MFBLP Method Forecast for Regional Load Demand System

Zuhairi Baharudin zuhairb@petronas.com.my
Electrical and Electronics Department
University Technology PETRONAS
31750 Tronoh, Perak, Malaysia

Nidal Kamel nidalkamel@petronas.com.my
Electrical and Electronics Department
University Technology PETRONAS
31750 Tronoh, Perak, Malaysia

Abstract

Load forecast plays an important role in planning and operation of a power system. The accuracy of this forecasted value is necessary for economically efficient operation and also for effective control. This paper describes a method of modified forward backward linear predictor (MFBLP) for solving the regional load demand of New South Wales (NSW), Australia. The method is designed and simulated based on the actual load data of New South Wales, Australia. The accuracy of discussed method is obtained and comparison with previous methods is also reported.

Keywords: Short term load forecasting (STLF), autoregressive (AR), Modified, Linear Predictor, Autoregressive moving average (ARMA), Burg.

1. INTRODUCTION

Short term prediction of future load demand is important for the economic and secure operation of power systems. Fundamental operation functions such as unit commitment, hydro-thermal coordination, interchange evaluation, scheduled maintenance and security assessment require a reliable short term load forecast (STLF).

Throughout the paper the term “short” is used to imply prediction times of the order of hours. The time boundaries are from the next hour, or possibly a half-hour, up to 168 h. The basic quantity of interest in STLF is the hourly integrated total system load. Owing to the importance of the STLF, research in this area in the past two decades has resulted in the development of numerous forecasting methods [1, 2, 2-10]. One of the STLF methods that received significant attention in literature for more than 20 years and a large number of estimation methods is the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA). Both models are also known as Box-Jenkins have 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 required of algorithms that must simultaneously estimate the moving average and autoregressive parameters of the models. This phenomenon finally produces low accuracy in forecast of the model algorithm based from the assumption below:

Though the all pole models have less degree of freedom than ARMA and ARIMA, they exhibit major advantages. First of all, all-pole models have been found to provide a sufficiently accurate representation for many different types of signals in many different applications. Another reason for the popularity is the special structure which leads to fast and efficient algorithms for finding the all-pole parameters. In this paper, the modified forward backward linear predictor (MFBLP) is introduced as autoregressive (AR) based solution to STLF problem. The method is discussed and the performance is tested and compared with Box-Jenkins ARIMA [5] (assuming that the performance of ARIMA is better than ARMA), artificial neural network (ANN) [11] and Burg's algorithm [12]. A load demand data record from New South Wales (NSW), Australia is used for the model designed purpose and validation process [13].

2. METHODOLOGY

The methods highlight in the paper is the non-weather type model. This means that, it takes only the historical load data as an input to the model. In general, the time dependent non-weather model uses past and latest load behavior to extrapolate the sample data prior to the forecast stage. The hourly load data are observed to identify the data behavior. Then, the data need to examine whether the data consist of seasonality or trend. The non stationary behavior should be remedied before the estimating process takes place.

In the previous literature [5, 14] had suggested that the method of difference the data can be one of the steps to achieve stationary white-noise in the data series. The next step is to examine the residuals of the estimation process. The residuals should be white-noise (stationary) [15] then, it possible to run the key steps in the methodology, which is the forecast. The summary of the steps taken in the methodology is depicted in the Figure 1.

2.1. Data observation

The historical load data used in this study has been gathered from the National Electricity Market Management Company Limited (NEMMCO) [13]. The data series is the actual hourly data of the system load demand for the region of New South Wales (NSW), Australia. The regional data series for NSW system demand covering the period from 1st January 2005 to 31st December 2007. Figure 2 and 3 shows the time series plot (sequence plot) for the estimation sample and forecast validation data respectively.

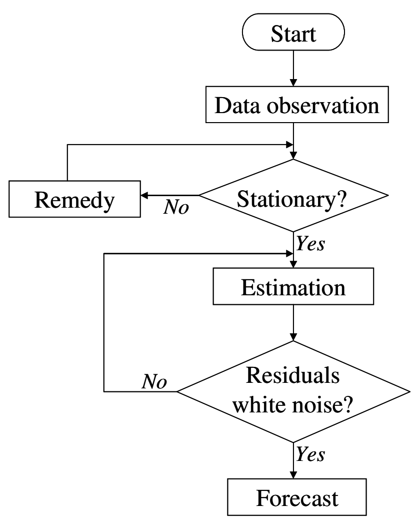

FIGURE 1: The general procedure of the STLF.

