[1]Z. Baharudin,
[2]N. S. Kamel, [3]I. Musirin

ARMA versus AR in Short Term Load Forecast

IJPEAI
International Journal of Power, Energy and Artificial Intelligence

Abstract Short-term load forecasting plays an important role in planning and operation of power system. The accuracy of this forecasted value is necessary for economically efficient operation and also for effective control. This paper describes a comparison of autoregressive moving average (ARMA) and autoregressive (AR) Burg's and modified covariance (MCOV) methods in solving short term load forecast. The methods are tested based on historical load data of New South Wales, Australia. The accuracy of discussed methods are obtained and reported.

Keywords - Short term load forecasting (STLF), autoregressive (AR), Autoregressive moving average (ARMA), Burg, MCOV, Durbin, MAPE

I. INTRODUCTION

Prediction of future events and conditions is called forecasts, and the act of making such predictions is called forecasting. It is essential to have accurate models to forecast the future load demand. Load demand forecasting is significant for efficient network operation, network planning, economic scheduling of generation units and also maintenance activities. Load demand forecasting is typically categorized into long, medium and short term prediction. Long term and medium term forecast usually cover for longer duration of time (monthly and yearly values), which is used for applications in capacity expansion of generation and transmission [1-4]. While short term load forecast (STLF) is normally carried out for an interval ranging from possibly half an hour or one hour to one week ahead. To supply the load demand over this particular duration of time involves the start up and shutdown of entire generating units, which will be determined by a number of generation control functions such as hydro scheduling, hydro thermal coordination, unit commitment and interchange evaluation [5]. These load information is obtained from the STLF system and it is vital to the operational of dispatch centre in order to dispatch load economically. It is a main goal for any utilities company to operate at cost as low as possible. One way to achieve this is to minimize the forecast error. It was estimated that an increase of operating cost associated with a 1% increase of forecast error was 10 million pounds per year [6].

Previously the time series methods [7-9] have been applied to STLF. The statistical approach in [10] had drawn that, the all-pole models algorithm exhibit less degree of freedom, thus proven to be an alternative solution to STLF [11,12].

[1, 2] Electrical & Electronics Department, Universiti Teknologi PETRONAS, 31750 Tronoh, Perak , Malaysia. Email: zuhairb@petronas.com.my.

[3] Universiti Teknologi MARA, 40000 Shah Alam, Malaysia. Email: ismailbm@salam.uitm.edu.my

With all-pole models algorithm, the STLF problem can be solved with only a single input i.e. historical load data. Recently, the hybrid method and artificial intelligence (AI) based techniques have been applied to STLF. The hybrid load forecast [13-16] had shown an improvisation of load forecast accuracy. AI based techniques such as artificial neural network (ANN) [17-19] uses weather ensemble data as one or more of the input to the developed models. The ANN technique is proven reliable in prediction error, however very large historical data is needed. This is contrast with the fuzzy logic. With such generic conditioning rules and expert knowledge [20,21] properly designed fuzzy logic systems can be very robust when used for forecasting. A simple prediction of fuzzy model in [22,23] proves that it could provide a very satisfying prediction error. Other fuzzy logic techniques and combined approach are discussed in [24-27]. The method of particle swarm [28], support vector machines [29] and neuro-fuzzy [30,31] had also shown relatively good forecast accuracy. However, these techniques are applicable only when a precise future weather data and expert knowledge are available.

In this paper, the methods of autoregressive moving average (ARMA) and autoregressive (AR) solving for STLF are presented. The methods which incorporate the statistical time series modeling [10] are drawn to predict future load demand. The methods are simulated and tested based from the historical load data of New South Wales, Australia. The models result of daily and weekly accuracy in STLF is obtained and compared.

II. AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODEL

A time-series model that approximate many discrete-time stochastic processes encountered in practice is presented by the filter linear difference equation of complex coefficients

$$x(n) = -\sum_{k=1}^{p} a_p(k)x(n-k) + \sum_{k=0}^{q} b_q(k)u(n-k) = \sum_{k=0}^{\infty} h(k)u(n-k) \quad (1)$$

in which $x(n)$ is the output sequence of a causal filter ($h(k)$ = 0 for k < 0) that models the observed data and $u(n)$ is an input driving white noise sequence. Eq. (1) determines the *autoregressive-moving average* (ARMA) model for the time series $x(n)$.

