Calendar variation model based on ARIMAX for forecasting sales data with Ramadhan effect

Muhammad Hisyam Lee[1]
Suhartono[1, 2]
[1]*Department of Mathematics, Universiti Teknologi Malaysia*
[2]*Department of Statistics, Institut Teknologi Sepuluh Nopember*
E-mail: mhl@utm.my, suhartono@statistika.its.ac.id

Nor Aishah Hamzah[3]
[3]*Institute of Mathematical Sciences University of Malaya*
E-mail: aishah_hamzah@um.edu.my

ABSTRACT

Many economics and business data are compiled each month and are available as monthly time series. Such time series may be subjected to two kinds of calendar effects, namely trading day and holiday effects. The objective of this research is to develop a calendar variation model based on ARIMAX method for forecasting sales data with Ramadhan effect. In most Islamic countries, the trade activities frequently contain calendar variation pattern due to the consumption to certain products usually follows lunar calendar, instead of following common solar calendar. Firstly, this research focuses on the development of model building procedure to find the best calendar variation model. In this ARIMAX method, the deterministic and stochastic trends are examined. The procedure for fitting the ARIMAX model consists of only modeling calendar variation effect, its error, and adaptively combining the calendar variation effect and ARIMA model. This procedure is applied in modeling and forecasting of sales data, i.e. the monthly sales of Moslem boys' clothes in Indonesia. The results show that the proposed model yields better forecast at out-sample data compared to the Decomposition method and Seasonal ARIMA, and Neural Networks.

Keywords: *ARIMAX, Calendar Variation, Lunar Calendar, Model Building Procedure.*

Introduction

Generally, many economics and business data are monthly time series which may be subjected to two kinds of calendar effects. Firstly, the levels of economics or business activities may change depending on the day of the week. Since the composition of days of the week varies from month to month and year to year, the observed series may be affected by such variation. Such effects, particularly due to the composition of trading days (or work days) in each month, are referred to as trading day effects (Liu, 1980, 1986; Hillmer, 1982; Pfeffermann and Fisher, 1982; Cleveland and Devlin, 1982; Bell and Hillmer, 1983; Balaban, 1995; Mills and Coutts, 1995; Sullivan, Timmermann, and White, 2001; Gao and Kling, 2005; Al-Khazali, Koumanakos, and Pyun, 2008; Alagidede, 2008a; Evans and Speight, 2010). Secondly, in addition to trading day variation, some traditional festivals or holidays, such as Easter, Chinese New Year, and Jewish Passover, are set according to lunar calendars and the dates of such holidays may vary between two adjacent months in the Gregorian calendar from year to year. Since business activities and consumer behavior patterns may be greatly affected by such holidays, the observed time series may vary substantially depending on whether a particular month contains such holidays or not. Such effects are referred to as holiday effects (Pfeffermann and Fisher, 1982; Cleveland and Devlin, 1982; Liu, 1986; Ariel, 1990; Brockman and Michayluk, 1998; Vergin and Mcginnis, 1999; Sullivan, Timmermann, and White, 2001; Seyyed, Abraham, and Al-Hajji, 2005; Alagidede, 2008b).

ISBN 978-967-363-157-5

Islamic calendar is also a lunar calendar based on twelve lunar months in a year of 354 or 355 days, used to date events in many Muslim countries (concurrently with the Gregorian calendar), and used by Muslims everywhere to determine Islamic holy days and festivities. Indonesia, a country with Moslem majority, also uses Islamic calendar, especially in determining religious holidays, such as Eid-holidays. The calendar variation arises from these holidays. The effects can be known from the amount of tradings around Eid-holidays, both pre-to-post Eid-holidays. Example given in this research is on the amount of tradings of Moslem boys' clothes from a company in Indonesia.

Time series analysis for data with calendar variation effect requires special treatment. Common classical time series analysis, such as Decomposition method and ARIMA model, maybe fail to capture such variation. In this paper, we propose a comprehensive procedure for modeling time series data in the presence of calendar variation. The proposed methods are ARIMAX methods. This paper is organised as follows: a brief literature review about the calendar variation effects on time series analysis, data description and the modeling method, results based on the model, evaluations of the the forecast accuracy, conclusions and recommendation.

Literature Review

Calendar variation effect on time series analysis has been investigated by many researchers since 1980s. Liu (1980) and Cleveland and Devlin (1980) were among the first researchers who studied the effect of calendar variation. Liu (1980) studied the effect of holiday variation on the identification and estimation of ARIMA models. He suggested modifications of ARIMA models by including holiday information as deterministic input variable(s) and used the monthly highway traffic volume in Taiwan as a case study. Whereas, Cleveland and Devlin (1980) proposed two sets of diagnostic methods for detecting calendar effects in monthly time series, i.e. by using spectrum analysis and time domain graphical displays. They stated that these methods could be used in an initial analysis to decide if calendar adjustment is necessary and can be used on an adjusted series to determine if the adjustment has properly removed all of the calendar effects. Three series were used as examples, i.e. the weekday series, bell system toll revenues, and international airline series.

