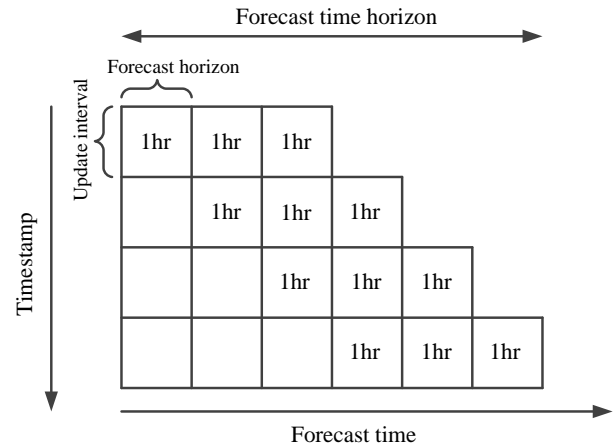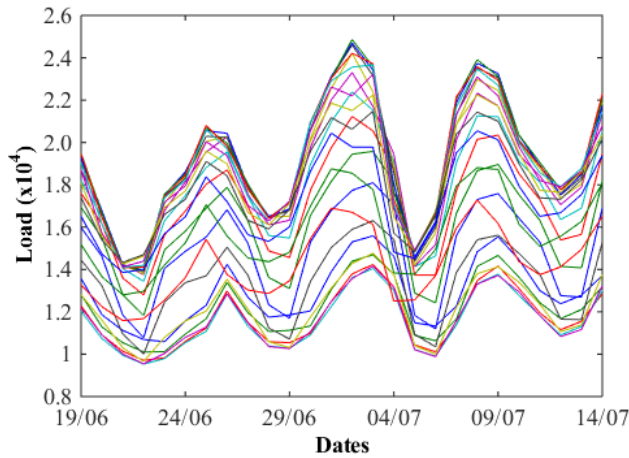

Fig. 6. Day-ahead load forecast structure

TABLE I. DATA REQUIREMENTS FOR STLF

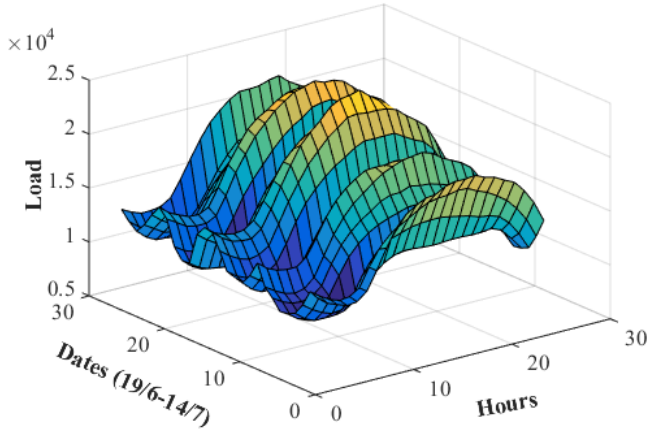| Data	Resolution	Forecast Horizon
Day-ahead load forecast	1hr	24hr
Hour-ahead load forecast	30min.	24hr
Real-time load forecast	5min.	1hr

Fig. 7. Load profile for date range (19/06/2014 – 14/07/2014)

Fig. 7 illustrates the load profile of selected dates. It depicts that, in working days the load demand is much higher than the off days (Saturday and Sunday) due to higher social activities. This weekly pattern is repeated more or less throughout the year.

IV. ERROR IMPLICATION

In STLF, accurate load forecasts are very important because they determine scheduling of generation units for next day or maybe few hours ahead. Any error in forecast results in suboptimal commitment of generation unit in day-ahead market, which avoided by utilities. Forecast error, as seen in Fig. 8, is defined as the difference between the actual load (L_A) and the forecasted load (L_F). It is evident from the figure that day-ahead load forecast error varies within the ±7% range. Few random overshoots in the range of +10% and -20% can be observed in the figure, which are excluded for error analysis. The uncertainty associated with the load forecast is one of the most influential factors influencing the resulting uncertainty.
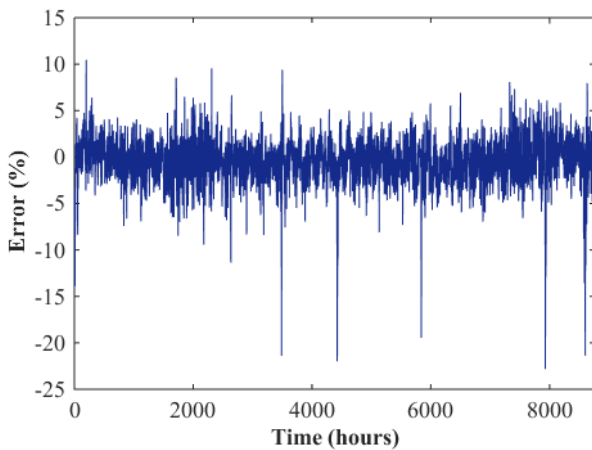