The $a_p(k)$ parameters form the autoregressive portion of the ARMA model. The $b_q(k)$ parameters form the moving average portion of the ARMA model. Thus, a wide sense stationary ARMA(p,q) process may be generated by filtering unit variance white noise $u(n)$ with a causal linear shift-invariant filter having p poles and q zeros.

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q(k) z^{-k}}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \qquad (2)$$

Therefore, a random process $x(n)$ may be modeled as an ARMA(p,q) process using the model shown in Fig. (1), where $u(n)$ is unit variance white noise.

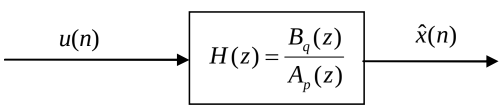


Fig. 1 Modeling a random process x(n) as the response of a linear shift-invariant filter to unit variance white noise.

If (1) is multiplied by $x^*(n-m)$ and the expectation taken, the result is known as Yule-Walker equation, given as [11].

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = c_q(k) \qquad (3)$$

where $r_x(k)$ is the autocorrelation sequence (ACS) of $x(n)$ and the sequence $c_q(k)$ is the convolution of $b_q(k)$ and $h^*(-k)$

$$c_q(k) = b_q(k) * h^*(-k) = \sum_{l=0}^{q-k} b_q(l+k) h^*(l) \qquad (4)$$

Since $h(n)$ is assumed to be causal, then $c_q(k) = 0$ for $k > q$ and the Yule-Walker equations for $k > q$ are a function only of the coefficients $a_p(k)$,

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = 0 \quad ; \; k > q \qquad (5)$$

The autoregressive parameters of an ARMA model are related by a set of linear equations to the autocorrelation sequence. Expressing (5) in matrix form for $k = q+1, q+2, \ldots, q+p$ we have

$$\begin{bmatrix} r_x(q) & r_x(q+1) & \cdots & r_x(q-p+1) \\ r_x(q+1) & r_x(q) & \cdots & r_x(q-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(q+p-1) & r_x(q+p-2) & \cdots & r_x(q) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} r_x(q+1) \\ r_x(q+2) \\ \vdots \\ r_x(q+p) \end{bmatrix} \qquad (6)$$

Which is a set of p linear equations in the p unknowns, $a_p(k)$.These equations, referred to as the Modified Yule-Walker equations.

Once the coefficients $a_p(k)$ have been determined, the next step is to find the moving average (MA) coefficients, $b_q(k)$. There are several approaches that may be used to accomplish this. Two of these approaches are describes briefly in the sequel of this section.

An MA(q) process may be generated by filtering unit variance white noise $u(n)$ with an FIR filter of order q as follows:

$$x(n) = \sum_{k=0}^{q} b_q(k) u(n-k) \qquad (7)$$

The Yule-Walker equations relating the ACS to the filter coefficients $b_q(k)$ are

$$r_x(k) = b_q(k) * b_q^*(-k) = \sum_{l=0}^{q-|k|} b_q(l+|k|\, b_q^*(l)) \qquad (8)$$

Note that, unlike the case for an autoregressive process, these equations are nonlinear in model coefficients $b_q(k)$. Therefore, even if the ACS were known exactly, finding the coefficients $b_q(k)$ may be difficult. Instead of attempting to solve the Yule-Walker equations directly, another approach is to perform a spectral factorization of the power spectrum $P_x(z)$. Specifically, since the autocorrelation of an MA(q) process is equal to zero for $|k|>q$, the power spectrum is a polynomial of the form

$$P_x(z) = \sum_{k=-q}^{q} r_x(k) z^{-k} = \sigma_0^2 Q(z) Q^*(1/z^*) \qquad (9)$$

where Q(z) is a minimum phase monic polynomial of degree q. $\sigma_0 = b_q(0)$ and $Q(z)$ is the minimum phase version of $B_q(z)$ that is formed by replacing each zero of $b_q(z)$ that lies outside the unit circle with one that lies inside the unit circle at a the conjugate reciprocal location [11]. Thus, given the autocorrelation sequence of an MA(q) process, we may find a model for $x(n)$ as follows. From the autocorrelation sequence $r_x(k)$ we for the polynomial $P_x(z)$ and factor it into a product of a minimum phase polynomial, $Q(z)$, and a maximum phase polynomial $Q^*(1/z^*)$. The process $x(n)$ may then be modeled as the input of the minimum phase FIR filter

$$H(z) = \sigma_0 Q(z) = \sigma_0 \sum_{k=0}^{q} q(k) z^{-k} \qquad (10)$$

driven by unit variance white noise.

