NFC-IEFR Journal of Engineering & Scientific Research

Journal URL: www.nijesr.iefr.edu.pk/

A New Forward-Backward Linear Algorithm for Electrical Load Demand Prediction

Badar ul Islam[1]*, Abdul Sattar Saand[2], Q. D. Memon[1], Tahir Raza[1], Amir Haider[1]

[1]*Department of Electrical Engineering, NFC Institute of Engineering and Fertilizer Research, Faisalabad*
[2]*Quaid-e-Awam University of Engineering, Science & Technology Nawabshah, Sindh, Pakistan*

ARTICLE INFO

***Corresponding Author:**
badar.utp@gmail.com
DOI:
10.24081/nijesr.2017.1.0008

Keywords:
Load Forecasting, Computation time, Mean Absolute percentage error, Hybrid intelligent system, Performance analysis.

ABSTRACT

One hour-ahead load forecasts are critically important for the reliable and cost-effective operation of modern power systems, especially, in the deregulated economy, where effective on-spot price fixing is always a major concern. In this paper, a time series modeling for these very short term forecasts is proposed by using a new auto-regressive algorithm, specifically designed for appropriate handling of large data records. The model divides the bulky data record into short segments and searches for the AR coefficients that simultaneously model the data with the least mean squared error. This approach can accurately forecast the one hour-ahead and one day-ahead loads of the weekdays. The proposed method can provide more accurate results than the conventional techniques, such as, standard AR-based algorithms, Burg and the seasonal Box-Jenkins ARIMA (SARIMA) and many of the intelligent hybrid approaches, where artificial neural networks are combined with evolutionary search methods. Obtained results from extensive testing of the proposed model confirm the validity of the developed approach. Three years load demand data of New South Wales (NSW) Australian power grid from 2011 to 2013 are used in the experimentation.

I. Introduction

Electrical load demand prediction is important to modern power system planning, operation, and control. These forecasts can be used for several purposes in accordance with the difference in their lead times. In particular, very short-term load forecasting with the lead time from one to several hours is crucial to economical and reliable operation, security analysis and unit commitment for both power generation and distribution facilities. Hence, the improvements in the accuracy of these forecasts will not only increase the suitability of scheduling and planning but also assures the reduced operational cost and reliable operation of the power systems.

Many techniques have been developed for these forecasts, which can be broadly classified into three categories; conventional approaches, artificial intelligence (AI)-based techniques and hybrid methods. The conventional approaches are linear mathematical/statistical models including, linear regression methods [1], exponential smoothing [2], Box–Jenkins approaches [3] and Kalman filters [4]. Artificial intelligence based techniques, including neural network models [5], expert system models [6], fuzzy inference [7] and support vector machines [8] have achieved good results for short-term load forecasting. All these techniques have their individual advantages and disadvantages; however the hybrid methods based on ANN and other computational intelligent method have produced high prediction accuracy in the recent past. Most of these techniques are mainly based on ANN combined with genetic algorithm (GA), particle swarm optimization (PSO) or simulated annealing (SA). Moreover, combinations between ANN, GA, Fuzzy logic and ARMA are also used to solve the STLF problem [9, 10].

The choice of the STLF model selection is also related with the availability and choice of input variables. In general, majority of the single variable based techniques rely on autoregressive moving average (ARMA) method and other parametric models by using historical record of the power load demand as a single variable. On the other hand, a combination of the weather variable with the power load demand is referred as multivariate approach. Intelligent hybrid forecast methods also referred as multivariate STLF techniques mostly considered as a good choice in case of multiple input variables.

Owing to the importance of ARMA for the STLF, a large number of estimation methods for ARMA model parameters have been proposed over the last 40 years. ARMA model has more degrees of freedom than the autoregressive, so greater latitude in its ability to generate diverse time-series shapes is therefore expected of its estimators. Unfortunately, this is not always the case, because of the nonlinear nature requirement of the algorithms that must simultaneously estimate the moving average and autoregressive parameters of the ARMA model. Indeed, all existing solutions to this problem appear to suffer from one or more drawbacks, as explained briefly in the followings:

- Several methods may end up in a hard failure mode. This means that the identification algorithm may return an invalid model or the algorithm cannot be carried out to completion because as a result of a step during its execution, a parameter set outside the class of permissible ones arises for which no provisions have been adopted. This may happen for moment fitting procedures, as well as to methods that first estimate the AR parameters and then the MA.