Later, Cleveland and Devlin (1982) developed new procedures for modeling and adjustment of calendar variation effects. They used international airline passengers series as a case study and compared the results to the X-11 calendar estimation procedures. Moreover, Pfeffermann and Fisher (1982) also proposed a new method of preadjusting time series for the effects of movements in festival dates and the variability in the number of working days. They applied the method into two series, i.e. the index of industrial production and workseekers at labour exchange in Israel. Furthermore, Hillmer (1982) proposed the modeling of time series which contain trading day variation and an approach to building models for such time series. He used the monthly Wisconsin telephone data as a case study. Then, Bell and Hillmer (1983) studied about the modeling of time series data that include calendar variation by using a sum of ARIMA and regression models. The ARIMA models were used for modeling the autocorrelation, trends, and seasonality, whereas trading day variation and Easter holiday variation were modeled by regression-type models. As a case study, they used the series retail sales of lumber and building materials that be obtained from the U.S. Census Bureau.

In addition, Liu (1986) proposed a comprehensive and easy-to-use method for the identification of the degree of differencing and appropriate ARMA model in univariate ARIMA modeling when the effects of calendar variation, such as trading day and holiday effects, were presented on the monthly time series data. He used the monthly highway traffic volume recorded by the Taiwan Highway Bureau as an example of the presence of the holiday effect and the monthly outward station movements (disconnections) of the Wisconsin Telephone Company as an example of the presence of the trading day effect.

Besides that, many reports that have investigated the effects of calendar variation on financial time series, especially in stock market. Thaler (1987) discussed about the seasonal anomalies movements in security prices, including the effects of weekend, holiday, turn of the month and intraday on stock market. Then, Ariel (1990) showed that returns observed on days preceding a public holiday are, on average, many times greater than returns on other trading days. This result was supported by Mills and Coutts (1995) who studied about calendar effects in the London Stock Exchange FT-SE indices, Brockman and Michayluk (1998) who investigated the persistent holiday effect on the NYSE, AMEX, and NASDAQ exchanges, and also by Vergin and McGinnis (1999) who studied about holiday effect in stock market returns in the United States.

Similarly, Balaban (1995) investigated day of the week effects in an emerging stock market of a developing country, namely Turkey. Whereas, Brooks and Persand (2001) examined the evidence for a day-of-the-week effect in five Southeast Asian stock markets: South Korea, Malaysia, the Philippines, Taiwan and Thailand. Later, Holden, Thompson, and Ruangrit (2005) carried out similar research on calendar effects of daily stocks returns in the Thai Stock Market using several models, such as OLS linear regression, General Autoregressive Conditional Heteroscedasticity (GARCH) and Threshold ARCH (TARCH). Furthermore, Gao and Kling examined calendar effects in Chinese stock market, particularly monthly and daily effects. Recently, anomaly in Africa's largest stock markets have been studied by Alagidede (2008a) for the day of the week effects and Alagidede (2008b) for the month-of-the-year and pre-holiday seasonality effects. In addition, Al-Khazali, Koumanakos and Pyun (2008) examined calendar anomalies in the Athens Stock Exchange by using stochastic dominance analysis. More recently, Evans and Speight (2010) provided an analysis of intraday periodicity, calendar and announcement effects in Euro exchange rate volatility.

Otherwise, Sullivan, Timmermann and White (2001) showed that the case of calendar effects in stock returns implied the dangers of data mining. They proved that there appeared to be very substantial evidence of systematic abnormal stock returns related to the day of the week, the week of the month, the month of the year, the turn of the month, holidays, and so forth. Additionally, the effects of moving calendar events such as the Muslim holy month of Ramadhan in stock returns also has studied by Seyyed, Abraham and Al-Hajji (2005).

Data

This research uses monthly sales of Moslem boys' clothes from a Moslem garment company in Indonesia as a case study for the period 2000: M1 until 2008: M12. The first 8 years data are used as in-sample data, while the remaining data as out-sample data. Figure 1 illustrates the data in a time series plot.

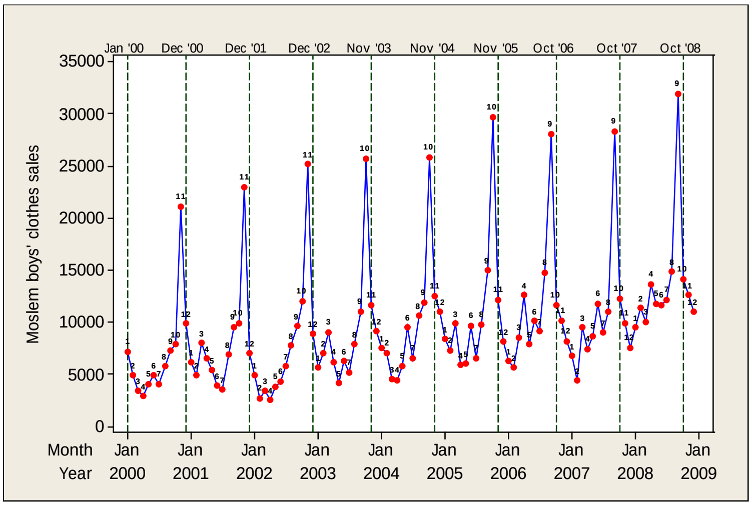

Figure 1. Monthly sales from January 2000 – December 2008

This plot shows the presence of calendar variation due to the celebration of Eid ul-Fitr every year. A vertical (dotted line) was included in this plot to emphasize the months of Eid holidays that occurred during this period. The exact dates of the events are shown in Table 1.