As an alternative to spectral factorization, a moving average model for a process $x(n)$ may also be developed using Durbin's method [11,33]. This approach begins by finding a high-order all pole model $A_p(z)$ for the moving average process. Then by considering the coefficients of the all-pole model $a_p(k)$ to be a new "data set", the coefficients of a qth-order moving average model are determined by finding a qth-order all-pole model for the sequence $a_p(k)$. Once the high-order all-pole model for $x(n)$ has been found, it is then necessary to estimate the MA coefficients $b_q(k)$ from all-pole coefficients $a_p(k)$. We see that since

$$A_p(z) \approx \frac{1}{B_q(z)} = \frac{1}{b_q(0) + \sum_{k=1}^{q} b_q(k) z^{-k}} \qquad (11)$$

then $1/B_q(z)$ represents a qth-order all-pole model for the "data" $a(k)$. The coefficients of the all-pole model for $a(k)$ are taken as the coefficients of the moving average model. Typically, the model order p is chosen so that it is at least four times the order q of the moving average process [34].

III. THE AUTOREGRESSIVE (AR) MODELS

A wide-sense stationary autoregressive process of order p is a special case of an ARMA(p,q) process in which q = 0. An AR(p) process may be generated by filtering unit variance white noise, $u(n)$ with an all-pole filter of the form

$$H(z) = \frac{b_q(0)}{1 + \sum_{k=1}^{p} a_p(k) z^{-k}} \qquad (12)$$

Just as with ARMA process, the autocorrelation sequence of an AR process satisfies the Yule-Walker equations

$$r_x(k) + \sum_{l=1}^{p} a_p(l) r_x(k-l) = \left|b_q(0)\right|^2 \delta(k) \quad ; \ k \geq 0 \qquad (13)$$

Writing these equations in matrix form for k =1,2, …, p, using the conjugate symmetry of $r_{xx}(k)$, we have

$$\begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) & \cdots & r_x^*(p-1) \\ r_x(1) & r_x(0) & r_x^*(1) & \cdots & r_x^*(p-2) \\ r_x(2) & r_x(1) & r_x(0) & \cdots & r_x^*(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & r_x(p-3) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ a_p(3) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \vdots \\ r_x(p) \end{bmatrix} \qquad (14)$$

Therefore, given the autocorrelation $r_x(k)$ for k =0,1, .., p we may solve (14) for the AR coefficients. These equations may be solved recursively using Levinson-Durbin Recursion [33] which led to a number of important discoveries including the lattice filter structure.

A close relationship exists between a linear prediction filter and an AR process. If the random process $x(n)$ is generated as an AR(p) process and the order of the linear predictor $m = p$, then the predictor coefficients will be identical to the AR parameters. This relationship is exploited by several algorithms in finding the AR coefficients through linear prediction [33].

Consider the forward linear prediction estimate

$$\hat{x}^f(n) = -\sum_{k=1}^{p} a_p^f(k) x(n-k) \qquad (15)$$

of the sample $x(n)$, where $a_p^f(k)$ is the forward linear prediction coefficients at time index k. The hat ^ is used to denote an estimate and the superscript f is used to denote that this is a forward estimate. The prediction is forward in the sense that the estimate at time index n is based on p samples indexed earlier in time. The complex forward linear prediction error is

$$e_p^f(n) = x(n) - \hat{x}_p^f(n) \qquad (16)$$

has a real variance

$$\rho^f = \mathrm{E}\left\{\left|e_p^f(n)\right|^2\right\} \qquad (17)$$

where **E**{.} denotes the expected value.

In similar way to forward prediction a backward linear prediction error estimate

$$\hat{x}_p^b(n) = -\sum_{k=1}^{p} a_p^b(k) x(n+k) \qquad (18)$$

may also be formed, in which $a_p^b(k)$ is the backward linear prediction coefficient at time index k. A superscript b is used to tag elements associated with the backward linear prediction estimate. The prediction is backward in the sense that the estimate at time index n is based on m samples indexed later in time. The backward linear prediction error is

$$e_p^b(n) = x(n-m) - \hat{x}_p^b(n-m) \qquad (19)$$

and has real variance

$$\rho^b = \mathbf{E}\left\{\left|e_p^b(n)\right|^2\right\} \qquad (20)$$

If the Levinson-Durbin recursion is substituted for $a_p^f(k)$ or $a_p^b(k)=a_p^{f*}(k)$ in definitions (15) and (18) for the forward and backward linear prediction errors, then it is simple to see that

$$\begin{aligned} e_{j+1}^f(n) &= e_j^f(n) + \Gamma_{j+1} e_j^b(n-1) \\ e_{j+1}^b(n) &= e_j^b(n-1) + \Gamma_{j+1}^* e_j^f(n) \end{aligned} \qquad (21)$$

