RESEARCH PAPER | Statistics | Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

Weighted Average ARMA Model

KEYWORDS: ARMA, WARMA, MSE, K-S test, Theil's U-Statistic.

P. Ramakrishna Reddy	* B. Sarojamma
Department of Statistics, S.V. University, Tirupati – 517502, Andhra Pradesh, India	Department of Statistics, S.V. University, Tirupati – 517502, Andhra Pradesh, India. *Corresponding Author

ABSTRACT *In the present study, we divided the data into two parts, for each set of data we have fitted suitable ARMA model. By using these ARMA models we fitted Weighted Average ARMA (WARMA) model for part I & part II separately. WARMA model is compared with ARMA (3, 2) model by using Mean Square Error (MSE) criteria. MSE criteria is used for choosing best model between the two models, Theil's U-Statistic is used for testing accuracy of models and Kolmogrov-Smirnov (K-S) test is used for goodness of fit.*

1. Introduction

Autoregressive moving average plays an important role in forecasting. As compared with Auto regression (AR) models and moving average (MA) models, Autoregressive moving average model is the best. Upon changing orders of Auto regression (p) and moving average (q), we get several ARMA models. The development of forecasting ARMA model requires a suitable data so that correct relationships can be established. Forecasting methodology is based mainly upon model building.

ARMA model:

ARMA model is the best model compared with MA and AR models. The basic elements of autoregressive and moving average models can be combined to produce a great variety of models. Combination of pth order autoregressive model and qth order moving average model called an ARMA (p, q) model and is expressed as

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

2. Methodology

Forecast models are fitted for data to estimate future values. For estimation of future values, we have to fit appropriate models, some of them are moving average, auto regression, exponential smoothing, autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) with various orders and also improved models like autoregressive conditional heteroscedastisity (ARCH), Generalized autoregressive conditional heteroscedastisity (GARCH), Exponentially generalized autoregressive conditional heteroscedastisity (EGARCH), Integrated generalized autoregressive conditional heteroscedastisity (IGARCH), etc. All these models are improved models of one after another by modifying basic models. Modification of model is because of fitting the best model for data than the previous one.

Now we improving ARMA model, and if data is large, split data into two parts, perform autoregressive moving average to both parts. If the first part of data possesses auto regression of order 'p' with moving average of order 'q' and second part of data possess ARMA (r, s) equations are as follows

Part I: ARMA (p, q)

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_{1,t} - \theta_1 e_{1,t-1} - \theta_2 e_{1,t-2} - \cdots - \theta_q e_{1,t-q}$$

Part II: ARMA (r, s)

$$y_t = c + \delta_1 y_{t-1} + \delta_2 y_{t-2} + \cdots + \delta_r y_{t-r} + e_{2,t} - \varphi_1 e_{2,t-1} - \varphi_2 e_{2,t-2} - \cdots - \varphi_s e_{2,t-s}$$

Root Mean Square Error (RMSE) is used for ARMA (p, q) and ARMA (r, s) of models for selecting the best model for part I and part II. A minimum RMSE value of various ARMA (p, q) is the best model for part I and similarly we estimate a suitable ARMA (r, s) model for part II by using RMSE

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}}$$

By choosing the best ARMA models for part I and part II of data are ARMA (p_i, q_i) and ARMA (r_i, s_i) respectively. We now estimating weighted ARMA model.

$$WARMA = \alpha\, ARMA(p_i, q_i) + \beta\, ARMA(r_i, s_i) \qquad \text{for part I}$$
$$= \beta\, ARMA(p_i, q_i) + \alpha\, ARMA(r_i, s_i) \qquad \text{for part II}$$

where $\alpha + \beta = 1$

Weighted autoregressive moving average is calculated for ARMA (p_i, q_i), ARMA (r_i, s_i) and are consider as variables, and their weights are α, β. For part I data, we taking high weight 'α' for ARMA (p_i, q_i), and ARMA (r_i, s_i) possessing low weight 'β'. For part II data, ARMA (r_i, s_i) take high weight 'α' and ARMA (p_i, q_i) takes low weight 'β'.

where $\alpha + \beta = 1$, and α takes highest values like 0.9, 0.8, 0.7 etc. and β takes 0.1, 0.2, 0.3, etc. respectively. By giving more weight to part I best ARMA model and least weight to best ARMA of part II with weights α and β, we get the weighted ARMA model for part I. By giving least weight to best ARMA model of part I and more weight to best ARMA model of part II, we get weighted ARMA model for part II.