Table 1. Date of Eid holidays since 2000 to 2008

Year	Date	Year	Date
2000	08-09 January	2004	14-15 November
	28-29 December	2005	03-04 November
2001	17-18 December	2006	23-24 October
2002	06-07 December	2007	12-13 October
2003	25-26 November	2008	01-02 October

Thus, the graph clearly shows a highest sales in 1 month before Eid holidays for every year. Also, we could see that the second and third highest sales are commonly occurred at 2 month before and during the month of Eid celebration respectively.

Modeling Method

Regression in time series context has the same form as general linear regression. By assuming output or dependent series, y_t, $t = 1,2,\ldots,n$, is being influenced by a collection of possible inputs or independent series, where the inputs are fixed and known, this relation can be expressed as linear regression model (Shumway and Stoffer, 2006). If there exist a trend in the data, then we can model as follows:

$$y_t = \beta_0 + \beta_1 t + w_t \tag{1}$$

where w_t is the error component, which is usually sequence of noise process and independently identically distributed normal with mean 0 and variance σ_w^2.

Similarly, data with calendar variation can also be modeled by using regression. Linear regression model for data having calendar variation is

$$y_t = \beta_0 + \beta_1 V_{1,t} + \beta_2 V_{2,t} + \cdots + \beta_p V_{p,t} + w_t \tag{2}$$

where $V_{p,t}$ is dummy variable for p-th calendar variation effect. The number of calendar variation effects can be identified based on the time series plot of the data. Ljung-Box statistics may be employed to test whether the sequence σ_w^2 is white noise.

ARIMAX Model for Calendar Variation

ARIMA model is a common model for data forecasting. Seasonal ARIMA model belongs to a family of flexible linear time series models that can be used for modeling many different types of seasonal as well as non seasonal time series. The seasonal ARIMA model can be expressed as (see Wei, 2006; Box, Jenkins, and Reinsel, 2008; Cryer and Chan, 2008):

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \tag{3}$$

where

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

$$\Phi_P(B^S) = 1 - \Phi_1 B - \Phi_2 B^{2S} - \cdots - \Phi_P B^{PS}$$

$$\theta_q(B) = 1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q$$
$$\Theta_Q(B^S) = 1-\Theta_1 B-\Theta_2 B^{2S}-\cdots-\Theta_Q B^{QS},$$

and S is the seasonal period, B is the backshift operator, and ε_t is a sequence of white noise with zero mean and constant variance. Box and Jenkins (1976) proposed a set of an effective model building strategies for seasonal ARIMA based on the autocorrelation structures in a time series.

ARIMAX model is ARIMA model with additional variable (see Cryer and Chan, 2008). In this research, there are two kind of additional variables, i.e. the dummy variable(s) for calendar variation effects only, and dummy variable(s) for calendar variation effects and deterministic trend. The first model is known as ARIMAX with stochastic trend by implementing a difference non seasonal and/or seasonal, and the second with deterministic trend (without differencing order). The general seasonal ARIMA model, which is shown in Eq. (3), can also be written as

$$y_t = \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t . \tag{4}$$

Therefore, ARIMAX model with stochastic trend is

$$y_t = \beta_1 V_{1,t} + \beta_2 V_{2,t} + \cdots + \beta_p V_{p,t} + \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t . \tag{5}$$

Whereas, ARIMAX model with deterministic trend is

$$y_t = \gamma\, t + \beta_1 V_{1,t} + \beta_2 V_{2,t} + \cdots + \beta_p V_{p,t} + \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)^D}\varepsilon_t . \tag{6}$$

The proposed ARIMAX model building procedure, in the presence of calendar variation effects in the time series data is described as follows:

Step 1: Determination of dummy variable for calendar variation period.

Step 2: Remove the calendar variation effect from the response by fitting Eq. (2) for model with stochastic trend, or fitting Eq. (1) and (2) simultaneously for model with deterministic trend, in order to obtain the error, w_t.

Step 3: Model w_t using ARIMA model (see Box-Jenkins procedure).

Step 4: The order of ARIMA model found from Step 3 is used to model the real data and dummy variables of calendar variation effect as input variables simultaneously as Eq. (5) and (6) for model with stochastic and deterministic trends respectively.

Step 5: Test parameter significance and perform diagnostic checks until the process is stationary and ε_t appears as white noise processes.

The flowchart of this proposed procedure is illustrated on appendix.

Results, analysis and evaluation

The initial step which is based on the time series plot is used to identify calendar variation period affecting the data. The data are fitted with a linear regression model where y_t as the response and the calendar variation effects as the predictors. For this, three dummy variables are used for evaluating calendar variation effect, i.e. 2 months before, 1 month before, and during the month of Eid celebration. Linear regression model parameters are estimated by Ordinary Least Square (OLS) estimation which yields

$$y_t = 6883 + 3488\, V_t + 19000\, V_{t-1} + 4814\, V_{t-2} + w_t . \tag{7}$$

where V_t is the dummy variable during Eid celebration, V_{t-1} is the dummy variable a month before and V_{t-2} is dummy variable two months before Eid celebration.