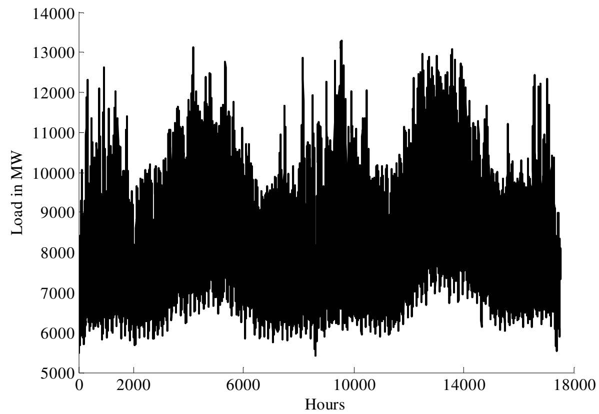

FIGURE 2: NSW hourly data from 1.1.2005 to 31.12.2006.

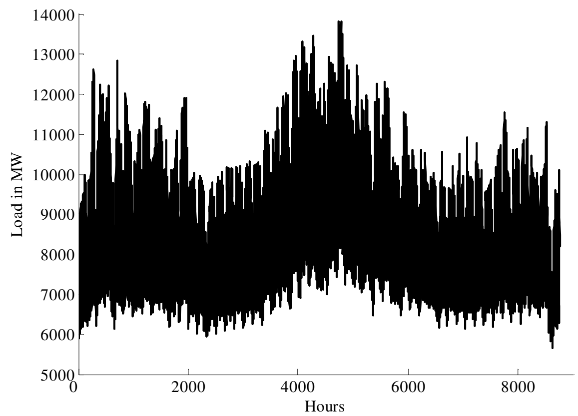

FIGURE 3: NSW hourly data from 1.1.2007 to 31.12.2007

Figure 2 and 3 clearly depict that the data series contains seasonality. Moreover, in Figure 2, the data show slightly positive trends. The seasonality is formed by the cycle of the load demand behavior, which is contributed from daily, weekly and yearly load demand patterns. It is clearly seen that, the data series is having non-zero mean and non-constant variance [12]. This suggests that a further remedy needs to apply to the data in order to achieve the zero mean and constant variance.

2.2. Data pre-processing

The important issue in time series forecast is to clarify the behavior of zero mean and constant variance of the given data series, which is the main assumption in this paper. According to Figure 1, the estimation sample data in Figure 2 need the necessary remedy to remove the seasonality and trend in the data. One of the effective solutions for this is differentiating the data as suggested in [5, 14]. Consequently, suggest the triple difference as a method to remove the non-stationary behavior in the data.

By assuming the actual load demand data series as L_t , hence, the considered difference is as follows:

$$x_t = L_t - L_{t-1} \qquad (1)$$

$$y_t = x_t - x_{t-24} \qquad (2)$$

$$z_t = y_t - y_{t-168} \qquad (3)$$

The first difference when $k = 1$ eliminate the small trend that develops in Figure 2. The difference of $k = 24$ and later by $k = 168$ eliminate the daily and weekly seasonality respectively. Tentatively the model can be described as follows:

$$\nabla_1 \nabla_{24} \nabla_{168} L_t = e_t \qquad (4)$$
$$(1-B)(1-B^{24})(1-B^{168})L_t = e_t \qquad (5)$$

Where B denotes the normal back-shift operator, i.e. $BL_t = L_{t-1}$ or $B^k L_t = L_{t-k}$ and the term e_t is the white noise process of the residuals. For simplicity the expression in (5) can be reduced as:

$$z_t = e_t \qquad (6)$$

The sequence plot and the autocorrelation plot (ACF) of z_t are depicted in Figure 4 and 5 respectively. It is clearly seen from the Figure 4, the new data series show stationary and most importantly the data are now zero mean and constant variance. The data autocorrelation is considered negligible, and it is well in the bound of 95% confidence level. The confidence level is given by [15],

$$Confidence\ level_{95\%} = \frac{\pm 1.96}{\sqrt{n}} \qquad (7)$$

Where n is the number of samples.

