International Journal of Applied Mathematics and Statistics,
Int. J. Appl. Math. Stat.; Vol. 55; Issue No. 2; Year 2016,
ISSN 0973-1377 (Print), ISSN 0973-7545 (Online)

Bayesian Multiperiod Forecasting for Indonesia Inflation Using ARMA Model

Zul Amry

Department of Mathematics,
State University of Medan, Indonesia
email : amryzul77@gmail.com

ABSTRACT

This paper present a Bayesian approach to find the Bayesian forecasting using ARMA model under Jeffrey's prior assumption with quadratic loss function applied to monthly national inflation of Indonesia from January 2000 to December 2014. The forecasting model be obtained based the marginal conditional posterior predictive density. To conclude whether the model is adequate is used the Ljung-Box Q statistic, while to look at the accuracy of the model is used the Root Mean Square error (RMSE)

Key word: ARMA model, Bayes theorem, national inflation

Mathematical Subject Classification: 62C10

1. INTRODUCTION

The classical forecasting have been depeloved by Box and Jenkins (1976). There are three steps are accomplished in the process of fitting the ARMA (p,q) model to a time series: identification of the model, estimation of the parameters, and application to forecast. In the classical approach, the parameters are considered fixed but unknown, whereas the Bayesian approach considers the parameters as random variables, which are described by their probability density function. Several of works relating to Bayesian forecasting in the ARMA model are Liu (1995), Amry and Baharum (2015) Fan & Yao (2008), Kleibergen & Hoek (1996), and Uturbey (2006). This paper focuses to find the Bayesian forecasting model for ARMA model using Jeffrey's prior with quadratic loss function.

2. MATERIALS AND METHODS

Bayes theorem calculates the posterior distribution as proportional to the product of a prior distribution and the likelihood function, the prior distribution is a probability model describing the knowledge about the parameters before observing the currently by the available data. Forecasting in the Bayesian approach is based on the construction of the marginal conditional posterior predictive distribution. Main idea of Bayesian forecasting is the predictive distribution of the future given the fast data follows directly from the joint probabilistic model, predictive distribution is derived from the sampling predictive density weighted by posterior distribution (Bijak, 2010). The method in this paper is study of literature by applying the Bayesian analysis under Jeffrey's assumption, whereas the materials are some theories in mathematics and statistics such as the ARMA model, Bayes theorem, integration, distribution of statistics and inflation data from January 2000 to December 2014. The forecasting model will be applied to forecast the data from January 2014 to December 2014 based on data from January 2000 to December 2013.

3. RESULTS

3.1. Likelihood function

The k–step–ahead point forecast of y_{n+k} is defined by:

$$\hat{y}(k) = E(y_{n+k} \mid S_n^*) \tag{3.1}$$

where $S_n^* = (y_1, y_2, \dots, y_{n+k-1})$

The ARMA (p,q) model is defined by :

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j} + e_t \tag{3.2}$$

where $\{e_t\}$ is sequence of $i\,i\,d$ normal random variables with $e_t \sim N(0, \tau^{-1})$, $\tau > 0$ and unknown, ϕ_i and θ_j are parameters. Residuals are written as:

$$e_t = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j e_{t-j} \tag{3.3}$$

By conditioning the first p observations and letting $e_p = e_{p-1} = \dots = e_r = 0$, where r = min(0, p+1–q), one may approximate by Box & Jenkins [4], the likelihood function for $\Psi = (\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q)$ and τ based S_n^* is:

$$L(\Psi, \tau \mid S_n^*) \propto \tau^{\frac{(n+k-1)-p}{2}} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=p+1}^{n+k-1}\left(y_t - \sum_{i=1}^{p}\phi_i y_{t-i} - \sum_{j=1}^{q}\theta_j e_{t-j}\right)^2\right]\right\} \tag{3.4}$$

Furthermore the equation (3.4) can be expressed as:

$$L(\Psi, \tau \mid S_n^*) \propto \tau^{\frac{(n+k-1)-p}{2}} \exp\left\{-\frac{\tau}{2}\left[\sum_{t=p+1}^{n+k-1} y_t^2 - 2\Psi' \sum_{t=p+1}^{n+k-1} y_t B_{t-1} + \sum_{t=p+1}^{n+k-1} (\Psi' B_{t-1})^2\right]\right\} \tag{3.5}$$