All of these coefficients are significant, but the process is non-stationary and the error does not appear as white noise processes. This is due to trend and seasonal pattern in the data. The time series plot of the error is shown in Figure 2. The ACF of this error tends to die down slowly in seasonal pattern and this suggests seasonal differencing. The differencing error of seasonal order with period 12 results ACF and PACF plot in Figure 3. As such there are four orders of ARIMA models that may be appropriate for the error data, i.e. $(1,0,0)(1,1,0)^{12}$, $(1,0,0)(0,1,1)^{12}$, $(0,0,1)(1,1,0)^{12}$, and $(0,0,1)(0,1,1)^{12}$. From the results of parameter estimation and diagnostic checking only two models yield error that appear to be white noise, i.e. ARIMA$(1,0,0)(1,1,0)^{12}$ and $(1,0,0)(0,1,1)^{12}$.

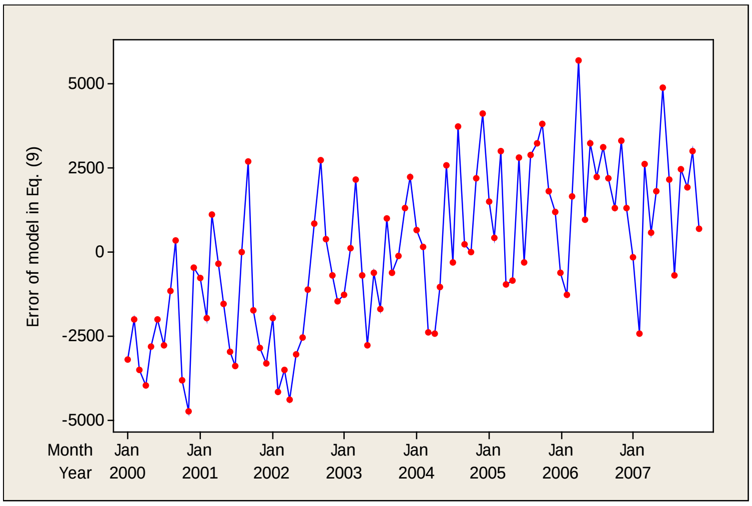

Figure 2. Time series plot of the error of model in Eq. (7)

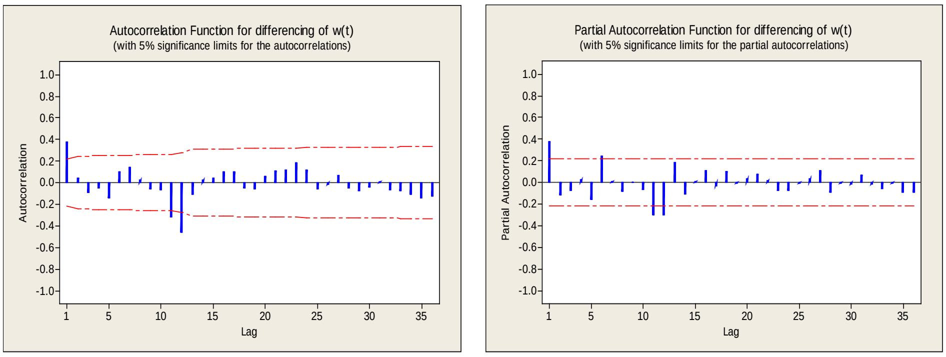

Figure 3. ACF and PACF plots for seasonal differencing of the error in Eq. (7)

Using these two models, we now develop the ARIMAX models with stochastic trend. By adding calendar variation effect as dummy variables in ARIMA model for the given data, ARIMAX model with stochastic trend could be developed. By using SAS, the models which yield all significantly parameters and white noise errors are then considered. The estimated models are as follows.

(i). Model 1: ARIMA$(1,0,0)(1,1,0)^{12}$ and V_{t-1} without a constant, i.e.

$$y_t = 13980.3 V_{t-1} + \frac{1}{(1-0.57236B)(1+0.66800B^{12})(1-B^{12})}\varepsilon_t . \tag{8}$$

(ii). Model 2: ARIMA(1,0,0)(0,1,1)^{12} and V_{t-1} without a constant, i.e.

$$y_t = 14010.2\,V_{t-1} + \frac{(1-0.54755B^{12})}{(1-0.57860B)(1-B^{12})}\varepsilon_t\,. \tag{9}$$

The ARIMAX models with deterministic trend are developed by adding calendar variation effect as dummy variables and trend t simultaneously in ARIMA model for the given data. By applying the proposed model building procedure and using SAS, the ARIMAX models with deterministic trend which yield all significantly parameters and white noise errors are as follows.

(i). Model 3: ARIMA(1,0,0)(1,0,0)^{12}, t, V_t, V_{t-1}, V_{t-2} with a constant, i.e.

$$y_t = 4514.2 + 52.78989\,t + 2757.3\,V_t + 18027.1V_{t-1} + 3577.4\,V_{t-2} + \frac{1}{(1-0.35438B)(1+0.27445B^{12})}\varepsilon_t\,. \tag{10}$$

(ii). Model 4: ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2} with a constant, i.e.

$$y_t = 4374.5 + 54.91470\,t + 3114.4\,V_t + 18354.5\,V_{t-1} + 3864.9\,V_{t-2} + \frac{1}{(1-0.31586B)}\varepsilon_t\,. \tag{11}$$

Performance Evaluation

To evaluate the performance of the proposed calendar variation models, the root mean squared error (RMSE) is selected as an evaluation index. The RMSE of out-sample data is defined as

$$\mathrm{RMSE}_{\mathrm{out}} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}}\ ,$$

where n is the number of forecasts. Whereas, the RMSE of in-sample data for some methods such Seasonal ARIMA and ARIMAX is defined as

$$\mathrm{RMSE}_{\mathrm{in}} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n-p}}\ ,$$

where p is the number of parameters.