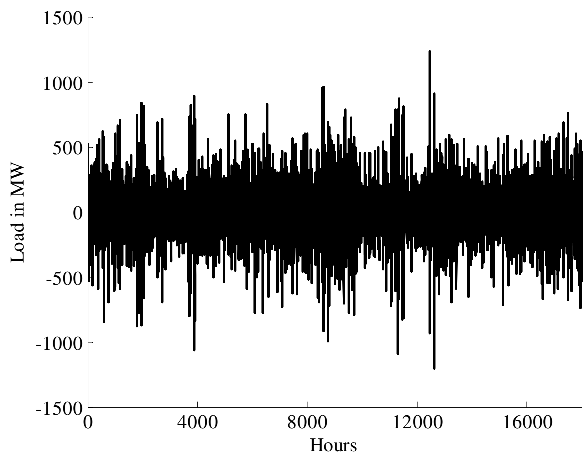

FIGURE 4: Sequence plot of z_t.

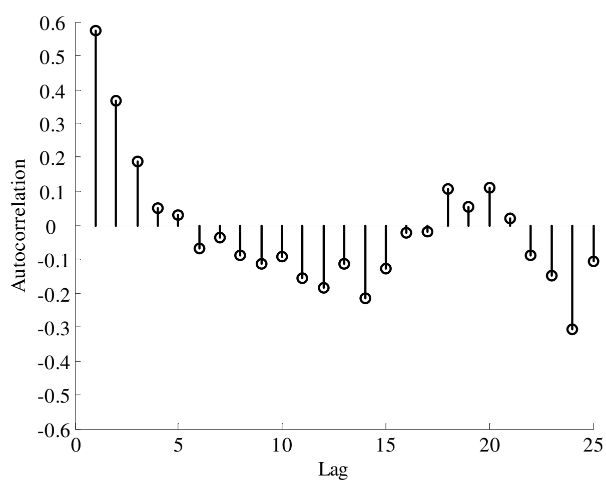

FIGURE 5: ACF of z_t for the first 25 lags.

2.3. The MFBLP algorithm

The *M*-point data sequence of *z*(1), *z*(2), … *z*(*M*) is to be used to estimate the *p*th AR filter coefficients. Since, with AR algorithms the order of the model is proportional to the length of data record and in order to avoid using large orders with long data records, let's consider 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, it seems reasonable to combine the linear prediction error statistics of both directions in order to generate more error points [16]. The net result should be an improved estimate of the autoregressive parameters. The (N-p) forward and the (N-p) backward linear prediction samples of the non-windowed case may be written in matrix form as follows:

$$\mathbf{D_q f} = \mathbf{w} \qquad (8)$$

Where the 2×(N-p) forward-backward linear prediction data matrix is defined as,

$$\mathbf{D_q} = \begin{bmatrix} x(N-1) & x(N-2) & \cdots & x(0) \\ x(N) & x(N-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(N-2) & x(N-3) & \cdots & x(N-p) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(N-p) & x^*(N-p+1) & \cdots & x^*(N-1) \end{bmatrix} \qquad (9)$$

Let **w** denotes the desired response at the predictor output, defined as,

$$\mathbf{w_q} = \left[x(p) \quad x(p+1) \quad \cdots \quad x(N-1) \quad x^*(1) \quad x^*(2) \quad \cdots \quad x^*(N-p-1) \right]^T \; (10)$$

Meanwhile the following vector notation for the forward linear prediction coefficients is given as,

$$\mathbf{f} = \left[a_p^f(1) \;\; a_p^f(2) \;\; \ldots \;\; a_p^f(p) \right]^T \qquad (11)$$

Since the forward and backward prediction coefficients for a stationary random process are simply complex conjugate of one another, the vector of the desired response at the predictor output, once again can be rewritten in matrix form as in Eq. (8). By forming the data matrix $\mathbf{D_q}$ in corresponds to each data segment, and arranging the resultant matrices in the following form,

$$\mathbf{D} = \begin{bmatrix} \mathbf{D_1} \\ \mathbf{D_2} \\ \vdots \\ \mathbf{D_Q} \end{bmatrix} \qquad (12)$$

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

$$\mathbf{w} = \begin{bmatrix} \mathbf{w_1} \\ \mathbf{w_2} \\ \vdots \\ \mathbf{w_Q} \end{bmatrix} \qquad (13)$$

From Eq.(12) and Eq.(13), the modified equations of the predictors, can be simplify as,

$$\mathbf{Df} = \mathbf{w} \qquad (14)$$

Where **D** is the data series, **f** is the predicted coefficients and **w** is the predicted response (signal). A well established method of least squares [17, 18] is used to obtain a solution to Eq. 14 for the predictor coefficient vector **f.** The solution will guarantee the minimum sum [16] of the squared values of the predicted errors (residuals).