where $B_t = (y_t, y_{t-1}, \dots, y_{t+1-p}, e_t, e_{t-1}, \dots, e_{t+1-q})$

By letting: $U = \begin{bmatrix} y_p & y_{p+1} & \cdots & y_{n+k-2} \\ y_{p-1} & y_p & \cdots & y_{n+k-3} \\ \vdots & \vdots & \vdots & \vdots \\ y_1 & y_2 & \cdots & y_{n+k-1-p} \\ \hat{e}_p & \hat{e}_{p+1} & \cdots & \hat{e}_{n+k-2} \\ \hat{e}_{p-1} & \hat{e}_p & \cdots & \hat{e}_{n+k-3} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{e}_{p+1-q} & \hat{e}_{p+2-q} & \cdots & \hat{e}_{n+k-1-q} \end{bmatrix}$, $x_0 = \begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_{n+k-1} \end{bmatrix}$, $W=UU^T$ and $V=U^T x_0$,

where $\hat{e}_t = y_t - \sum_{i=1}^{p}\hat{\phi}_i y_{t-i} - \sum_{j=1}^{q}\hat{\theta}_j \hat{e}_{t-j}$, $t = p+1, p+2, \dots, n$, $\hat{\phi}_i$ and $\hat{\theta}_j$ are maximum likelihood estimator of ϕ

and θ_j, the likelihood function in equation (3.5) can be expressed as:

$$L(\Psi, \tau | S_n^*) \propto \tau^{\frac{(n+k-1-p)}{2}} exp \left\{ -\frac{\tau}{2}\left[\sum_{t=p+1}^{n+k-1} y_t^2 - 2\Psi^T V + \Psi^T W \Psi \right] \right\} \quad (3.6)$$

3.2. Posterior distribution

Based the likelihood function in equation (3.6), the Jeffrey's prior is:

$$\pi_{Jeff}(\tau^{-1}) \propto \tau^{-1} \quad (3.7)$$

By applying the Bayes theorem to equation (3.6) and (3.7), the posterior of Ψ and τ^{-1} is:

$$\pi(\Psi, \tau^{-1} | S_n^*) = L(\Psi, \tau^{-1} | S_n^*) x \pi_{Jeff}(\tau^{-1})$$

$$\propto \tau^{\frac{(n+k-1-2p)+p}{2}-1} exp \left\{ -\frac{\tau}{2}\left[\Psi^T W \Psi - 2\Psi^T V + \sum_{t=p+1}^{n+k-1} y_t^2 \right] \right\} \quad (3.8)$$

3.3. Conditional posterior predictive

Based on $e_t = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j e_{t-j}$ with $e_t \sim N(0, \tau^{-1})$ will be gotten that

$$f(e_t | S_n^*, \Psi, \tau^{-1}) = (2\pi\tau^{-1})^{-\frac{1}{2}} exp\left\{ -\frac{\tau}{2}(e_t)^2 \right\}.$$

If it expressed in y_t, it will becomes:

$$f(y_t | S_n^*, \Psi, \tau^{-1}) = (2\pi\tau^{-1})^{-\frac{1}{2}} exp \left\{ -\frac{\tau}{2}\left[y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j e_{t-j} \right]^2 \right\} \quad (3.9)$$

Based the equation (3.9), the conditional predictive density of Y_{n+k} is:

$$f(y_{n+k} | S_n^*, \Psi, \tau^{-1}) = (2\pi\tau^{-1})^{-\frac{1}{2}} exp \left\{ -\frac{\tau}{2}\left[y_{n+k} - \sum_{i=1}^{p} \phi_i y_{n+k-i} - \sum_{j=1}^{q} \theta_j e_{n+k-j} \right]^2 \right\} \quad (3.10)$$

$$\propto \tau^{\frac{1}{2}} exp \left\{ -\frac{\tau}{2}\left[y_{n+k} - \sum_{i=1}^{p} \phi_i y_{n+k-i} - \sum_{j=1}^{q} \theta_j e_{n+k-j} \right]^2 \right\}$$

$$\propto \tau^{\frac{1}{2}} exp \left\{ -\frac{\tau}{2}\left[y_{n+k} - \left(\sum_{i=1}^{p} \phi_i y_{n+k-i} + \sum_{j=1}^{q} \theta_j e_{n+k-j} \right) \right]^2 \right\} \quad (3.11)$$