Two classical time series methods, namely decomposition and Seasonal ARIMA, and neural networks (NN) as modern time series method are also applied to forecast this time series data. In this paper, the most popular NN model, namely three-layer feed-forward neural networks (FFNN) is used for forecasting the data. For example, notation 3-2-1 means that architecture of FFNN is 3 inputs, 2 neurons in hidden layer, and 1 output. The inputs selection is based on the best lag AR of ARIMA model (Faraway and Chatfield, 1998), i.e. lag 12, 24, and 35.

The results of RMSEs obtained using the proposed calendar variation models, two other classical time series models, and neural networks, both in-sample and out-sample data, are listed in Table 2. Based on RMSE of in-sample data, neural networks with 10 neurons in hidden layer yields better prediction than other models. Moreover, based on RMSE of out-sample data, ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2}, with a constant yields better forecast than other models. Thus, these results show that the proposed calendar variation model, i.e. ARIMAX with deterministic trend, yield the best forecast. Additionally, these comparisons also show that NN tends to yield very accurate forecast only at in-sample data. It in lines with many previous comparison results about the ability of NN for time series forecasting.

Another technique to show the best model is visually from the plot of error (difference between actual out-sample data and forecasts). The following error plot in Figure 4 is among the overall best model based on each approach in out-sample data, i.e. additive decomposition, ARIMA $([35],0,0)(1,1,0)^{12}$, neural networks, and ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2}, with a constant.

Table 2. Comparison of forecast accuracy between models

Model	RMSE	
	In-sample	Out-sample
Decomposition		
Additive Decomposition	4254.52	5392.80
Multiplicative Decomposition	4216.07	5605.25
Seasonal ARIMA		
$\text{ARIMA}([35],0,0)(1,1,0)^{12}$	3333.68	2683.23
$\text{ARIMA}(0,0,[35])(0,1,1)^{12}$	3486.44	3171.31
$\text{ARIMA}(0,0,[35])(1,1,0)^{12}$	3409.11	3092.27
$\text{ARIMA}([35],0,0)(0,1,1)^{12}$	3420.09	2705.16
Neural Networks – Input lag 12, 24, 35		
3-1-1 (inputs-neurons-output)	2008.95	2762.73
3-2-1	1528.50	2320.17
3-3-1	1170.96	3161.81
3-4-1	998.14	2444.09
3-5-1	892.66	3704.16
3-6-1	703.80	4620.82
3-7-1	552.61	3188.30
3-8-1	481.16	5035.99
3-9-1	335.21	4956.48
3-10-1	**225.34**	3188.81
ARIMAX		
$\text{ARIMA}(1,0,0)(1,1,0)^{12}$, V_{t-1}, without a constant	1784.97	3197.92
$\text{ARIMA}(1,0,0)(0,1,1)^{12}$, V_{t-1}, without a constant	1887.85	3748.40
$\text{ARIMA}(1,0,0)(1,0,0)^{12}$, t, V_t, V_{t-1}, V_{t-2}, with a constant	1727.03	2452.08
ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2}, with a constant	1769.67	**2012.15**

Time series plots in Figure 4(a1-d1) aid comparison between actual and fitted models. The visible difference between these proposed, the classical, and modern models is the accuracy in detecting the calendar variation, i.e. Eid celebration effects. Figure 4(a1) shows that the fit is very poor, particularly during Eid for all years considered. Figure 4(b1) illustrates ARIMA$([35],0,0)(1,1,0)^{12}$ fits which is reasonably good except at Eid 2003. The proposed model, i.e. ARIMAX model (see Figure 4(d1)), and neural networks (see Figure 4(c1)) yields better results. The both predictions are in general closed to the actual including those with calendar effects.

Error plots on the right (see Figure 4(a2, b2, c2, and d2)) also show the accuracy of the fits. The effects of Eid celebration still exist in Figure 4(a2) which indicates that there exists some structure in the error component thus leading to inadequacy of the fit. Figure 4(b2) shows an extreme error due to the mistakes in prediction. This results from ignoring calendar variation effect in the model. Figures 4(c2) and 4(d2) showed that the error component are random with narrower prediction interval than those given in Figure 4(a2) and 4(b2). This means that the proposed model and neural networks work well for data with calendar variation effect.

In addition, Table 3 shows the out-sample forecast values in 2008 for each best model and the error could be seen in Figure 5. This figure shows that the model with overall error close to zero provides the best model. These results illustrate that the ARIMAX models are better than decomposition method, ARIMA model, and NN for forecasting the given time series. Thus, the performance evaluation of forecast accuracy shows that the ARIMAX model, i.e. ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2}, with a constant, as in Eq. (11) is the best model with smallest RMSE.