Fig. 8. Load forecast error (%)

The normal distribution is amongst the most common method to describe load forecasting errors. It can be described by the first two statistical moments, namely mean (μ) and variance (σ^2). The third and fourth moments, namely skewness and kurtosis, are often close to zero if the observed distribution is well represented by normal distribution. For error analysis and to check the efficiency of our neural network-based load forecasting method, two widely used performance metrics called Mean Absolute Percentage Error (MAPE) and Coefficient of Variation (CV) are used. Considering two time series, actual load $L_A(t)$ and forecasted load $L_F(t)$ for time-series $t = 1,2, ..., T$, the MAPE is defined as:

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{L_A(t) - L_F(t)}{L_A(t)}\right| \times 100\%$$

The coefficient of variation, also called relative standard deviation, measures the ratio of the forecasted error standard deviation (σ) to the signal mean (μ). It is defined as:

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(L_A(t) - L_F(t))^2}}{\frac{1}{T}\sum_{t=1}^{T} L_A(t)}$$

In our study, following results are obtained:

$$\mu = -61.89; \ \sigma = 351.35$$

$$MAPE \ (\%) = 1.74; CV = 5.67$$

Both, MAPE and CV are traditional relative error metrics traditionally reported in forecast-related literature although MAPE is more commonly reported. Following the traditional method of examining distributions, an error histogram is plotted as shown in Fig. 9. The histogram is plotted with bin size of 100 so as to exclude the unwanted distribution and focus on large forecast errors. Few observations from the histogram are:

- The dotted line shows a normal distribution with the same mean and standard deviation as the forecasted errors.

- The observed error distribution is more peaked, with narrower shoulders and larger tails than the normal distribution assumption would suggest.

- One of the most critical features of the observed distribution is the negative mean bias, represented by a mean value of 62MW.

- Histogram shows that load forecast errors are not normally distributed.

- The actual load forecast error distribution has more mass around zero than what a normal probability density function (PDF) would predict. A solution to this anomaly is choosing the logistic distribution, as it proved to fit the data better. The actual distribution used to model load forecast uncertainty is not crucial as long as it accurately represents the data.
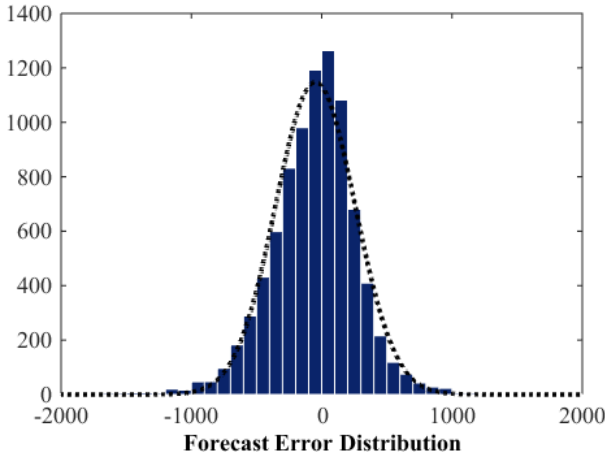
Fig. 9. Histogram and normal distribution fit of forecasted error with μ=-61.89 and σ=-351.35

In our error and uncertainty analysis, load forecast errors are summed for each dispatch interval in the past within a sliding window as shown in Fig. 6. The sliding window size is selected to collect sufficient statistical information regarding the forecast errors. The information can be accumulated separately for each forecast horizon; for instance, for the hour-ahead forecast, two hours ahead forecast, and so on. Based on the collected statistics, the approach evaluates the percentile intervals (also called confidence intervals or uncertainty ranges) for each forecast horizon and different level of confidence. These intervals are assumed to be the same in the future dispatch interval; that is, for the next hour, the hour after that, and so on. It is used, in later part, to analyse the truncated normal distribution of load forecast error.