Verify the expression $\sum_{i=1}^{p} \phi_i y_{n+k-i} + \sum_{j=1}^{q} \theta_j e_{n+k-j}$ become:

$$\sum_{i=1}^{p} \phi_i y_{n+k-i} + \sum_{j=1}^{q} \theta_j e_{n+k-j} = \phi_1 y_{n+k-1} + \phi_2 y_{n+k-2} + \dots + \phi_p y_{n+k-p} + \theta_1 e_{n+k-1} + \theta_2 e_{n+k-2} + \dots + \theta_q e_{n+k-q}$$

$$= (\phi_1 \quad \phi_2 \quad \dots \quad \phi_p \quad \theta_1 \quad \theta_2 \quad \dots \quad \theta_q) \begin{pmatrix} y_{n+k-1} \\ y_{n+k-2} \\ \vdots \\ y_{n+k-p} \\ e_{n+k-1} \\ e_{n+k-2} \\ \vdots \\ e_{n+k-q} \end{pmatrix} = \Psi^T B_{n+k-1}$$

where $B_{n+k-1} = (y_{n+k-1}, y_{n+k-2}, \dots, y_{n+k-p}, e_{n+k-1}, e_{n+k-2}, \dots, e_{n+k-q})$, the equation (3.11) can be written as:

$$f(y_{n+k} | S_n^*, \Psi, \tau^{-1}) \propto \tau^{\frac{1}{2}} exp\left\{ -\frac{\tau}{2}\left[y_{n+k} - \Psi^T B_{n+k-1} \right]^2 \right\}$$

$$\propto \tau^{\frac{1}{2}} exp\left\{-\frac{\tau}{2}\left[y_{n+k}^2 - 2\Psi^T B_{n+k-1} y_{n+k} + \left(\Psi^T B_{n+k-1}\right)^2\right]\right\}$$

$$\propto \tau^{\frac{1}{2}} exp\left\{-\frac{\tau}{2}\left[y_{n+k}^2 + \Psi^T R\Psi - 2\Psi^T B_{n+k-1} y_{n+k}\right]\right\} \quad (3.12)$$

where $R = B_{n+k-1} \otimes B_{n+k-1}^T$ and $\left(\Psi^T B_{n+k-1}\right)^2 = \Psi^T R\Psi$.

Based on the equation (3.9) and (3.12), the conditional posterior predictive density of Y_{n+k} is:

$$f_p\left(y_{n+k} \mid S_n^*, \Psi, \tau^{-1}\right) = \pi\left(\Psi, \tau^{-1} \mid S_n^*\right) f\left(y_{n+k} \mid S_n^*, \Psi, \tau^{-1}\right)$$

$$\propto \tau^{\frac{(n+k-1-2p)+p+1}{2}-r} exp.\left\{-\frac{\tau}{2}\left[\begin{array}{l}\Psi^T Z\Psi - \Psi^T(V + B_{n+k-1} y_{n+k}) - \\ (V^T + B_{n+k-1}^T y_{n+k})\Psi + y_{n+k}^2 + \sum_{t=p+1}^{n+k-1} y_t^2\end{array}\right]\right\} \quad (3.13)$$

where Z = W + R

3.4. Marginal conditional posterior predictive

The marginal conditional posterior predictive density of Y_{n+k} is obtained by integrating equation (3.13) with respect to Ψ and τ^{-1}, that is:

$$f_p\left(y_{n+k} \mid S_n^*\right) = \int\int f_p\left(y_{n+k} \mid S_n^*, \Psi, \tau^{-1}\right) d\Psi\, d\tau \propto \left[n+k-1-p + \frac{\left(y_{n+k} - \left(1 - B_{n+k-1}^T Z^{-1} B_{n+k-1}\right)^{-1}\left(B_{n+k-1}^T Z^{-1} V\right)\right)^2}{\frac{\sum_{t=p+1}^{n+k-1} y_t^2 - V^T Z^{-1} V}{(n+k-1-p)\left(1 - B_{n+k-1}^T Z^{-1} B_{n+k-1}\right)}}\right]^{-\frac{(n+k-p)}{2}} \quad (3.14)$$

The marginal conditional posterior predictive density of Y_{n+k} is a univariate student's t-distribution on $(n+k-1-p)$ degrees of freedom with mean $\mu = \left(1 - B_{n+k-1}^T Z^{-1} B_{n+k-1}\right)^{-1}\left(B_{n+k-1}^T Z^{-1} V\right)$.

3.5 Point forecast

For quadratic loss function, the point forecast of Y_{n+k} is the posterior mean of the marginal conditional posterior predictive, that is:

$$E\left(Y_{n+k} \mid S_n^*\right) = \left(1 - B_{n+k-1}^T Z^{-1} B_{n+k-1}\right)^{-1}\left(B_{n+k-1}^T Z^{-1} V\right) \quad (3.15)$$

4. APPLICATION

The forecasting model is applied for inflation monthly data based on Consumen Price Index (CPI), which have collected for the period of January 2000 to December 2014, where the data period of January 2000 - December 2013 are used to forecast the period of January-December 2014.