- Maximum-likelihood methods that depend on search over the parameter space involve significant computations and are not guaranteed to converge, or they may converge to the wrong solution.
- Finally, some methods may be inaccurate (e. g., significantly biased) in finite samples. This is the case with Durbin's two stage least-squares method [11-13], and with the approximate subspace methods that enforce positivity of the estimated MA spectrum based on [14].

In this paper, an AR algorithm designed for long data records is proposed. The algorithm divides the data record into segments and searches for AR coefficients that simultaneously model all of them with least means squared errors. In order to verify the proposed algorithm as a solution to STLF problem, its performance is compared with other AR-based algorithms, like Burg [15, 16] and the Modified Covariance (MCOV) [13], as well as with Durbin's as ARMA-based algorithm [13, 14]. The results of the proposed algorithm are found better in terms of forecast accuracy and computational time.

II. Modified Forward-Backward Linear Prediction (MF-BLP) Algorithm

In this section the proposed AR algorithm designed for long data records is described. The algorithm divides the data record into segments and searches for AR coefficients that simultaneously model all of them with least means squared errors. Assume the m-points data sequences x(1), x(2), …, x(m) are to be used to estimate the *p*-th AR parameters. Since, with AR algorithms the order of the model is proportional to the length of data record [13, 17, 18]. In order to avoid using large orders with long data records (three years half hourly data), it is considered that the segmentation of the m-points data sequence into Q segments of N samples each. Assume one segment of data out of the available Q segments. Because forward and backward linear predictions have similar statistical information [13], it seems reasonable to combine the linear prediction error statistics of both directions in order to generate more error points. The net result should be an improved estimate of the auto-regressive parameters.

To estimate the *p-th* AR parameters the m-points data sequences *x(1), x(2), …, x(m)* are assumed. Since, with AR algorithms the order of the model is proportional to the length of data record [13, 17, 18]. In order to avoid using large orders with long data records (three years load demand data), it is considered that the segmentation of the m-points data sequence into Q segments of N samples each.

Therefore, the matrix form of (N-p) forward and the (N-p) backward linear prediction samples representing the Q segments and N samples are formulated by the following steps of equations, and it is given by;

$$\mathbf{D}_p^f(q) = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(m-2) & x(m-3) & \cdots & x(m-p-1) \end{bmatrix} \quad (1)$$

$$\mathbf{D}_p^f(q) = \left[x_0^f(q)\ x_1^f(q) \cdots x_{m-p-1}^f(q)\right]^T \quad (2)$$

where;

$$\mathbf{z}_k^f(q) = \left[x(p+k-1)\ \ x(p+k-2) \cdots x(k)\right]^T \quad (3)$$

where; $k = 0, 1, 2, \ldots, m\text{-}p\text{-}1$.

The linear predicted array outputs corresponding to the forward data matrix $\mathbf{D}_p^f(q)$, can be described as;

$$w_p^f(q) = \left[\hat{x}(p)\ \hat{x}(p+1) \cdots \hat{x}(m-1)\right]^T \quad (4)$$

The similar approach from the above forward linear prediction, the (*m-p*) sub vectors in the backward direction is drawn as follows;

$$\mathbf{D}_p^b(q) = \begin{bmatrix} x^*(p-1) & x^*(p-2) & \cdots & x^*(0) \\ x^*(p) & x^*(p-1) & \cdots & x^*(1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-2) & x^*(m-3) & \cdots & x^*(m-p-1) \end{bmatrix} \quad (5)$$

$$\mathbf{D}_p^b(q) = \left[x_0^b(q)\ x_1^b(q) \cdots x_{m-p-1}^b(q)\right]^T \quad (6)$$

where;

$$x_k^b(t) = \left[x^*(k+1)\ x^*(k+2) \cdots x^*(k+L)\right]^T \quad (7)$$

$k = 0, 1, 2, \ldots, m\text{-}p\text{-}1$.

The predicted array outputs corresponding to the backward data matrix $\mathbf{D}_p^b(t)$, are given as;

$$w_p^b(t) = \left[\hat{x}^*(0)\ \hat{x}^*(1) \cdots \hat{x}^*(m-p-1)\right]^T \quad (8)$$

Therefore, with the combination of forward and backward linear predicted algorithm the MF-BLP data matrix can be described as;

$$\mathbf{D}^{fb} = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(p-2) & x(p-3) & \cdots & x(m-p-1) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-p) & x^*(m-p+1) & \cdots & x^*(m-1) \end{bmatrix} \quad (9)$$

Assume that w^{fb} is the desired response at the predictor output of the D^{fb}, which can be defined as;

$$\mathbf{w}^{fb} = \left[x(p)\ x(p+1) \cdots x(m-1)\ \hat{x}^*(1)\ \hat{x}^*(2) \cdots \hat{x}^*(m-p-1)\right]^T \quad (10)$$