Based on the results obtained, intensive approaches for the uncertainty analysis of forecast errors is carried out in this study following the work on wind forecast uncertainty [16]. The two approaches are described below:

1. *Distribution Fitting Approach*: Distribution fitting is fitting of probability distributions, and based on assumptions about a specific standard form of random variables; for example, normal, uniform or Poisson distributions. Based on the standard distributions and selected set of its parameters, they assign probability to the event that the random variable takes on a specific, discrete value, or falls within a specified range of continuous values. In this case, the error distribution is not a normal distribution as seen in Fig. 8. Hence, selecting a distribution model means choosing a standard probability distribution and then adjusting its parameters to fit the data. One solution to the above problem is fitting the error distribution data set with truncated normal distribution (TND). The PDF of such a truncated normal distribution is written as:

$$PDF(x; \mu, \sigma, a, b) = \frac{\frac{1}{\sigma} PDF_N\left(\frac{x-\mu}{\sigma}\right)}{CDF_N\left(\frac{b-\mu}{\sigma}\right) - CDF_N\left(\frac{a-\mu}{\sigma}\right)}$$

where, μ is the mean value of non-TND

σ is the standard deviation of the non-TND

a, b are upper and lower limits of the non-TND

$x \epsilon (a, b), -\infty \le a < b \le \infty$

$PDF_N(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$ is the PDF of standard normal distribution

$CDF_N(\cdot)$ is the CDF of standard normal distribution

The use of truncated normal distribution in wind forecast error has been reported in refs. [17-18]. One of the many reasons to use truncated normal distribution is because though normal distribution is a good fit for our forecasted error data, for physical reasons it is known that data can never be negative. Also, the values of a normal distribution can, in theory, assume any value over the range from $-\infty$ to $+\infty$, which may lead to significant computational errors in situations where the distribution's outcomes are constrained. For error analysis, it is desirable to consider data within a particular range of interest as per the planner, which we might symbolize as [A,B], or [A,+∞), or (-∞,B], depending on the truncation we apply. The large forecasting error distribution could be represented by extending the range of the truncated normal distribution function, i.e., a larger range than the 99.95% confidence interval of $[\mu - 3.5\sigma, \mu + 3.5\sigma]$ as studied in ref. [19]. Accordingly, the modified load forecast error, shown in Fig. 10., is represented by a truncated normal distribution, in which the mean is the hourly power forecast and the standard deviation is 5% of the mean:

$$f(x) = \begin{cases} 0, & x < \mu - 3.5\sigma \\ & \text{or } x > \mu + 3.5\mu \\ \frac{1}{\alpha \sqrt{2\pi \sigma}} . e^{-(x-\mu)^2 / 2\sigma^2}, & \mu - 3.5\sigma \le x \le \mu + 3.5\sigma \end{cases}$$

where, $\alpha = \int_{\mu - 3.5\sigma}^{\mu + 3.5\sigma} (1/\sqrt{2\pi \sigma}) . e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$
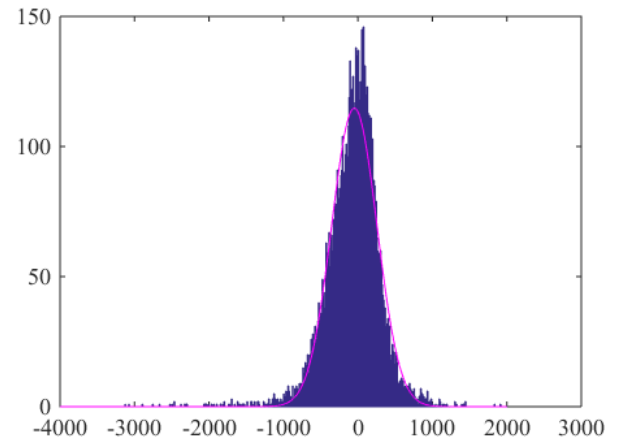

Fig. 10. Histogram of modified load forecast error

2. *Empirical Probability Approach*: Statistical analysis of load forecast error distribution using empirical probability approach is not studied extensively. Empirical probability is a type of non-parametric distribution that do not follow any standard probability distribution. In this approach, the models make no assumptions about the form of the underlying distribution, so no parameter estimates are needed [20]. An advantage of estimating probability distributions using the empirical modeling approach is that this procedure is relatively free of assumptions.

Compared to real-life with physical data, like load forecast error distribution, the empirical PDF is simply the histogram. Integrating the PDF using cumulative sum produces the empirical CDF. If a sample comes from a parametric distribution (such as a normal distribution), its empirical CDF will resemble the parametric distribution as in the case of load forecast error shown in Fig. 11. If not, the empirical distribution still gives an estimate of the CDF for the distribution. As seen in the figure, the continuous and stairs empirical CDF overlap on each other, thus proving that the distribution is not very diverse. The confidence intervals are spaced evenly from empirical CDF.
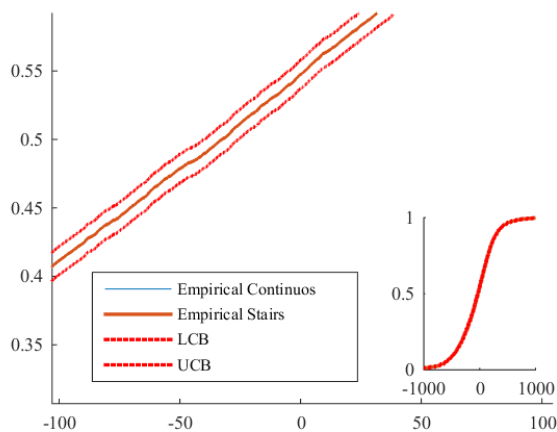

