eProceedings Chemistry 2 (2017) 288–290

eISSN 2550-1453 http://eproceedings.chemistry.utm.my/

Modeling El Niño southern oscillation index using time series forecasting

Nor Radwa Ismail and Shariffah Suhaila Syed Jamaludin
Department of Mathematical Sciences, Faculty of Science, UniversitiTeknologi Malaysia,,81310 Johor Bahru, Johor, Malaysia.
Corresponding Author: suhailasj@utm.my

Article history:
Received 11 October 2017
Accepted 25 November 2017

ABSTRACT

Southern Oscillation Index (SOI) is measured as the difference in air pressure across the Pacific Ocean, between Tahiti in the south-east and Darwin in the west. SOI is one of the indices that are often used to analyze and predict the changes in El Niño Southern Oscillation phenomenon. Many statistical models have been developed using SOI indices in forecasting. The objective of this study is to find the best method among the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA), Single Exponential Smoothing and Double Exponential Smoothing in forecasting the monthly SOI. SOI data from January 1990 to December 2015 with a total of 25 years were employed in this study. Akaike Information Criterion (AIC) and the Sum of Square Error (SSE) were used as goodness of fit test in selecting the best model. The result indicated that the Box-Jenkins ARIMA is a suitable method in forecasting SOI values compared to others based on the smallest SSE.

Keywords: Southern Oscillation Index; Box-Jenkins Autoregressive Integrated Moving Average (ARIMA); Single Exponential Smoothing; Double Exponential Smoothing; Akaike Information Criterion; Sum of Square Error.

1. INTRODUCTION

El Niño is defined by unusually warm ocean temperatures in the Equatorial Pacific, as opposed to La Niña, which is characterized by unusually cold ocean temperatures in the Equatorial Pacific. Based on the report from National Geographic that had been published by Howard [1], El Niño phenomenon which occurred in October 2015 was the warmest with 0.98°C (1.76°F) above the 20th century average for the month. The Southern Oscillation Index (SOI) is a standardized index which is measured based on the difference in observed sea level pressure between Tahiti and Darwin, Australia. The positive measures of the SOI indicate La Niña events while negative phase represent El Niño. El Niño Southern Oscillation (ENSO)phenomenon gives negative impacts on global and local scale. Many lives have been lost and livelihoods are affected by this natural disaster. Lal et al. [2] stated that the effect of ENSO on global climate, through inter-annual fluctuations in temperature, precipitation and atmospheric circulation at distant locations is termed as teleconnection. The effect of El Niño also affected on human health. Velayudhan [3] mentioned that it is expected to have more mosquitoes capable of spreading the virus Zika resulted from the occurrence of El Niño. According to Kovats [4], El Niño climate phenomenon is a complex and unfair situation, which ultimately led to the death and obliteration. Hence, it is important to conduct a statistical analysis on the effect of ENSO phenomenon, modeling as well as predicting the value of SOI. The main objectives of the study are to profile El Niño Southern Oscillation Index and rainfall using descriptive measures, to measure the relationship between El Niño Southern Oscillation Index and rainfall using correlation analysis, to find the best forecasting model for monthly SOI and lastly to forecast the monthly Southern Oscillation Index using the selected model. Minitab 16 was used to analyze the result.

2. STUDY AREA AND DATA

The selected rainfall stations in Sabah and Sarawak will be used in this study. Monthly rainfall of four principal meteorological stations namely Kota Kinabalu (KK), Tawau (T), Miri (M), and Bintulu (B) are chosen. Figure 2.1 and Figure 2.2 show the location of these stations. SOI data from January 1990 to December 2015 with a total of 25 years were used in this study.

Figure 2.1 Locations of Kota Kinabalu and Tawau stations

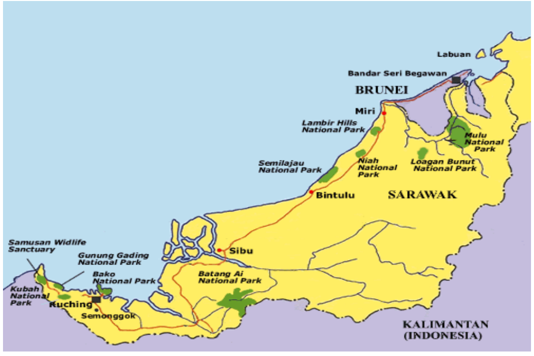

Figure 2.2 Locations of Miri and Bintulu stations

3. METHODOLOGY

The process of ARIMA, Single Exponential Smoothing and Double Exponential Smoothing were described in this research. ARIMA model predicts a value in a response time series as a linear combination of its own past values, past errors (also called shocks or innovations), and current and past values of other time series These methods are systematic way to choose the best possible models for forecasting. ARIMA models are univariate as they use only the history of the time series in order to express how the variables react with the earlier stochastic variation. The process for the method of Box Jenkins involves four stages which are model identification, parameter estimation, diagnostic checking and forecasting.

Let ∇^d the operator at difference value, d, z_t be the time series data/observation and e_t be the white noise and B the time –lagged operator/back shift operator. The general form of the ARIMA (p, q) process is

$$\Phi(B)z_t = \Phi_p \nabla^d z_t = \theta_q(B)e_t$$

where $\nabla^d = (1-B)^d$ is a difference operator, $\Phi_p(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_P B^P$ is an autoregressive operator at step p and $\Phi_q(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \cdots - \Phi_q B^q$ is an autoregressive operator at step q.