Theil's U-Statistics is used for accuracy, mean square error (MSE) is used for selection of best model among ARMA (3, 2) and WARMA.

Theils U- Statistic

Theil's U Statistic is used for measuring forecasting accuracy. It is a relative comparison of formal forecasting methods with naive approaches and also squares errors involved so that large errors are given much more weight

Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

than small errors.

Theil's U-Statistic U = $\sqrt{\dfrac{\sum_{t=1}^{n-1}\left(\dfrac{F_{t+1}-Y_{t+1}}{Y_t}\right)^2}{\sum_{T=1}^{n-1}\left(\dfrac{Y_{t+1}-Y_t}{Y_t}\right)^2}}$

Mean Square Error (MSE)

$$MSE = \frac{\sum_{t=1}^{n} e_t^2}{n}$$

$$MSE \text{ of } WARMA = \frac{\sum_{j=1}^{m} e_j^2}{m}$$

3. Empirical investigation

Empirical investigations are carrieed out by taking time points from 1976 to 2012 and time series values are annual average temperatures. We know that ARMA (3, 2) is a better model for the data among ARMA (1, 1), ARMA (1, 2), ARMA (1, 3), ARMA (2, 1), ARMA (2, 2), ARMA (2, 3), ARMA (3, 1), ARMA (3, 2) and ARMA (3, 3) models. Data contains 37 years of data. Data splitted into two parts, first part contains from 1976 to 1994 and another part contains from 1995 to 2012. We perform all 9 models of ARMA for part I data from 1976 to 1994 and also for part II data from 1995 to 2012 by using RMSE, we select the best model among 9 models.

We chose ARMA (1, 1) model with lowest RMSE 0.774 and is the best model compared with all other models of ARMA.

Best model for part I: ARMA (1, 1)

Y_t = - 0.999 Y_{t-1} – 0.997 e_{t-1} + e_t – 217.557

An ARMA model possesses lowest RMSE is the best model.

Best model for part II: ARMA (2, 1)

Y_t = 0.141 Y_{t-1} – 0.219 Y_{t-2} – 0.998 e_{t-1} + e_t – 53.839

ARMA models	(1, 1)	(1, 2)	(1, 3)	(2, 1)	(2, 2)	(2, 3)	(3, 1)	(3, 2)	(3, 3)
RMSE of Part I	0.774	0.804	0.817	0.784	0.827	0.805	0.795	0.814	0.835
RMSE of Part II	1.364	1.184	1.191	1.071	1.098	1.132	1.099	1.147	1.222

We now fitting weighted ARMA model with weights α = 0.9 and β = 0.1 for ARMA (1, 1), and ARMA (2, 1) for part I. For part II data, weight is switching that is weighted ARMA (2. 1) takes 0.9 and weighted of ARMA (1, 1) takes 0.1.

Weighted ARMA = 0.9 ARMA (1, 1) + 0.1 ARMA (2, 1) for part I

= 0.1 ARMA (1, 1) + 0.9 ARMA (2, 1) for part II

Theil's U-Statistics: The calculated values of Theil's U-Statistic for ARMA (3, 2) and WARMA models are as follows

Theil's U-Statistics of ARMA (3, 2) is 0.0070

Theil's U-Statistics of WARMA is 0.0054

The weighted ARMA model is the best model for forecasting.

4. Comparison of ARMA with WARMA:

By using mean square error criteria, we compare ARMA and WARMA models for the entire data (without splitting).

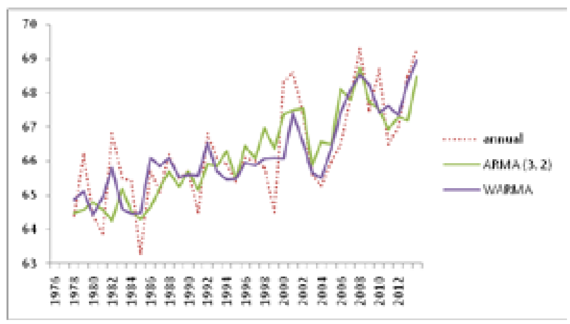
MSE of WARMA: 0.6504

MSE of ARMA (3, 2): 0.8203

MSE of WARMA < MSE of ARMA (3, 2), therefore we conclude that weighted autoregressive moving average model is better when compared with autoregressive moving average (3, 2).