Since the lattice filter provides an alternative parameterization of the all-pole filter, i.e., in terms of its reflection coefficients, we may also consider formulating the all-pole signal modeling problem as one of finding the reflection coefficients that minimize some error. In the following section we look at two such lattice methods for signal modeling including Burg's method, and the modified covariance method.

A. Burg's Method

The method determine the reflection coefficients and can be computed sequentially by minimizing the mean-square of the forward and backward prediction error [11].

$$\varepsilon_j^{fb} = \varepsilon_j^f + \varepsilon_j^b = \mathbf{E}\left\{\sum_{n=j}^{N}\left|e_j^f(n)\right|^2 + \sum_{n=j}^{N}\left|e_j^b(n)\right|^2\right\} \quad (22)$$

Now, we may find the value of the reflection coefficients Γ_j^{fb}, that minimizes ε_j^{fb} by setting the derivates of ε_j^{fb} with respect to $(\Gamma_j^{fb})^*$ equal to zero as follows.

$$\frac{\partial}{\partial\left(\Gamma_j^{fb}\right)^*}\varepsilon_j^{fb} = \frac{\partial}{\partial\left(\Gamma_j^{fb}\right)^*}\mathbf{E}\left[\sum_{n=j}^{N}\left\{\left|e_j^f(n)\right|^2 + \left|e_j^b(n)\right|^2\right\}\right] \quad (23)$$

$$= \mathbf{E}\left[\sum_{n=j}^{N}\left\{e_j^f(n)\left[e_{j-1}^b(n-1)\right]^* + \left[e_j^b(n)\right]^* e_{j-1}^f(n)\right\}\right] = 0$$

Substituting the error update equations for $e_j^f(n)$ and $\left[e_j^b(n)\right]^*$, which are similar to those given for $e_{j+1}^f(n)$ and $\left[e_{j+1}^b(n)\right]^*$ in (21), and solving for Γ_j^{fb} we find that the value of Γ_j^{fb} that minimizes ε_j^{fb} is

$$\Gamma_j^{fb} = -\frac{2\sum_{n=j}^{N} e_{j-1}^f(n)\left[e_{j-1}^b(n-1)\right]^*}{\sum_{n=j}^{N}\left\{\left|e_{j-1}^f(n)\right|^2 + \left|e_{j-1}^b(n-1)\right|^2\right\}} \quad (24)$$

B. The Modified Covariance Method

In the previous section we described Burg recursion, which finds the reflection coefficients for an AR model by sequentially minimizing the mean of the squared forward and backward prediction errors. In this section we look at the modified covariance method or forward-backward algorithm for AR signal modeling. As with Burg algorithm, the modified covariance method minimizes the mean of the squares of the forward and backward prediction errors,

$$\varepsilon_p^{fb} = \varepsilon_p^f + \varepsilon_p^b \quad (25)$$

The difference, however, between the two approaches is that, in the modified covariance method, the minimization is not performed sequentially. In other words, for a *p*th-order model, the modified covariance method finds the set of reflection coefficients or, equivalently, the set of transversal filter coefficients $a_p(k)$, that minimize ε_p^{fb}.

To find the filter coefficients that minimizes ε_p^{fb} we set the derivates of ε_p^{fb} with respect to $a_p^*(l)$ equal to zero for l = 1, 2, …, p. Since

$$e_p^f(n) = x(n) + \sum_{k=1}^{p} a_p^f(k)x(n-k) \quad (26)$$

and

$$e_p^b(n) = x(n-p) + \sum_{k=1}^{p} a_p^{f*}(k)x(n-p+k) \quad (27)$$

then

$$\frac{\partial\varepsilon_p^{fb}}{\partial a_p^*(l)} = \mathbf{E}\left[e_p^f(n)\frac{\partial\left[e_p^f(n)\right]^*}{\partial a_p^*(l)} + \left[e_p^b(n)\right]^*\frac{\partial e_p^b(n)}{\partial a_p^*(l)}\right] \quad (28)$$

$$= \mathbf{E}\left[e_p^f(n)x^*(n-l) + \left[e_p^b(n)\right]^* x(n-p+l)\right] = 0$$