Fig. 11. Empirical histogram of forecasted error with μ=-61.89 and σ=-351.35 [*Sub-image is truncated to ±1000. LCB: Lower Confidence Bound and UCB: Upper Confidence Bound*]

V. Conclusion

A neural-network based load forecasting model was designed, implemented, trained with real data and results obtained with high degree of accuracy. A set of optimized weights and the associated biases after training the network from load data obtained from the power utility company were also obtained. The accuracy of the forecasts was verified by comparing the simulated outputs from the network. A single day was considered in demonstrating the accuracy of forecast model. Further, the forecasted load was analysed for error and uncertainty with various performance based statistical analysis tool. One important observation is that the error distribution does not follow a normal distribution. Hence, performance

analysis was performed and results indicate that truncated normal distribution is a more accurate means of modeling the distribution.

References

[1] G. Heinemann, D. Nordman, and E. Plant, "The relationship between summer weather and summer loads: A regression analysis," *IEEE Trans. Power App. Syst.,* vol. PAS-85, no. 11, pp. 1144-1154, 1966.

[2] S. R. Khuntia, B. W. Tuinema, J. L. Rueda, and M. A. M. M. van der Meijden, "Time-horizons in the planning and operation of transmission networks: an overview," *IET Generation, Transmission & Distribution*, vol. 10, no. 4, pp. 841-848, 2016.

[3] H. L. Willis, and J. E. Northcote-Green, "Spatial electric load forecasting: a tutorial review," *Proc. IEEE*, vol. 71, no. 2, pp. 232-253, 1983.

[4] M. De Felice, and X. Yao, "Short-term load forecasting with neural network ensembles: A comparative study [Application Notes]," *IEEE Computational Intelligence Magazine*, vol. 6, no. 3, pp. 47-56, 2011.

[5] G. George, and F. D. Galiana, "Short-term load forecasting," *Proc. IEEE*, vol. 75, no. 12, pp. 1558-1573, 1987.

[6] I. Moghram, and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Trans. Power Syst.,* vol. 4, no. 4, pp. 1484-1491, 1989.

[7] J. W. Taylor, and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Trans. Syst.*, vol. 22, no. 4, pp. 2213-2219, 2007.

[8] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44-55, 2001.

[9] Y. Rui, and A. A. El-Keib, "A review of ANN-based short-term load forecasting models," *Proc. Twenty-Seventh Southeastern Symp. Syst. Theory*, pp. 78-82, 1995.

[10] A. Baliyan, K. Gaurav, and S. K. Mishra, "A review of short term load forecasting using artificial neural network models," *Procedia Computer Science*, vol. 48, pp. 121-125, 2015.

[11] H. Chen, Y. Du, and J. N. Jiang, "Weather sensitive short-term load forecasting using knowledge-based ARX models", Presented at the IEEE Power Engineering Society General Meeting, pp 190-196, 2005.

[12] H. Yoo, and R. L. Pimmel, "Short-term load forecasting using a selfsupervised adaptive neural network," *IEEE Trans. Power Syst.*, vol. 14, no. 2, pp. 779-784, 1999.

[13] M. Peng, N. F. Hubele, and G. G. Karady, "Advancement in the application of neural networks for short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 7, pp. 250-257, 1992.

[14] L. M. Saini, and M. K. Soni, "Artificial neural network based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods," *IEE Gen. Trans. Dist.*, vol. 149, no. 5, pp. 578-584, 2002.

[15] MATLAB version 2014b Natick, Massachusetts: The MathWorks Inc., 2014.

[16] Y.V. Makarov et al. Incorporating wind generation and load forecast uncertainties into power grid operations. Report PNNL-19189, PNNL, Jan 2010.

[17] L. Xie, Y. Gu, X. Zhu, and M. G. Genton, "Power system economic dispatch with spatio-temporal wind forecasts," Presented at the IEEE Energytech, 2011.

[18] N. Lu, R. Diao, R. P. Hafen, N. Samaan, and Y. V. Makarov, "A comparison of forecast error generators for modeling wind and load uncertainty," Presented at the IEEE Power and Energy Society General Meeting, 2013.

[19] R. Billinton, and R. Allan, Reliability Evaluation of Power Systems, 2nd ed. New York, NY, USA: Plenum, 1996

[20] K. P. Murphy, Machine learning: a probabilistic perspective. MIT press, 2012.