2.4. Estimation

By referring to the Eq. (14), the linear predicted errors are given by,

$$\mathbf{e} = \mathbf{w} - \mathbf{Df} \qquad (15)$$

and the sum of squared errors is given by,

$$\zeta = \mathbf{e}^{\mathbf{H}}\mathbf{e} \qquad (16)$$

By substituting Eq. (14) into Eq. (16), the equation may explicitly express the dependence of the sum of squared errors on the predictor coefficients, as follows

$$\zeta = \mathbf{w}^{\mathbf{H}}\mathbf{w} - \mathbf{w}^{\mathbf{H}}\mathbf{Df} - \mathbf{f}^{\mathbf{H}}\mathbf{D}^{\mathbf{H}}\mathbf{w} + \mathbf{f}^{\mathbf{H}}\mathbf{D}^{\mathbf{H}}\mathbf{Df} \qquad (17)$$

Now, by differentiating Eq. (17) with respect to the **f**, hence the following gradient vector is determined,

$$\begin{aligned}\frac{\partial\zeta}{\partial\mathbf{f}} &= -2\,\mathbf{D}^{\mathbf{H}}\mathbf{w} + 2\,\mathbf{D}^{\mathbf{H}}\mathbf{Df} \\ &= \mathbf{-2}\,\mathbf{D}^{\mathbf{H}}\mathbf{e}\end{aligned} \qquad (18)$$

As it is clear the sum of squared errors reaches its minimum value when the gradient vector is zero. Then from Eq. (18), the equation immediately deduces to,

$$\mathbf{D}^{\mathbf{H}}\mathbf{Df} = \mathbf{D}^{\mathbf{H}}\mathbf{w} \qquad (19)$$

Thus, the predictor coefficients that give the least squared errors are obtained as a solution to Eq. (19). However, this solution is unique only when the matrix **D** is full rank. When this condition is satisfied the matrix $\mathbf{D}^{\mathrm{H}}\mathbf{D}$ is nonsingular and the solution is unique, given as,

$$\mathbf{f} = \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathbf{H}}\mathbf{w} \qquad (20)$$

Thus the predictor coefficients that give the least squared errors are obtained as a solution to Eq. (20). However, this solution is unique only when the nullity of the matrix **D** is zero [19]. The nullity of a matrix denoted as null (.) is defined as the dimension of the matrix null space. In other words the least-squares solution is unique when the matrix **D** is of full rank. When this condition is satisfied, the *p*-by-*p* matrix $\mathbf{D}^{\mathrm{H}}\mathbf{D}$ is nonsingular and the solution is unique, given as,

$$\begin{aligned}\mathbf{f} &= \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathbf{H}}\mathbf{w} \\ &= \mathbf{D}^{\#}\mathbf{w}\end{aligned} \qquad (21)$$

Where **D** is called the pseudo-inverse of the matrix **D**, given as,

$$\mathbf{D}^{\#} = \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathrm{H}} \qquad (22)$$

2.5. Forecast

Once the process estimates the suggested model, the predicted coefficients, **f**, are determined. With the coefficients, the forecast load for the region of NSW will be determined. One year actual data from 1.1.2007 to 31.12.2007 are used as a validation data or model forecast performance indicator. This validation data series is the actual hourly load data and can be described as,

$$v(t) = \left[v(1)\, v(2) \cdots v(N)\right] \qquad (23)$$

The forecast horizons are set as follows: 1-hour, 24-hour, 48-hour and 168-hour step ahead. This means that the coefficients, **f**, are applied in determining the *t*-th step ahead by using the autoregressive (AR) method. The AR a model for *p*-th order of coefficients, **f**, for the 1-hour ahead is given by,

$$f(1) = \left[a(1)z(t) + a(2)z(t-1) \cdots a(p)z(t-p+1)\right] \qquad (24)$$

The process is reversed to determine the actual forecast value for, $F(1)$, and is given by,

$$F(1) = (1-B)(1-B^{24})(1-B^{168})\,f(1) \qquad (25)$$

Thus the forecast error is stated as,

$$e_t = v(1) - F(1) \qquad (26)$$

Therefore, for the *t*-th step ahead forecast, the AR (*p*) model can be described as,

$$F(t) = (1-B)(1-B^{24})(1-B^{168})\,f(t) \qquad (27)$$

Where

$$f(t) = a_p(B)z_t \qquad (28)$$

The term *B* is the backshift operator, defined by $Bz_t = z_{t-1}$, and hence $B^k z_t = z_{t-k}$.