Substituting Eqs. (26) and (27) into (28) and simplifying we find that the normal equations for the modified covariance method are given by

$$\sum_{k=1}^{p}\left[r_x(l,k) + r_x(p-k,p-l)\right]a_p^f(k) = -\left[r_x(l,0) + r_x(p,p-l)\right];\ l = 1,...,p \quad (29)$$

where

$$r_x(l,k) = \sum_{n=p}^{N} x(n-k)x^*(n-l) \quad (30)$$

For the modified covariance error we may use the orthogonality condition in (28) to express ε_p^{fb} as follows

$$\varepsilon_p^{fb} = E\left[e_p^f(n)x^*(n) + \left[e_p^b(n)\right]^* x(n-p)\right] \quad (31)$$

Substituting the expression given in Eqs. (26) and (27) for $e_p^f(n)$ and $e_p^b(n)$ and simplifying, we have

$$\varepsilon_p^{fb} = r_x(0,0) + r_x(p,p) + \sum_{k=1}^{p} a(k)\left[r_x(0,k) + r_x(p,p-k)\right] \quad (32)$$

IV. APPLICATIONS AND RESULTS

The performance of the discussed ARMA and AR methods for short term load forecast is compared with the load demand data from the New South Wales (NSW), Australia [35]. Fig. 2 shows the load variation in New South Wales demand over one year of data record.

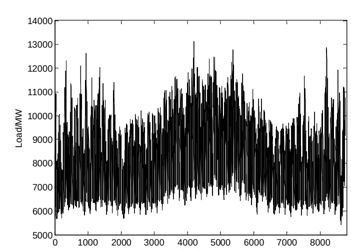


Fig. 2: Demand data for the NSW grid between 01 January 2005 and 31 December 2005.

Fig. 3 depicts the load demand pattern for two weeks. The figure represent daily load pattern, starting from Monday to

Sunday. It can be seen here that the weekday's pattern are almost similar. Mean while the weekends pattern are slightly lower than weekdays load. The significant difference of this load pattern is because during weekends, the load usage from the government sector, business, and factory are lower compared to weekdays. Hence, result in the complexity of the load demand data and making the prediction is tough to achieve high accuracy.

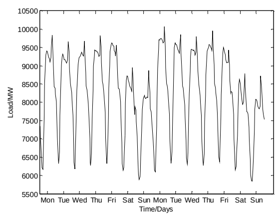


Fig. 3: Hourly load curve for NSW grid for two weeks.

In this simulation, we investigate the performance of ARMA and AR models [10] in several case studies. All case studies are using 2352 of observation data and 168 for forecast data (validation data). The observation data is used as an input to the model. After that, the model determines the reflection coefficients of the data. The multiplication [11, 12] is done by using the reflection coefficients to determine the next hour of load forecast. After forecasting this first hour load, its value is used to forecast the second hour load. This process is repeated iteratively until the 24 hours load forecast is achieved. After 24 hours forecast is completed, the observation data will be updated. This is mean that, the first 24 hours observation data is deleted and the actual data of previous 24 hours data is added. Hence, with this approach will keep the total of observation data to be consistent at 2352 data points. The mentioned case studies; recorded and forecast period are depicted in TABLE I.

TABLE I
RECORDED AND FORECAST PERIOD

Case	Recorded period	Forecast period
1	1 Jan 05-3 April 05	4 April 05-10 April 05
2	7 Feb 05 – 8 May 05	9 May 05 – 15 May 05
3	7 Mar 05 – 5 Jun 05	6 Jun 05 – 12 Jun 05
4	14 Mar 05 – 12 Jun 05	13 Jun 05 – 19 Jun 05
5	4 Apr 05 – 3 Jul 05	4 Jul 05 – 10 Jul 05

The Mean Absolute Percentage Errors (MAPE), calculated from 24 forecast value is used as a performance indicator. The MAPE is given by

$$MAPE = \frac{1}{24}\sum_{i=1}^{24}\frac{\left|\hat{x}_i - x_i\right|}{x_i} \qquad (33)$$

where x_i is the actual data and $\hat{x}_i$ is the forecast value.