5. Graphical representation of ARMA and WARMA models

By taking time 't' on x-axis and time series values (annual production), the estimated values of autoregressive moving average (ARMA (3, 2)) model and estimated values of weighted autoregressive moving average (WARMA) are on Y-axis gives the following graph.



6. Summary and conclusions

We estimated weighted ARMA model. If data is large then divide the data into two parts and estimate the best ARMA models for two parts of data. By taking the best ARMA models of part as variables and their weights are α and β. We computing weighted ARMA model. For two parts of data, we estimated ARMA equations which suit for each part using RMSE.

Weighted autoregressive moving average is calculated as ARMA (p_i, q_i), ARMA (r_i, s_i) and these are variables and their weights are α, β. For part I data, we take high weight 'α' for ARMA (p_i, q_i), ARMA (r_i, s_i) possessing low weight 'β'. For part II data, ARMA (r_i, s_i) take high weight 'α' and for ARMA (p_i, q_i) takes low weight 'β'.

$$WARMA = \alpha\, ARMA(p_i, q_i) + \beta\, ARMA(r_i, s_i) \qquad \textit{for part I}$$
$$= \beta\, ARMA(p_i, q_i) + \alpha\, ARMA(r_i, s_i) \qquad \textit{for part II}$$

Data is divided into two parts, first part contains 1976 to 1994 and second part contains 1995 to 2012. We performed 9 ARMA models for two parts and by using RMSE, we choose the best models of two parts.

We choose ATRMA (1, 1) as the best model for Part I, ARMA (2, 1) is the best for Part II.

For the given data the best estimated model for two parts are

ARMA (1, 1) for Part I: Y_t = - 0.999 Y_{t-1} – 0.997 e_{t-1}

Volume : 4 | Issue : 8 | August 2014 | ISSN - 2249-555X

$+ e_t - 217.557$

ARMA (2, 1) for Part II: $Y_t = 0.141\ Y_{t-1} - 0.219\ Y_{t-2} - 0.998\ e_{t-1} + e_t - 53.839$

We have fitted other models for two parts and they are

ARMA (2, 1) for Part I: $Y_t = 0.904\ Y_{t-1} - 0.576\ Y_{t-2} - e_{t-1} + e_t - 217.568$

ARMA (1, 1) for Part II: $Y_t = 0.145\ Y_{t-1} - 0.984\ e_{t-1} + e_t - 53.840$

By using the above 4 ARMA equations, we fitting weighted autoregressive moving average (WARMA) model.

Weighted ARMA = 0.9 ARMA (1, 1) + 0.1 ARMA (2, 1) for part I

= 0.1 ARMA (1, 1) + 0.9 ARMA (2, 1) for part II

We checked the best model among fitted models and old ARMA by using MSE.

In WARMA model, we computed MSE of WARMA and ARMA models.

MSE of WARMA: 0.6504

MSE of ARMA: 0.8203

MSE of WARMA < MSE of ARMA (3, 2)

Therefore, we conclude that WARMA model is better model compared with ARMA (3, 2) model.

We tested accuracy of fitted models by using Theil's U test statistic. We showed empirically that new fitted WARMA models are accurate models.

Finally we used Kolmogrov-Smirnov test for testing goodness of fit for new fitted ARMA models.

REFERENCE Holt, C.C. (1957), "Forecasting seasonal and trends by exponentially weighted moving averages", ONR research Memorandum 52, and R.J. (2004), International Journals of Forecasting, 20, 5-13. | Spyros Makridakis, Steven C. Wheelwright and Rob J. Hyndman, (1998) "Forecasting methods and applications", Third Edition, Wiley india Pvt. Ltd., New Delhi. | Winters, P.R. (1960), "Forecasting sales by exponentially weighted moving averages", Management Science, 6, 324-342. | Box, G.E.P. and G.M. Jenkins (1970), time series analysis: Forecasting and control, San Francisco: Holden-Day. | Box, G.E.P. and D.A. Pierce (1970), Distribution of the residual autocorrelations in autoregressive-integrated moving-average time series models, Journal of the American Statistical Association, 65, 1509-1526. | Mc Kenzie, E. (1984), General exponential smoothing and the equivalent ARIMA process, Journal of Forecasting, 3, 333-334. | Mc Kenzie, E. (1986), Error analysis for Winters' additive seasonal forecasting system, International Journal of Forecasting, 2, 373-382. | http://www.srh.noaa.gov/graphicast.php?site=fwd