Based on the plot of series, ACF, PACF in Figure 1, 2 and 3 can be concluded that the collection of data is stationary and based on the value of $\hat{\phi}$, $\hat{\theta}$, and AIC in Table 1, a suitable model for the data is ARMA *(0,1)* with $\hat{\theta}$ =0.2696.

Figure 1: plot of time series data

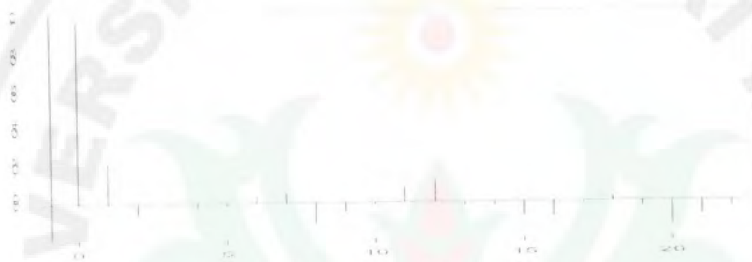

Figure 2: plot of autocorrelation function

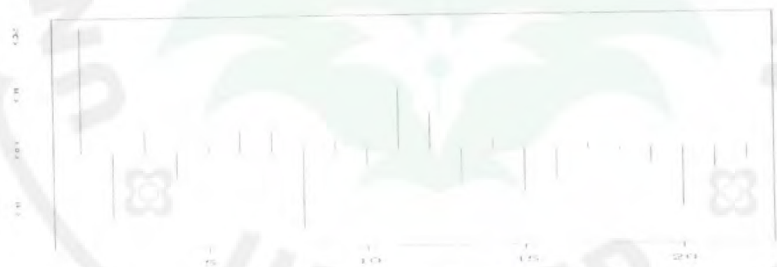

Figure 3: plot of partial autocorrelation function

Table 1: value of $\tilde{\phi}$, $\tilde{\theta}$ and *AIC*

Model	$\tilde{\phi}$	$\tilde{\theta}$	*AIC*
ARMA(*1,0*)	0.2205	-	424.53
ARMA(*0,1*)	-	0.2696	422.62
ARMA(*1,1*)	-0.2738	0.5250	423.80

By using the Bayesian forecasting model in equation (3.15), the results of point forecast are presented in Table 2 as follows:

Table 2: Result of point forecast

Period: *j*	Observation: y_j	Forecast $\hat{y}_j$
169	1.07	0.2630176
170	0.26	0.1398343
171	0.08	0.0697564
172	-0.02	0.0361444
173	0.16	0.0183071
174	0.43	0.1939937
175	0.93	0.5924600
176	0.47	0.2870595
177	0.27	0.1515332
178	0.47	0.2759849
179	1.50	0.7344441
180	2.46	0.8568429

Testing the result of point forecast on the forecasting model to conclude whether is adequate, an investigation to autocorrelations of residuals by using the Ljung-Box Q statistic (Wei, 1994) is:

$$Q = n(n+2)\sum_{k=1}^{K}\frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(K-p-q) \qquad (4.1)$$

If $Q > \chi^2_{1-\alpha}(K-p-q)$ the adequacy of the model is reject at the level α. where n is the sample size, ρ_k is the autocorrelation of residuals at lag k and K is the number of lags being test. By supporting the calculation in the Table 3 it can be obtained that $Q = 10.8059$, whereas $\chi^2_{95\%}(9) = 16.9190$ and $Q < \chi^2_{95\%}(9)$, so can be concluded that the results of point forecast are adequate.

Table. 3: Computation of Ljung-Box Q statistic

Number of lags: k	Autocorrelation ρ_k	ρ_k^2
1	0.321	0.103041
2	-0.062	0.003844
3	-0.150	0.022500
4	-0.056	0.003136
5	0.029	0.000841
6	-0.024	0.000576
7	-0.166	0.027556
8	-0.301	0.090601
9	-0.255	0.065025
10	-0.057	0.003249
Q = 10.8059		

One measure to determine the accuracy of a forecasting model is Root Mean Square (RMSE), (Assis et all. 2010), that is:

$$RMSE = \sqrt{\frac{ESS}{n}} \qquad (4.2)$$

where ESS = the error sum of square, and n = the number of observations. By calculation in Table 4, colum 1 through 5 contain the period (t), observation (y_t), result of forecast ($\hat{y}_t$), residual (e_t), and squares of e_t (e_t^2) is obtained RMSE = 0.336964.