For the case 1; Fig. 4 and TABLE II, show the accuracy of the studied ARMA and AR algorithms. In Fig. 4, it is clearly shown that, AR give better accuracy than ARMA during peak load time. Mean while, TABLE II shows the performances of the ARMA and AR through their MAPE values over the week of forecast.

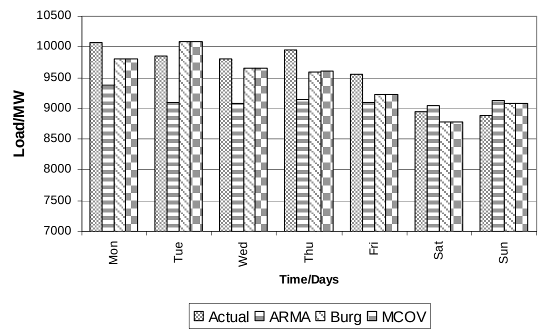


Fig. 4: Load forecast at 7 pm over the period 04 -10 April 2005 for NSW load demand.

TABLE II
MAPE VALUES OF ONE-WEEK AHEAD FORECASTS OF NSW LOAD DEMAND.

	MAPE (%)		
	ARMA	AR	
Day	Durbin	Burg	MCOV
Mon	4.61	2.75	2.77
Tue	4.93	1.83	1.81
Wed	4.50	1.35	1.34
Thu	3.05	1.65	2.05
Fri	5.12	2.23	2.24
Sat	5.95	1.63	1.62
Sun	5.40	1.77	1.76
Average	**4.80**	**1.89**	**1.94**

As the results in TABLE II indicates AR's models are showing a relatively consistent performance independent to some extent from the level of fluctuation in the observation data, with averaged MAPE value of less than 2%. On the other hand, the results in TABLE II show poor performance by Durbin's' ARMA with MAPE values varying between 3.05% on Thursday and 5.95% on Saturday, producing an averaged MAPE value of nearly 5%. Comparing the two performances, we may say that AR's models are outperforming Durbin's ARMA in NSW grid by roughly two times.

TABLE III
MAPE RESULTS FOR CASE 2

	MAPE (%)		
	ARMA	AR	
Day	Durbin	Burg	MCOV
1 (Mon)	6.23	2.58	2.60
2 (Tue)	3.56	1.92	1.98
3 (Wed)	2.23	1.35	1.40
4 (Thu)	2.51	1.54	1.55
5 (Fri)	4.89	2.12	2.32
6 (Sat)	5.68	1.87	1.93
7 (Sun)	4.79	1.65	1.66
Average	**4.27**	**1.86**	**1.92**

The simulation is continued for the case 2 with the observation data is kept consistent at 2352 data points. Hence, the results for this simulation are depicted in TABLE III.

TABLE IV
MAPE RESULTS FOR CASE 3, 4 AND 5.

	MAPE (%)		
	ARMA	AR	
Case	Durbin	Burg	MCOV
Case 3	4.98	1.79	1.87
Case 4	5.12	1.85	1.91
Case 5	4.75	1.72	1.79
Case 4	**4.96**		**1.82**

Finally, TABLE IV depicts the MAPE for the one week forecast value for the case 3, 4 and 5. It is clear from TABLE III and IV, AR outperform ARMA by some margin. For daily and weekly comparison, AR definitely shows the better accuracy.

V. CONCLUSION

In this paper, the performance of ARMA is compared with AR in short term load forecast. Two AR's models, Burg, and the Modified covariance are discussed and compared with Durbin's ARMA. Twenty four hours forecast based on 2352 observation data is simulated. The forecast values are obtained and 168 data is used to validate the performances of the ARMA and AR models. The Mean Absolute Percentage Errors (MAPE), calculated from 24 forecasts over the day is used as statistical indicator to algorithms performances. For both load cases, AR's models show better performance than Durbin's ARMA by about two times. The results show clearly that Durbin's ARMA fails to great extent to cope with the rapid variation in load data, developing in high MAPE rates reaching a maximum value of 6.23% (case 2). On the other hand, AR's models show relatively consistent performance with significantly lower value reaching a maximum of about 2.77% (case 1) MAPE. The better results of AR show that, univariate method of all pole modeling is one of the alternative solutions to short term load forecast. The studies are being extended and possibly the modified version of forward-backward linear prediction could give promising results to this problem.

VI. ACKNOWLEDGMENT

The authors gratefully acknowledge Universiti Technologi PETRONAS, Malaysia for providing the facilities and financial support.