It is noticed that the Eq. 25 are assumed as a 'starter' forecast value, where this formulation is used to determine the next value of forecast whenever $t > 1$. The accuracy of the total forecast is depending on this starter value and how accurate the process estimates in producing the AR coefficients, *a* (*p*). As mentioned earlier, by using the MFBLP algorithm and estimation by the least squares methods, the errors of the forecast are considered minimum.

3. THE EXPERIMENT

The simulations considered in this paper consist of computing the forecast of 1, 24, 36 and 168 hours ahead. The system load demand data that are applied in this work are gathered from NEMMCO. For the estimation process, the data are range from 1.1.2005 to 31.12.2006 and for the forecast validation the data are taken from 1.1.2007 to 31.12.2007. Both mentioned data are hourly sequence load demand data. This data series is considered as a raw data because no attempts are done to smooth the data significant from the abrupt change in weather, extreme weather condition or even the effect of public holiday which are very significant for the particular affected week of the load demand. The main objective for this is to test and confirm that the non-weather AR-model of Eq. 14 and Eq. 28 could deliver an acceptable forecast result. For a comparison purpose, several previous methods are in comparison to benchmark the proposed method described in the paper. All models in this work are simulated in a computer with a MATLAB version 7.2.0.232 and Pentium 4 CPU 3.20GHz with 0.5GB of RAM.

3.1. Errors indicator

The error analysis in estimation and forecast processes are represented by the mean absolute percentage errors (MAPE), and it is given by,

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|x_i - \hat{x}_i\right|}{x_i}\times 100\% \qquad (29)$$

3.2. The segmentation parameter

The main significant finding for the MFBLP is the formulation of data segmentation. The data are segmented into *Q* segments of *N* samples each. As described previously, the total point's data after it has been remedied to remove the seasonality and trends consist of 17520 points of data. Consequently, provide several combinations of segmentation parameters (SP) of *Q* and *N*. Four SP combinations are investigated and the results for the average MAPE during an estimation process for each SP combination are depicted in the Table I.

Table I: Average MAPE for four combinations of SP.

Segmentation parameters		Estimation
Q	N	Average MAPE, %
2	8760	0.97
12	1460	0.82
24	730	0.75
48	365	0.91

From Table 1 it is clearly found that the best SP combinations are Q = 24 and N = 730. The assumption for the SP combination finding is that Q^*N must equal to total number of data points, which in this case 17520.

3.3 Estimation the number of order coefficients

The MFBLP in Eq. 9 and the simplified Eq. 14 are used to determine the estimation of the data. By implementing the least squares, it can provide the minimum sum squares of the errors. As mentioned before, the works in this paper highlight the non-weather AR (AR) methods in defining the model order number and its coefficients. In [20] suggested that for the optimum number of model order, it should not exceed 10% of the total estimation sample data. This would result from the initial draw of the parameters for the model should be less than 1752. However, in the actual experiment, the model order number is much less than the suggested one [20]. Figure 6 depicts the relation between the estimation errors and the model order number.

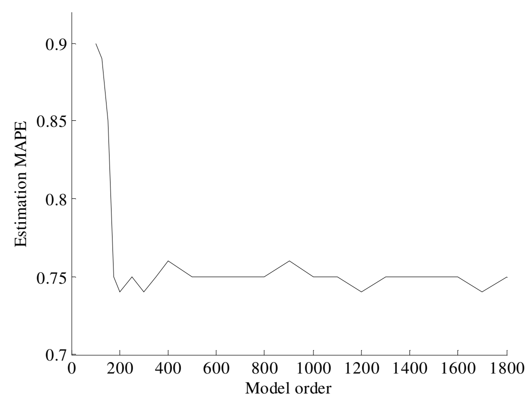

FIGURE 6: The relation of estimation MAPE and the model order number.

It is clearly seen that, the MAPE converges after a certain number of order. The lowest MAPE is 0.74% at order 200. There is a slight ripple of MAPE averaging at 0.75%. By using the AICC [21] the order with the minimum value is selected, and it is found that the best number of the order model is 175 as indicated in the Figure 7. The residuals of the estimated model are calculated and examined with the selected model order. Figure 8 depicts the residuals for the estimated samples and Table II describes the model performance with the specified parameters. The white noise of the residuals is supported by checking the ACF value which is well inside the 95% confidence interval bound. This is a clear indication that the residuals are stable, controllable and most importantly stationary with zero mean and constant variances.