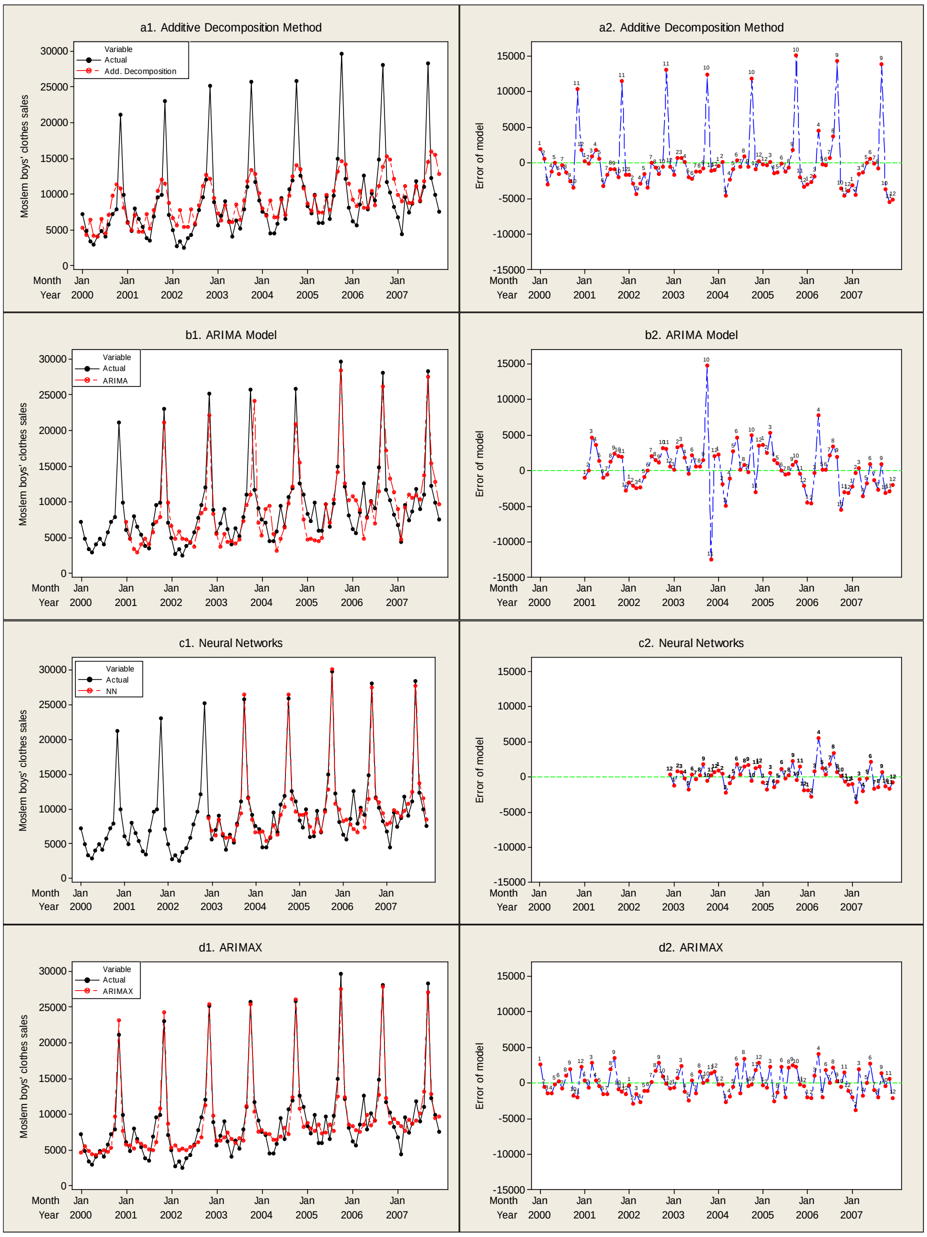

Figure 4. Time series plot of (a) Additive Decomposition, (b) ARIMA, (c) Neural Networks, and (d) ARIMAX (left figures) and their errors (right figures)

Table 3. Comparison the forecast values at out-sample data

t	Y_t	Add. Decomposition	SARIMA	Neural Networks	ARIMAX
97	9514	10600.9	6746.7	7013.5	9046.0
98	11351	9625.7	7783.7	10764.9	9549.2
99	10075	11746.7	9615.4	9307.8	9745.7
100	13644	9444.0	10708.0	9880.6	9845.3
101	11802	9410.0	9766.9	9354.6	9914.3
102	11651	11833.9	11552.8	10195.5	9973.7
103	12192	9793.1	9441.8	9716.6	10030.0
104	14818	12418.5	15731.8	17720.7	13950.3
105	31963	15159.0	31555.7	30157.4	28495.0
106	14165	16653.3	12017.0	12916.7	13309.8
107	12611	16154.1	8292.8	9611.7	10250.3
108	10973	13424.1	6368.1	8370.1	10305.3
	RMSE	5392.80	2683.23	2320.17	2012.15