Table. 4 : Computation of RMSE

t	y_t	$\hat{y}_t$	e_t	e_t^2
169	1.07	0.2630176	0.8069824	0.6512205939
170	0.26	0.1398343	0.1201657	0.0144397955
171	0.08	0.0697564	0.0102436	0.0001049313
172	-0.02	0.0361444	-0.0561444	0.0031521937
173	0.16	0.0183071	0.1416929	0.0200768779
174	0.43	0.1939937	0.2360063	0.0556989736
175	0.93	0.5924600	0.3375400	0.1139332516
176	0.47	0.2870595	0.1829405	0.0334672265
177	0.27	0.1515332	0.1184668	0.0140343827
178	0.47	0.2759849	0.1940151	0.0376418590
179	1.50	0.7344441	-0.2344441	0.0549640360
180	2.46	0.8568429	0.6031571	0.3637984873
ESS = 1.362533				
RMSE = 0.336964				

5. CONCLUSION

This paper focus on the study of mathematics in the Bayesian forecasting of ARMA model under the Jeffrey's prior assumption with quadratic loss function. By observing the behavior of ACF graph shaped the damped sine wave in Figure 2 and based on the smallest AIC value in Table 1 can be concluded that the time series data for monthly national inflation of Indonesian starting on January 2000 until December 2013 was stationary a suitable model for the data is ARMA (0,1) with $\tilde{\theta} = 0.2696$. Mathematically the point forecast in equation (3.15) is obtained based the posterior predictive mean of the marginal conditional posterior predictive density. Finally, based an investigation to autocorrelations of residuals by using the Ljung-Box Q statistic, can be conclude that the forecasting model is adequate with accuracy measure RMSE=0.336964.

REFERENCES

Amry, Z. and Baharum, A. (2015). Bayesian Multiperiod Forecasting for Arma Model under Jeffrey's Prior. *Mathematics and Statistics 3(3), 65-70.*

Assis, et all. (2010). Forecasting Cocoa Bean Prices Using Univariate Time Series Models. *Journal of Arts & Commerce. ISSN 2229-4686, Vol.- I, Issue- I, 71-80*

Bain, L.J. and Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*, 2nd. Duxbury Press, Belmont, California.

Bijak, J. (2010). Bayesian Forecasting and Issues of Uncertainty. *Centre for Population Change. University of Southampton, 1-19.*

Box, G.E.P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden Day, San Francisco.

Enders, W. (1995). *Applied Econometric Time Series* John Wiley & Son. Inc. New York.

Fan, C. and Yao, S. (2008). Bayesian Approach for ARMA Process and its Application *International Business Research, 1(4), 49-55.*

Gelman, A. (2008), Objections to Bayesian Statistics, *Bayesian Analysis 3(3), 4 45–450.*

Geweke, J and Whiteman, C., (2004). *Bayesian Forecasting*. Department of Economics, UniverSity of Iowa.

Hussein, H. M. A. (2008). Bayesian Analysis of the Autoregressive Moving Average Model with Exogenous Inputs Using Gibb Sampling. *Journal of Applied Sciences Research, 4(12),1885-1892.*

Ihaka, R. (2005). *Time Series Analysis*. Statistics Department University of Auckland.

Kleibergen, F and Hoek, H (1996). Bayesian Analysis of ARMA model using Non informative Prior. *Paper, Econometric Institute, Erasmus University, Rotterdam, 1–24*

Liu, S. I. (1995). Bayesian Multiperiod Forecasts for ARX Models. *Ann. Statist. Math. Vol. 47, no. 2, 211-224.*

Liu, S. I. (1995). Comparison of Forecasts for AR Models Between A Random Coefficient Approach and A Bayesian Approach.*Commun. Statist.-Theory Meth., 24(2), 319-333*

Pole, A., West, M. and Harrison, J. (1994). *Applied Bayesian Forecasting and Time Series Analysis* Chapman and Hall, New York.

Ramachandran, K. M. and Tsokos, C.P. (2009). *Mathematical Statistics with Applications*, Elsevier Academic Press. San Diego, California.

Safadi, T. and Morettin, P. A. (2000). Bayesian Analysis of Threshold Autoregressive Moving Average Models. *Sankhya: The Indian Journal of Statistics, Volume 62, Series B, Pt. 3, 353-371.*

Uturbey, W. (2006). Identification of ARMA Model by Bayesian Methods Applied to Streamflow Data *9th International Conference Probabilistic Method Applied to Power Systems KTH (11- 15 June 2006, Stockholm, Sweden), 1-7.*

Wei, W.W. S.(2006). *Time Series Analysis Univariate and Multivariate Methods*. Addison Wesley Publishing Company, Inc. Canada.