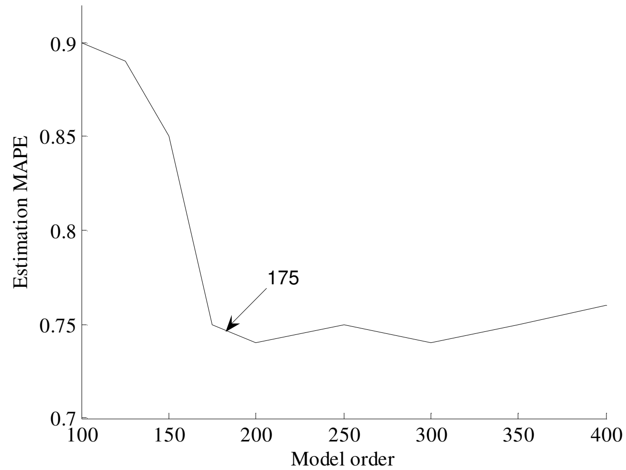

FIGURE 7: The estimation MAPE and the model order number up to 400.

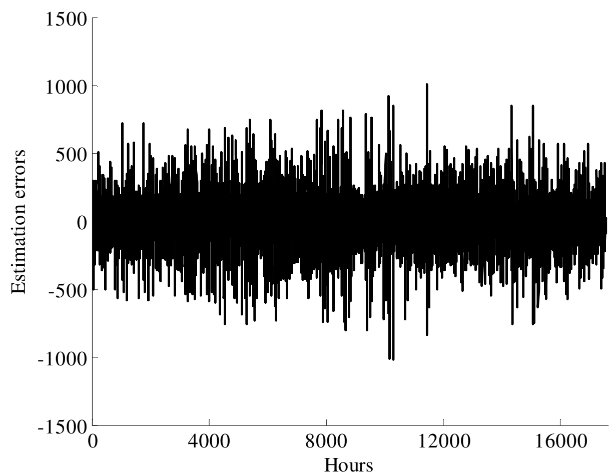

FIGURE 8: The residuals of the estimated models.

TABLE II: The number of order parameter selected and the performance of estimated models.

No. of samples	No. of order parameter	Residuals
17520	175	0.75%

3.4. Model forecast

The forecast model is described in previous section and can be represented by the Eq. 14, Eq. 16 and Eq. 19. The models emphasize the methods of non-weather AR with the application of high number of order coefficients. The model order coefficients are defined as,

$$a(p) = \{a(1), a(2), ..., a(p-1), a(p)\} \qquad (30)$$

From Table II it is found that the model order coefficients p , is equal to 175. Therefore, the forecast values are calculated by using the back shift operator as described in Eq. 27.

The forecast horizons simulated in this experiment are 1, 24, 36 and 168-hours ahead. For every forecast horizon, the value of forecast is determined and being validated with the actual (out of sample) data covering the period from 1.1.2007 to 31.12.2007. For the comparison purpose, the proposed method is compared with the multiplicative ARIMA [5], artificial neural network (ANN) [11], Burg's algorithm [12] and also with the naïve-random walk ARIMA (0,1,0) as a basic benchmarking.

4. APPLICATION RESULTS

Figure 9 depicts the model performance for the forecast duration of the year 2007. Meanwhile in Figure 10 describes the seasonal MAPE of 168-hours (one week) ahead forecast.

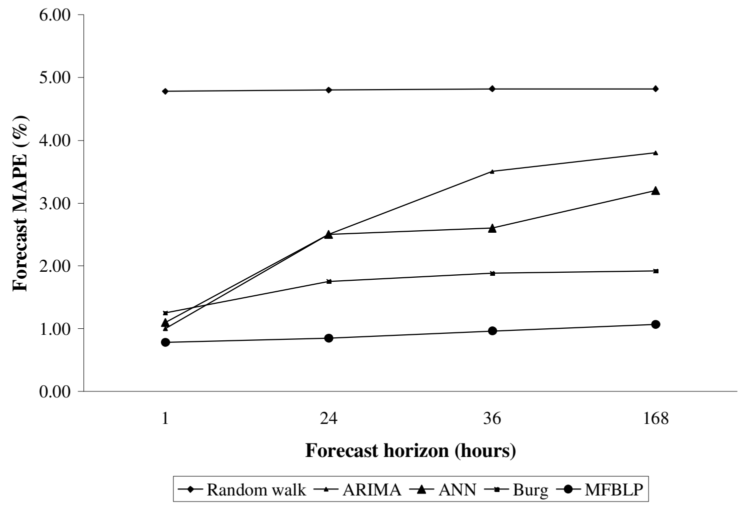

Figure 9: The MAPE plotted against forecast horizons for the NSW region of one year forecast period.