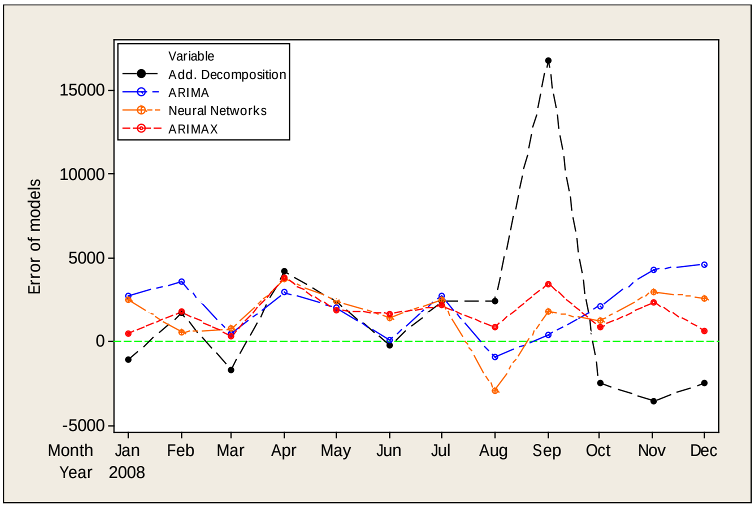

Figure 5. Plot of error comparison between Additive Decomposition, ARIMA, Neural Networks, and ARIMAX models

Conclusion and future works

In general, time series data that is affected by lunar calendar needs special treatment. The application of ordinary Decomposition method and ARIMA model usually yield spurious results, particularly about seasonal pattern and the presence of outliers. The proposed models, i.e. ARIMAX models, yield better prediction for out-sample data, compared to those of decomposition method, ARIMA model, and neural networks. For the Moslem boy's clothes data, ARIMAX model with deterministic trend, i.e. ARIMA(1,0,0), t, V_t, V_{t-1}, V_{t-2}, with a constant, provides a reasonably good model, especially for out-sample forecasting while neural networks perform well for in-sample data.

In addition, this study proposes a procedure for building a calendar variation model based on ARIMAX method. The results show that the proposed procedure has been applied well at certain real data. Further research is needed to validate this procedure to other real data sets and to compare the forecast accuracy with other more intelligent methodologies such as fuzzy time series or genetic algorithm.

Acknowledgment

This research was supported by Research Management Centre, Universiti Teknologi Malaysia (UTM), Malaysia, under the Post-Doctoral Fellowship Scheme for the Project: "*New Hybrid Wavelet Neural Networks for Time Series Forecasting*".

References

Alagidede, P. (2008a). Day of the week seasonality in African stock markets. *Applied Financial Economics Letters*, 4, 115-120.

Alagidede, P. (2008b). Month-of-the-year and pre-holiday seasonality in African stock markets. *Stirling Economics Discussion Paper* 2008-23, Department of Economics, University of Stirling.

Al-Khazali, O.M., Koumanakos, E.P., & Pyun, C.S. (2008). Calendar anomaly in the Greek stock market: Stochastic dominance analysis. *International Review of Financial Analysis*, 17, 461–474.

Ariel, R.A. (1990). High Stock Returns before Holidays: Existence and Evidence on Possible Causes. *The Journal of Finance*, 45(5), 1611-1626.

Balaban, E. (1995). Day of the week effects: new evidence from an emerging stock market. *Applied Economics Letters*, 2, 139-143.

Bell, W.R., & Hillmer, S.C. (1983). Modeling Time Series with Calendar Variation. *Journal of the American Statistical Association*, 78, 526-534.

Box, G.E.P., & Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. San Fransisco: Holden-Day, Revised edn.

Box, G.E.P., Jenkins, G.M., & Reinsel, G.C. (2008). *Time Series Analysis, Forecasting and Control*. 4th edition, New Jersey: John Wiley & Sons, Inc.

Brockman, P., & Michayluk, D. (1998). The Persistent Holiday Effect: Additional Evidence. *Applied Economics Letters*, 5, 205-209.

Brooks, C., & Persand, G. (2001). Seasonality in Southeast Asian stock markets: some new evidence on day-of-the week effects. *Applied Economics Letters*, 8, 155-158.

Cleveland, W.S. & Devlin, S.J. (1980). Calendar Effects in Monthly Time Series: Detection by Spectrum Analysis and Graphical Methods. *Journal of the American Statistical Association*, 75(371), 487-496.

Cleveland, W.S. & Devlin, S.J. (1982). Calendar Effects in Monthly Time Series: Modeling and Adjustment. *Journal of the American Statistical Association*, 77(379), 520-528.

Cryer, J.D., & Chan, K.S. (2008). *Time Series Analysis: with Application in R*. 2nd edition, New York: Springer-Verlag.

Evans, K.P., & Speight, A.E.H. (2010). Intraday periodicity, calendar and announcement effects in Euro exchange rate volatility. *Research in International Business and Finance*, 24, 82-101.

Faraway, J., & Chatfield, C. (1998). Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics*, 47, 231-250.

Gao, L., & Kling, G. (2005). Calendar Effects in Chinese Stock Market. *Annals of Economics and Finance*, 6, 75-88.

Hillmer, S.C. (1982). Forecasting Time Series with Trading Day Variation. *Journal of Forecasting*, 1, 385-395.

Holden, K., Thompson, J., & Ruangrit, Y. (2005). The Asian crisis and calendar effects on stock returns in Thailand. *European Journal of Operational Research*, 163, 242–252.

Liu, L.M. (1980). Analysis of Time Series with Calendar Effects. *Management Science*, 2, 106-112.

Liu, L.M. (1986). Identification of Time Series Models in the Presence of Calendar Variation. *International Journal of Forecasting*, 2, 357-372.