VII. REFERENCES

[1] Kandil, M.S.; El-Debeiky, S.M.; Hasanien, N.E., "Long-term load forecasting for fast developing utility using a knowledge-based expert system," Power Systems, IEEE Transactions on , vol.17, no.2, pp.491-496, May 2002.

[2] Qia Ding, "Long-Term Load Forecast Using Decision Tree Method," Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES , vol., no., pp.1541-1543, Oct. 29 2006-Nov. 1 2006.

[3] Gavrilas, M.; Ciutea, I.; Tanasa, C., "Medium-term load forecasting with artificial neural network models ," Electricity Distribution, 2001. Part 1: Contributions. CIRED. 16th International Conference and Exhibition on (IEE Conf. Publ No. 482) , vol.6, no., pp.5 pp. vol.6-, 2001.

[4] Asber, D.; Lefebvre, S.; Saad, M.; Desbiens, C., "Modeling of Distribution Loads for Short and Medium-Term Load Forecasting," Power Engineering Society General Meeting, 2007. IEEE , vol., no., pp.1-5, 24-28 June 2007.

[5] G. Gross, F.D. Galiana, " Short term load forecasting," Proc. IEEE, Vol. 75, No. 12, pp. 1558-1573, 1987.

[6] D.W. Bunn and E.D. Farmer, Comparative Models for Electrical Load Forecasting, John Wiley & Son, New York, 1985.

[7] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Time series analysis: Forecasting and control 3rd ed. Englewood Cliffs, NJ, Prentice Hall, 1994.

[8] M. T. Hagan and S. M. Behr,"The time series approach to short term load forecasting," paper 87 WM 044-1, presented at the IEEE Power Engineering Society Winter Meeting, New Orleans, LA, Feb. 1987.

[9] Amjady, N., "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," Power Systems, IEEE Transactions on , vol.16, no.4, pp.798-805, Nov 2001.

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

[11] Z. Baharudin and N. Kamel," One week ahead short term load forecasting," In Proceedings of the 7th IASTED International Conference on Power and Energy Sytems (EuroPES 2007), pages 91–96, Spain, Aug 2007.

[12] Z. Baharudin and N. Kamel, "AR Burg in one week ahead short term load forecasting," ," In Proceedings of the 2nd International Conference on Intelligent and Advanced Systems (ICIAS 2007), Malaysia, Nov 2007.

[13] Fan, S.; Chen, L., "Short-term load forecasting based on an adaptive hybrid method," Power Systems, IEEE Transactions on , vol.21, no.1, pp. 392-401, Feb. 2006.

[14] Kyung-Bin Song; Seong-Kwan Ha; Jung-Wook Park; Dong-Jin Kweon; Kyu-Ho Kim, "Hybrid load forecasting method with analysis of temperature sensitivities," Power Systems, IEEE Transactions on , vol.21, no.2, pp. 869-876, May 2006.

[15] Villalba, S.A.; Bel, C.A., "Hybrid demand model for load estimation and short term load forecasting in distribution electric systems," Power Delivery, IEEE Transactions on , vol.15, no.2, pp.764-769, Apr 2000.

[16] El Desouky, A.A.; Elkateb, M.M., "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA," Generation, Transmission and Distribution, IEE Proceedings- , vol.147, no.4, pp.213-217, Jul 2000.

[17] Taylor, J. W.; Buizza, R., "Neural Network Load Forecasting with Weather Ensemble Predictions," Power Engineering Review, IEEE , vol.22, no.7, pp.59-59, July 2002.

[18] Senjyu, T.; Takara, H.; Uezato, K.; Funabashi, T., "One-Hour-Ahead Load Forecasting Using Neural Networks," Power Engineering Review, IEEE , vol.22, no.1, pp.70-70, Jan. 2002

[19] Charytoniuk, W.; Chen, M.-S., "Very short-term load forecasting using artificial neural networks ," Power Systems, IEEE Transactions on , vol.15, no.1, pp.263-268, Feb 2000

[20] Wu, H. C.; Lu, C. N., "A Data Mining Approach for Spatial Modeling in Small Area Load Forecast," Power Engineering Review, IEEE , vol.22, no.3, pp.60-60, March 2002.

[21] Liang, R.-H.; Cheng, C.-C., "Combined regression-fuzzy approach for short-term load forecasting ," Generation, Transmission and Distribution, IEE Proceedings- , vol.147, no.4, pp.261-266, Jul 2000.