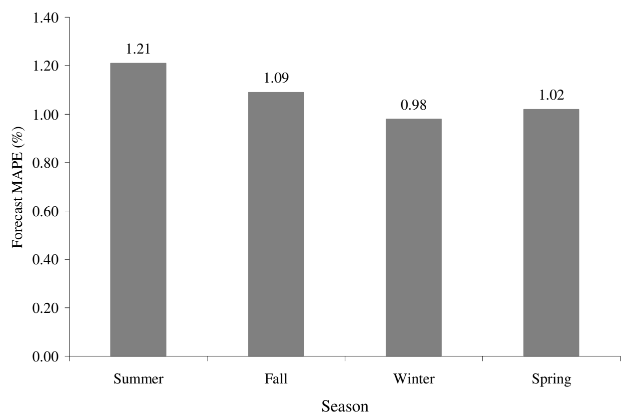

Figure 10: The MAPE plotted against seasons for the NSW one year forecast period.

It is clearly shown that from Figure 9 and 10, the methods of non-weather AR capable to describe its performance outstandingly. It is worth to investigate the model performance by weekly results. In order to obtain the results, only one selected season is chosen. The summer season is selected because, during this duration the load demand variations are high. Therefore, the forecast process is somehow difficult. It is assumed that other seasons show slightly simpler in forecasting and their MAPE results are considered lower than the summer season. Figure 11 describes the model accuracy for the random selected one week forecast during the summer season. The forecast horizons apply for 1, 24, 36 and 168-hours ahead. Meanwhile, the MAPE of the daily accuracy for the 168-hours ahead forecast horizon is depicted in Table III. The minimum and maximum for the 24-hours forecast MAPE during the week is highlighted in Table IV and Table V respectively.

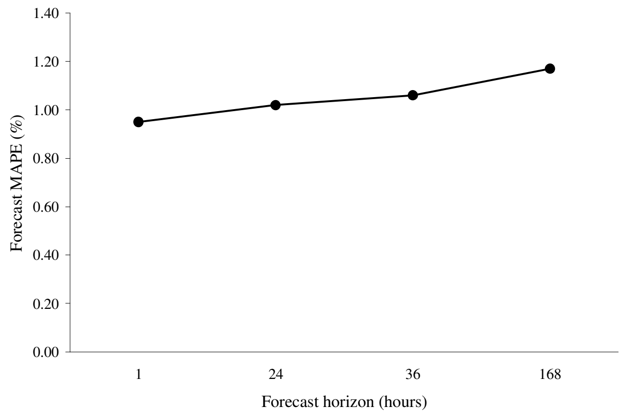

Figure 11: The forecast MAPE plotted against forecast horizons for the one week forecast duration in summer.

TABLE III: Daily MAPE for one week in the summer.

Forecast day	**MAPE (%)**
Day 1	1.25
Day 2	0.58
Day 3	0.78
Day 4	1.33
Day 5	1.38
Day 6	1.44
Day 7	1.34

TABLE IV: The hourly MAPE minimum of the week.

Forecast hour	**MAPE (%)**	**Forecast hour**	**MAPE (%)**
1	0.30	13	0.47
2	0.68	14	0.81
3	0.70	15	0.17
4	0.73	16	0.39
5	1.11	17	0.58
6	0.86	18	0.76
7	0.02	19	0.04
8	0.28	20	0.23
9	0.66	21	0.05
10	0.85	22	0.59
11	0.80	23	1.33
12	0.68	24	0.75

TABLE V: The hourly MAPE maximum of the week.