With the **f**, matrix for the MF-BLP prediction coefficients, where **f**, is given by;

$$\mathbf{f} = \left[a_p^{fb}(1)\ a_p^{fb}(2)\cdots\ a_p^{fb}(p)\right]^T \tag{11}$$

Hence the MF-BLP model can be written in matrix form as;

$$\mathbf{D}^{fb}\mathbf{f}_p^{fb} = \mathbf{w}^{fb} \tag{12}$$

For the simplicity, the Eq. (12) can be reduced to;

$$\mathbf{D}\ \mathbf{f} = \mathbf{w} \tag{13}$$

where **D** is the data series, **f** is the predicted coefficients and **w** is the predicted response (signal).

The (N-p) forward and the (N-p) backward linear prediction formulation in Eq. (13) associated with the Q-segments of data series can be redefined as;

$$\mathbf{D}_q\mathbf{f} = \mathbf{w}_q \tag{14}$$

where; the $2\times(N-p)$ forward-backward linear prediction data matrix is given by;

$$\mathbf{D}_q = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(p-2) & x(p-3) & \cdots & x(m-p-1) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-p) & x^*(m-p+1) & \cdots & x^*(m-1) \end{bmatrix} \tag{15}$$

Let **w** denotes the desired response at the predictor output, given by;

$$\mathbf{w}_q = \left[x(p)\ x(p+1)\cdots\ x(m-1)\ \hat{x}^*(1)\ \hat{x}^*(2)\cdots\ \hat{x}^*(m-p-1)\right]^T \tag{16}$$

Therefore, the coefficients of MF-BLP algorithm is described as;

$$\mathbf{f} = \left[a_p(1)\ a_p(2)\cdots\ a_p(p)\right]^T \tag{17}$$

By forming the data matrix $\mathbf{D}_q$ in corresponds to each data segment, Q, and arranging the resultant matrices in the following form;

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \vdots \\ \mathbf{D}_Q \end{bmatrix} \tag{18}$$

The corresponding predicted vector to matrix **D** is defined as;

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_Q \end{bmatrix} \tag{19}$$

Hence, the Eq. (17), (18) and (19) can be rewritten and summarized as;

$$\mathbf{D}\ \mathbf{f} = \mathbf{w} \tag{20}$$

where; $\mathbf{f} = [\mathrm{f}_L(1)\ \mathrm{f}_L(2)\ldots\ \mathrm{f}_L(\mathrm{L})]^T$.

A well known criterion called the Least-Squares, will be used to obtain a solution to Eq. (20) for the predictor coefficient vector **f**. This solution will guarantee the minimum sum of squared values of the predicted errors [19]. Since the number of the predicted values has been significantly increased by segmentation of the data, then it is expected that the least-squares solution for the predictor coefficients will provide effective solution to power load forecast.

The MF-BLP algorithm applied is applied to power load demand data in the following manner:

1. The power load demand data $x(1), x(2), \ldots, x(n)$ is formed.
2. The data series is segmented into Q segments and *N* number of samples for each segments as indicated in Eq. (18).
3. From the data matrix, the optimum number for Q and *N* are obtained.
4. The estimation data matrix (in-sample data) is then estimated.
5. Finally, the value of f_L are calculated using Eq. (21).

III. Results and Discussion

In this section we will find the best values for Q, N, and L. Three years data load demand record collected by NEMMCO in NSW, Australia, between the beginning of 2011 and the end of 2013, is used for this study [20-23]. The first two years hourly data (17520 samples) are used for model extraction and the remaining year of data are used for model validation. The mean absolute percentage error (MAPE) is used as a metric indicate the accuracy of the model in predicting data.

In the first experiment, the two years 17520 data samples are segmented into Q segments of different lengths N. The different arrangements of Q and N are shown in Table 1.

Table 1: Q and N for the NSW two-year hourly load demand data.

Segment No.	Q	N
Segmentation 1	2	8760
Segmentation 2	5	3504
Segmentation 3	10	1752
Segmentation 4	24	730
Segmentation 5	48	365

Next, the predictor order (L) is varied from 0.1N to 0.25N and the best order of the MF-BLP predictor is defined. As indicator of performance, the MAPE value is calculated for each L over the validation year of data. Since the MAPE values are function of the leading forecast time, they are calculated with one hour-ahead leading times over the validation year of data:

The MATLAB results are included in Tables 2, which shows the MAPE values of MF-BLP algorithm as a function of L in case of 1 hour-ahead forecast. The MAPE values are calculated from 350 × 24 data samples over the validation year.