[22] Z. Baharudin, M. F. Jamaluddin, and N. Saad.," Fuzzy Logic Technique for Short Term Load Forecasting,"In Proceeding of the BICET, Brunei Darussalam, Aug 2005.

[23] Z. Baharudin, N. Saad, and R. Ibrahim.," A Fuzzy Logic Technique for Short Term Load forecasting,"In Proceeding of the 2nd ICAIET, Sabah, Malaysia, 2003.

[24] Al-Hamadi, H.M.; Soliman, S.A., "Fuzzy short-term electric load forecasting using Kalman filter," Generation, Transmission and Distribution, IEE Proceedings- , vol.153, no.2, pp. 217-227, 16 March 2006.

[25] S. J. Kiartzis and A. G. Bakirtzis. A Fuzzy Expert System for Peak Load Forecasting: Application to the Greek Power System. Proceedings of the 10th Mediterranean Electrotechnical Conference, 3:1097-1100, 2000.

[26] V. Miranda and C. Monteiro. Fuzzy Inference in Spatial Load Forecasting. Proceedings of IEEE Power Engineering Winter Meeting, 2:1063-1068, 2000.

[27] S. E. Skarman and M. Georgiopoulous. Short-Term Electrical Load Forecasting using a Fuzzy ARTMAP Neural Network. Proceedings of SPIE, 181-191, 1998.

[28] Chao-Ming Huang; Chi-Jen Huang; Ming-Li Wang, "A particle swarm optimization to identifying the ARMAX model for short-term load forecasting," Power Systems, IEEE Transactions on , vol.20, no.2, pp. 1126-1133, May 2005.

[29] Bo-Juen Chen; Ming-Wei Chang; Chih-Jen lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," Power Systems, IEEE Transactions on , vol.19, no.4, pp. 1821-1830, Nov. 2004.

[30] Ling, S.H.; Leung, F.H.F.; Lam, H.K.; Tam, P.K.S., "Short-term electric load forecasting based on a neural fuzzy network," Industrial Electronics, IEEE Transactions on , vol.50, no.6, pp. 1305-1316, Dec. 2003.

[31] Khotanzad, A.; Enwang Zhou; Elragal, H., "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," Power Systems, IEEE Transactions on , vol.17, no.4, pp. 1273-1282, Nov 2002.

[32] J. Durbin, "The fitting of time series models," Rev. Internat. Statist., vol. 23, pp. 233-244, 1960.

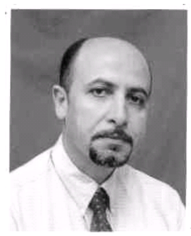
[33] C. W. Therrien, Discrete Random Signals and Statistical Signal Processing, Prentice-Hall, Englewood Cliff, NJ, 1992.

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

[35] National electricity market management company (2006, December). National electricity market management company. Retrieved January 2, 2007, from http://www.nemmco.com.au

VIII. BIOGRAPHIES

Zuhairi Baharudin is a faculty member at the Universiti Teknologi PETRONAS, Malaysia. He is currently pursuing his PhD at Universiti Teknologi PETRONAS. His areas of interests are Power Economics Operation and Control, Energy and Load Forecasting, Time Series Applications and Artificial Intelligence applications.



Nidal S. Kamel received his PhD degree (Hons) in statistical signal processing from the Technical University of Gdansk, Poland, in 1994. His research is focused on signal processing, parameter estimation, pattern recognition and mobile communications. He is currently with Universiti Teknologi PETRONAS, Malaysia. He is senior member of IEEE.

Dr. Ismail Musirin obtained Diploma of Electrical Power Engineering in 1987, Bachelor of Electrical Engineering (Hons) in 1990; both from Universiti Teknologi Malaysia, MSc in Pulsed Power Technology in 1992 from University of Strathclyde, United Kingdom and PhD in Electrical Engineering from Universiti Teknologi MARA (UiTM), Malaysia in 2005. He is currently an Associate Professor and the Chair, Centre of Electrical Power Engineering Studies (CEPES), Faculty of Electrical Engineering, UiTM. He has published 2 books and more than 100 technical papers in the international and national conferences, and international journals. He is a reviewer of IEEE transactions and IET journals. He is also appointed as the permanent reviewer for the World Scientific and Engineering Academy and Society (WSEAS) centered in Greece. His research interest includes power system stability, optimization techniques, distributed generator and artificial intelligence. He is also a member of IEEE, IEEE Power Engineering Society and Artificial Immune System Society (ARTIST).