The simple exponential smoothing method is applied for time series that have no trend (stationary) and the mean (or level) of the time series y_t is slowly changing over time. The equation model is given as

$$F_{t+1} = \alpha y_t + (1-\alpha)F_t$$

where F_{t+1} represent forecast for the next period, α be the smoothing constant, , y_t be the observed values of series in period t and F_t be the old forecast for period t.

Double exponential smoothing is used when the data have trend and do not have a seasonal component. The general formula for double exponential smoothing is

$$F_{t+m} = L_t + mb_t$$

where $L_t = \alpha y_t + (1-\alpha)(L_{t-1} + b_{t-1})$ is a level series, $b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$ is trend estimates, α be smoothing constant for the level $(0 \le \alpha \le 1)$, γ be smoothing constant for the trend $(0 \le \gamma \le 1)$ and m be periods to be forecast into the future.

4. RESULT AND ANALYSIS

Statistics of ENSO and rainfall are summarized using descriptive statistics. Based on the results, Kota Kinabalu shows the highest monthly mean rainfall occur in October which is nearly 384.0 mm while the lowest in February with approximately 74.4 mm whereas in Tawau, the highest monthly mean rainfall occur in July which is nearly 206.17 mm while the lowest in February with 108.68 mm. Miri station deduced that the highest monthly mean rainfall occur in December which is 359.66 mm while the lowest in February with nearly 152.4 mm. We found that the highest monthly mean rainfall occurred in January with 479.17 mm while the lowest was recorded in May with 215.18 mm stated in Bintulu. Therefore, we can conclude that the highest monthly mean rainfall was recorded at Bintulu in January with nearly 479.17 mm due to the Northeast Monsoon flow from October to March and expected to bring more rainfall. The lowest mean rainfall is recorded in February at Kota Kinabalu with nearly 74.4mm. Correlation analysis is conducted to represent the strength of relationship between SOI and parameters tested. From the analysis, Kota Kinabalu shows high correlation between rainfall amount and SOI with r value of 0.73 in February. The correlation is significant since the p value= 0.0001which is less than 0.05. Tawau records the weakest correlation with the value, r = -0.32 in July. This correlation is significant at $\alpha = 0.05$ since the p values is 0.12. Miri was recorded weak correlation with the r values of -0.05 in February. Lastly in Bintulu, the result shows moderate correlation with the r values of 0.50 in September. Therefore, we can conclude that the highest correlations values between monthly rainfall and monthly SOI are recorded at Kota Kinabalu, $r = 0.73$ in February while the lowest correlation with $r = -0.32$ are computed in July at Tawau.

The Box-Jenkins ARIMA method can be classified into four steps; model identification, model estimation, diagnostic checking and forecasting. The data were analysed by using Minitab 16. ARIMA, Single Exponential Smoothing, and Double Exponential Smoothing method were used in data analysis. The values of SSE are compared for each method. ARIMA model is found to be the best method in forecasting SOI values based on the smallest value of SSE. The results are shown in Table 4.1.

Table 4.1 Sum of Square (SSE) values using ARIMA, Single Exponential Smoothing, and Double Exponential Smoothing method.

MONTH	ARIMA	SINGLE EXPONENTIAL SMOOTHING	DOUBLE EXPONENTIAL SMOOTHING
JAN	29.9370	44.7150	49.2178
FEB	47.8476	57.1134	61.0420
MAR	31.5815	34.1561	36.4226
APR	12.8463	19.9082	20.8987
MAY	6.6476	11.1751	13.0864
JUN	8.6849	10.3228	12.3955
JUL	11.7348	13.8004	16.9903
AUG	15.2764	17.5669	21.2608
SEP	17.3334	22.3738	24.7487
OCT	25.4309	22.3738	30.3356
NOV	12.0536	17.6134	19.9894
DEC	35.5735	37.7093	42.5959

5. CONCLUSION

Results indicated that the Box-Jenkins ARIMA is the best method in forecasting the SOI values compared to Single Exponential Smoothing and Double Exponential Smoothing. In future, it is recommended that forecasting models such as Generalized Autoregressive Conditional Heteroscedasticity (GARCH), time series regression, and Generalized Linear Model (GLM)can be tested to find the best forecast method of SOI values. It is suggested that SOI values can be analyzed according to daily, weekly or yearly in order to get a clear pattern and trend of data. In addition, a longer data set is recommended to increase the accuracy of the forecasted values.

REFERENCES

[1] Howard,B.C. (2015).How El Nino Affect Weather. National Geographic.
[2] Lal, A., Ikedaz, T., French, N., Baker, M.G., Hales, S. (2013).Climate Variability, Weather and Enteric Disease Incidence in New Zealand: Time Series Analysis. PLOS ONE, Vol 8(12): e83484.
[3] Velayudhan. (2016).El Nino may increase breeding grounds for mosquitoes spreading Zika virus, World Health Organization, Humanitarian Health Action, Geneva.
[4] Kovats, R.S. (2000). El Nino and Human Health. Bulletin of the World Health Organization, 78 (9).
[5] Trenberth, K.E., Stepaniak, D.P. (2000).Indices of El Nino Evolution. Journal of Climate, Vol 14, 1697-1701.