Table 2: The MAPE values of the MFBLP as a function of L in 1-hour ahead forecast

Segment No.	0.1N	0.15N	0.2N	0.25N
Segmentation 1	1.21	1.16	1.05	0.84
Segmentation 2	1.22	1.19	1.07	0.88
Segmentation 3	1.21	1.18	1.07	0.89
Segmentation 4	1.20	1.18	1.05	0.85
Segmentation 5	1.20	1.19	1.05	0.83

The MATLAB results in terms of MAPE as shown in Table 2, clearly indicate that the MF-BLP is reaching its best performance when L = 0.25N. Now, if we look into the 0.25N columns in Tables 2, we can easily find that the MF-BLP is achieving its best performance with the 5th segmentation, namely when Q = 48. Having obtained the best values for Q and N at L = 0.2N, would the MF-BLP algorithm show any further improvement with increased L? In order to answer this question, L is varied over the range from 0.1N to 0.65N and the MAPE values are calculated and depicted in Figure 1.

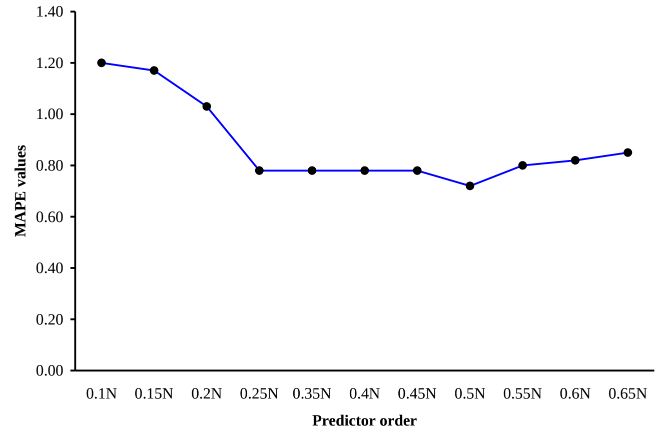

Figure 1: MAPE values of the MF-BLP algorithm as a function of predictor order.

Figure 1 shows a MATLAB plot of the predictor order verses MAPE values. It can be clearly seen that the MF-BLP is achieving its best performance for L = 0.25N. For higher values than L = 0.25N, the algorithm is showing a relatively constant behavior till L = 0.55N, where it starts to show some signs of deterioration with increased L. Thus, we chose L = 0.25N as the best value for the predictor order.

In the second experiment, the computational time of the MF-BLP is found for the different segmentation schemes. The algorithm is written in MATLAB Software and run on Intel Core i5 machine. L is considered to be 0.25N and N = 17520 samples (two years). The results of computational time in seconds are shown in Table 3.

Table 3: The computational time of the MFBLP algorithm with L = 0.25N and N = 17520 samples as a function of the different segmentations.

Segment No.	Computational time (seconds)
Segmentation 1	21
Segmentation 2	13
Segmentation 3	10
Segmentation 4	5
Segmentation 5	3

The MATLAB results in Table 3 show clearly that Segmentation-5 is giving the least computational time for the MFBLP algorithm. This will add another important factor to the smallest MAPE value which obtained in the first experiment, in order to consider Segmentation-5 as the best one for the MF-BLP algorithm.

Based on the obtained results in the two above experiments it is found that the MF-BLP is achieving its optimum performance in terms of MAPE and computational time with Q = 48 (in case of 17520 data samples) and L = 0.25N.

To verify that the proposed model has the generalization capacity and is not on over fitting, it has been tested for two more cases. In the first case, four years historical electrical load demand data (from 2009 to 2012) of Faisalabad Electricity Supply Corporation (FESCO) are employed. The similar data segmentation was adopted with the same number of samples as adopted in the first case. The proposed model produced the reasonable forecast accuracy in this case with a minor average MAPE rise of 0.1315% for all the five data segments. The peak demand of FESCO has been observed as 2500 MW, which is 5 times lesser as compared to the NSW Australian grids [24]. This minor rise in the forecast error is due to lesser load demand of FESCO grids as compared to NSW Australian grids. Additionally, the proposed model has been tested for its generalization capacity by using the ISO UK grid data [25]. In this case study, the error from segment 1 to segment 5 were recorded as; 0.85%, 0.91%, 0.83%, 0.89% and 0.92%, respectively. The least computational time for the segmentation 5 was computed as 3.5 seconds. These results show that the proposed model is reasonably adaptable in European and Pakistani electrical markets, provided that the load demands are in the range of few thousands of megawatts.