Mills, T.C. & Coutts, J.A. (1995). Calendar effects in the London Stock Exchange FT-SE indices. *The European Journal of Finance*, 1, 79-93.

Pfeffermann, D. & Fisher, J.M. (1982). Festival and Working Days Prior Adjustments in Economic Time Series. *International Statistical Review*, 50(2), 113-124.

Seyyed, F.J., Abraham, A., & Al-Hajji, M. (2005). Seasonality in stock returns and volatility: The Ramadan effect. *Research in International Business and Finance*, 19, 374-383.

Shumway, R.H., & Stoffer, D.S. (2006). *Time Series Analysis and Its Applications with R Examples.* 2nd edition, Berlin: Springer.

Sullivan, R., Timmermann, A., & White, H. (2001). Dangers of data mining: The case of calendar effects in stock returns. *Journal of Econometrics*, 105, 249-286.

Thaler, R. (1987). Anomalies: Seasonal Movements in Security Prices II: Weekend, Holiday, Turn of the Month, and Intraday Effects. *The Journal of Economic Perspectives*, 1(2), 169-177.

Vergin, R.C., & McGinnis, J. (1999). Revisiting the Holiday Effect: Is it on Holiday?. *Applied Financial Economics*, 9(5), 477-482.

Wei, W.W.S. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. 2nd edition, California: Addison-Wesley Publishing Company, Inc.

Appendix

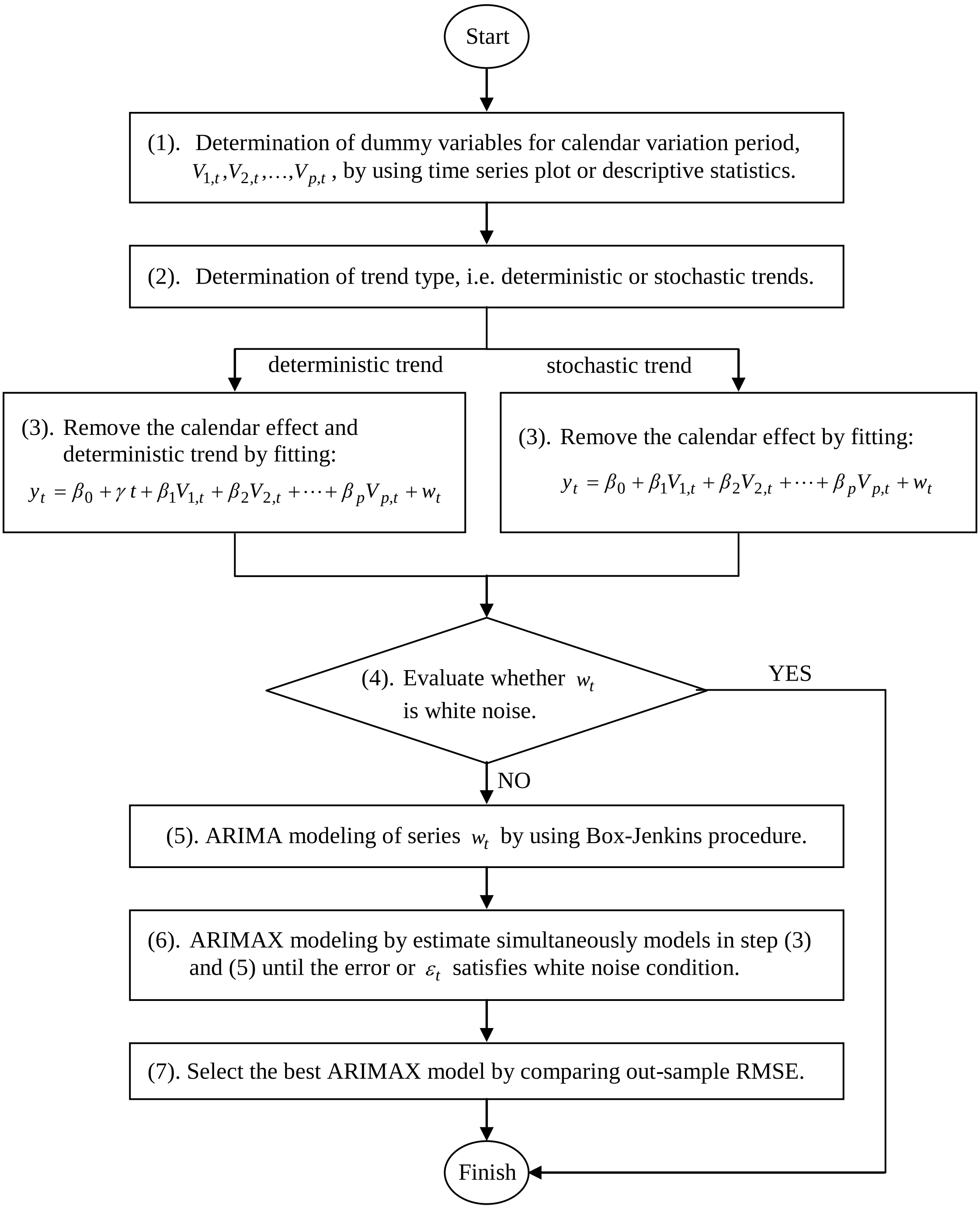

Figure 6. Model building procedure for calendar variation model based on ARIMAX method.