Forecast hour	MAPE (%)	Forecast hour	MAPE (%)
1	0.09	13	0.36
2	0.86	14	0.53
3	0.46	15	1.26
4	0.47	16	0.10
5	0.23	17	0.85
6	0.09	18	0.74
7	1.87	19	1.85
8	3.60	20	0.49
9	4.16	21	2.49
10	2.49	22	1.70
11	4.30	23	1.60
12	1.79	24	2.07

5. CONCLUSIONS

In this paper the non-weather autoregressive (AR) methods are proposed and discussed. The methods highlight the application of MFBLP and the estimation of model order coefficients by using the least squares approach. The experiment uses the hourly electricity load demand data gathered form NEMMCO. The estimation process uses two years hourly load data and one year hourly data for the forecast validation purpose. The forecast horizon for 1, 24, 36 and 168-hours ahead are examined to explain the workable of model forecast. It is worth mentioning that, there is no attempt to smooth the data prior to the estimation which is significant from the weather change and public holiday. This proves that the methods manage to find the accurate model coefficients to estimate and forecast the given data series.

The MAPE versus the forecast horizons for previous suggested approaches are also determined. The results show that the proposed MFBLP method is superior in others. For all forecast horizons MFBLP method performs remarkable results than others. Perhaps it could be a strong indicative of represent the performance of all suggested method. MFBLP shows a very promising result and perhaps, it can be used to solve any of the time series forecasting problems. In future work, it would be interesting to consider the public holidays and the obvious outliers prior to the estimation-forecast processes. This probably would enhance the forecast MAPE. Consequently, it would benefit the utility in reducing the operational cost and most importantly could avoid the energy being waste. Hence the daily usage power could be optimized significantly from less fuel being burnt and helping to reduce the pollution in our environment.

6. References

[1] Qia Ding, "Long-Term Load Forecast Using Decision Tree Method," *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES,* pp. 1541-1543, 2006.

[2] D. Asber, S. Lefebvre, J. Asber, M. Saad and C. Desbiens, "Non-parametric short-term load forecasting," *International Journal of Electrical Power & Energy Systems,* vol. 29, pp. 630-635, 10. 2007.

[3] D. W. Bunn and E. D. Farmer, *Comparative Models for Electrical Load Forecasting.* United States: John Wiley and Sons Inc.,New York, NY, 1985 Jan 01, pp. Pages: 186.

[4] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proceedings of the IEEE,* vol. 88, pp. 163-169, 2000.

[5] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time Series Analysis: Forecasting & Control (3rd Edition).* New York: Prentice Hall; 3rd edition (February 28, 1994), 1994, pp. 592 pages.

[6] Z. Baharudin and N. S. Kamel, "AR modified covariance in one week ahead short term load forecast," *IASTED Proceeding Power and Energy Systems,* vol. 606-070, 2008.

[7] M. El-Telbany and F. El-Karmi, "Short-term forecasting of Jordanian electricity demand using particle swarm optimization," *Electric Power Systems Research,* vol. 78, pp. 425-433, 3. 2008.

[8] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas and P. Liatsis, "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models," *Energy,* vol. 33, pp. 1353-1360, 9. 2008.

[9] N. Amjady, "Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method," *Power Systems, IEEE Transactions on,* vol. 22, pp. 333-341, 2007.

[10] D. Asber, S. Lefebvre, M. Saad and C. Desbiens, "Modeling of Distribution Loads for Short and Medium-Term Load Forecasting," *Power Engineering Society General Meeting, 2007. IEEE,* pp. 1-5, 2007.

[11] G. A. Darbellay and M. Slama, "Forecasting the short-term demand for electricity: Do neural networks stand a better chance?" *International Journal of Forecasting,* vol. 16, pp. 71-83, 0. 2000.

[12] 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.

[13] NEMMCO. January 5, 2008). National electricity market management company; australia wholesale market and power system operator. *2008(January),*

[14] G. E. P. Box and G. C. Reinsel, *Time Series Analysis Forecasting and Control.* San Francisco: Holden Day, 1976,

[15] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting.* ,2nd ed.Springer-Verlag New York, Inc., 2002,

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

[17] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Review of linear least squares theory," in *Time Series Analysis Forecasting and Control* ,Third Edition ed.Anonymous Prentice Hall, 1994, pp. 286-304.

[18] M. H. Hayes, "The least squares method," in *Statistical Digital Signal Processing and Modeling* ,First ed.Anonymous USA: John Wiley & Sons, Inc., 1996, pp. 131-133.

[19] S. Haykin, *Adaptive Filter Theory.* New Jersey: Prentice Hall Inc, Englewood Cliff, 1986,

[20] 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.

[21] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on,* vol. 19, pp. 716-723, 1974.