IV. Conclusion

The proposed MF-BLP algorithm is thoroughly described as a solution to power load forecast problem, and its optimum parameters are experimentally obtained. The algorithm is based on the segmentation of the power load demand data samples into Q segments and finding the forward back linear prediction data matrix for each segment. From the Q data matrices an overall matrix is obtained. Three years data load demand record collected by NEMMCO in NSW, Australia, between the beginning of 2011 and the end of 2013, are used to find the best values for the number of segments (Q) and predictor order (L) as a function of segment length (N). Two metrics are used as performance indicators of the MF-BLP, which are the MAPE and the computational time. The results shows that the best value of Q with two years data samples (17520 samples) is 48 and the best value for the predictor order as a function of segment length is 0.25N.

V. Acknowledgment

The authors wish to thank NFC-IEFR administration for providing the opportunity to conduct this research.

References

[1] Fumo, N., & Biswas, M. R. (2015). Regression analysis for prediction of residential energy consumption. Renewable and Sustainable Energy Reviews, 47, 332-343.

[2] Taylor, James W. "Short-term electricity demand forecasting using double seasonal exponential smoothing." Journal of the Operational Research Society 54.8 (2003): 799-805.

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

[4] Al-Hamadi, H. M., and S. A. Soliman. "Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model." Electric power systems research 68.1 (2004): 47-59.

[5] Hippert, H. S., Pedreira, C. E., & Castro, R. (2001). Neural networks for shortterm load forecasting: A review and evaluation. IEEE Transactions on Power System, 16(1), 44–55.

[6] Kim, Kwang-Ho, Hyoung-Sun Youn, and Yong-Cheol Kang. "Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method." IEEE Transactions on Power Systems 15.2 (2000): 559-565.

[7] Mori, H., & Kobayashi, H. (1996). Optimal fuzzy inference for short-term load forecasting. IEEE Transactions on Power System, 11, 390–396.

[8] Chen, B., Chang, M., & Lin, C. (2004). Load forecasting using support vector machines: A study on EUNITE competition 2001. IEEE Transactions on Power System, 19(4), 1821–1830.

[9] Zhang, G. P. (2003). Times series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175

[10] Kim, K. H., Park, J. K., Hwang, K. J., & Kim, S. H. (1995). Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems. IEEE Transactions on Power System, 10(3), 1534– 1539.

[11] Chen, J. F., Wang, W. M., & Huang, C. M. (1995). Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting. Electric Power Systems Research, 34(3), 187–196.

[12] J. Durbin, "The fitting of time series models," Rev. Internet. Statist., Vol. 23, Pp. 233-244, 1960, 1960.

[13] P. M. T. Broersen, "Autoregressive model orders for Durbin's MA and ARMA estimators," Signal Processing, IEEE Transactions on, vol. 48, pp. 2454-2457, 2000.

[14] N. S. Kamel, "High resolution sonar signal processing for detection of arrivals estimation," June 1993.

[15] M. H. Hayes, Statistical Digital Signal Processing and Modeling Wiley, 1996.

[16] N. Kamel and Z. Baharudin, "Short term load forecast using Burg autoregressive technique," Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on, pp. 912-916, 2007.

[17] J. Burg, "Maximum entropy spectral analysis," Proc. 37th Meeting of Soc. of Exploration Geophysicists, Oklahoma City, October 1967.

[18] S. de Waele and P. M. T. Broersen, "Order selection for vector autoregressive models," Signal Processing, IEEE Transactions on, vol. 51, pp. 427-433, 2003.

[19] P. M. T. Broersen, "Let the Data Speak for Themselves," Instrumentation and Measurement, IEEE Transactions on, vol. 56, pp. 814-823, 2007.

[20] Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L. and Yang, S., 2008. RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. IEEE Transactions on power systems, 23(3), pp.853-858.
[21] Fan, S. and Hyndman, R.J., 2012. Short-term load forecasting based on a semi-parametric additive model. IEEE Transactions on Power Systems, 27(1), pp.134-141.
[22] Xiao, L., Wang, J., Hou, R. and Wu, J., 2015. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy, 82, pp.524-549.
[23] <Last seen on 02/05/2017> http://www.nemmco.com.au/data/market_data
[24] <Last seen on 02/05/2017> www.ntdc.com.pk/Files/fesco.pdf
[25] <Last seen on 02/05/2017> https://www.iso-ne.com/isoexpress/web/reports/load-and-demand

1.