

Forecasting the Amount of the Lung Diseases by the Method of ARIMA-ARCH

B K Mbau

Magister of Information System, Faculty of Postgraduate, Universitas Komputer Indonesia, 112-116 Dipatiukur Street, Lebak Gede, Coblong, Bandung, West Java, 40132, Indonesia

*billysid.mbau2@hotmail.com

Abstract. This paper aimed at forecasting the number of cases of Tuberculosis in XYZ hospital. The lungs are one of the vital organs in the human body. This research method using ARIMA-ARCH model. Test results into the influential heteroskedastisitas forecasting the number of cases of tuberculosis, where when the model initially has no effect ARCH but still continued into the ARIMAARCH value produces a MAPE of better evidenced by the model of ARIMA (0, 1, 3) the ARCH (1) have a better accuracy level, where has a value of MAPE 2.09 % whereas the model ARIMA (0, 1, 3) has a value of MAPE 2.31%.

1. Introduction

Forecasting is an activity to predict what will happen in the future [1]. Forecasting is a vital part of every business organization and management decision making which is very significant. Forecasting of time series or a Time Series Forecasting is a series of observations of a variable taken from time to time and are recorded in chronological order according to the time sequence [2,3].

In the Time Series forecasting technique, it is not rarely the data shows a trend pattern, where the pattern of data shows a tendency to increase or decrease. Holt-Winter's Exponential Smoothing method is a forecasting method with an exponential smoothing approach based on forecasting results in the previous period. This method also adds parameters to handle seasonal data patterns. There are two main models in Holt-Winter's Exponential Smoothing method, namely multiplicative model and additive model. The determination of this model was chosen based on the seasonal pattern [4, 5].

Forecasting is an important tool in planning effective and efficient [6]. According to Makridakis, forecasting techniques are divided into two parts, the first method of forecasting the subjective and objective forecasting methods. Forecasting method of subjective qualitative models and methods have a forecasting model has two objectives, namely time series models and causal models [7,8]. Qualitative model strives to incorporate subjective factors in forecasting model, this model would be very helpful if an accurate quantitative data difficult to obtain [9-11]. In previous research, doing research with honda's motorcycle sales method holt winters additive with the amount of data that many produce high accuracy for a period of long-term forecasting. aimed at forecast the number of cases of Tuberculosis in XYZ hospital using ARIMA-ARCH model.



2. Experimental method

This research uses ARIMA-ARCH model in forecasting amount of lung disease.

2.1. ARIMA Model

ARIMA forecasting method is to complete a series of intervals to analyze coherently time.

Generally, ARIMA notation is:

ARIMA (p, d, q) (P, D, Q)

With:

p, d, q : The non-seasonal part of the model

(P, D, Q) s : The seasonal part of the model

The general formula of ARIMA (p, d, q) (P, D, Q) is as follows:

$$\Phi_p B^s \phi_p(B)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\Theta_q(B^S)a_t \quad (1)$$

With :

$\phi_p B$: AR Non Seasonal

$\Phi_p B^S$: AR Seasonal

$(1-B)^d$: differencing non seasonal

$(1-B^S)^D$: differencing seasonal

$\theta_q(B)$: MA non seasonal

$\Theta_q(B^S)$: MA seasonal

2.2. Autoregressive Conditional Heteroscedasticity (ARCH) method

Heteroskedastisitas occurs because of data time series showing the diversity of the pattern element generated from the rest of the previous variants.

Generally, ARCH notation is:

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \dots + \alpha_p e_{t-p}^2 \quad (2)$$

2.3. Measures of accuracy

It is necessary select a particular measure of accuracy in order to examine the accuracy of the forecasting method. Some of the measuring method which generally used are Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE).

2.3.1. Mean Absolute Percentage Error (MAPE). MAPE is the average value of the absolute value of the difference between the results of the forecasting and an actual value, the value is shown in percentage

$$MAPE = \sum_{t=1}^T \frac{|forecast - actual|}{actual} * 100 \quad (3)$$

3. Results and Discussion

3.1. Model ARIMA

3.1.1. Model identification. The process of identifying the data model is done by looking at the plot of the actual data.

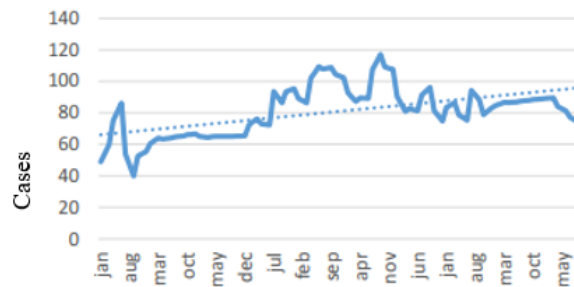


Figure 1. Graph monthly data the number of tuberculosis cases.

3.1.2. *Stationary Test Range.* Stationary test performed on a variety of training data the number of cases of tuberculosis.

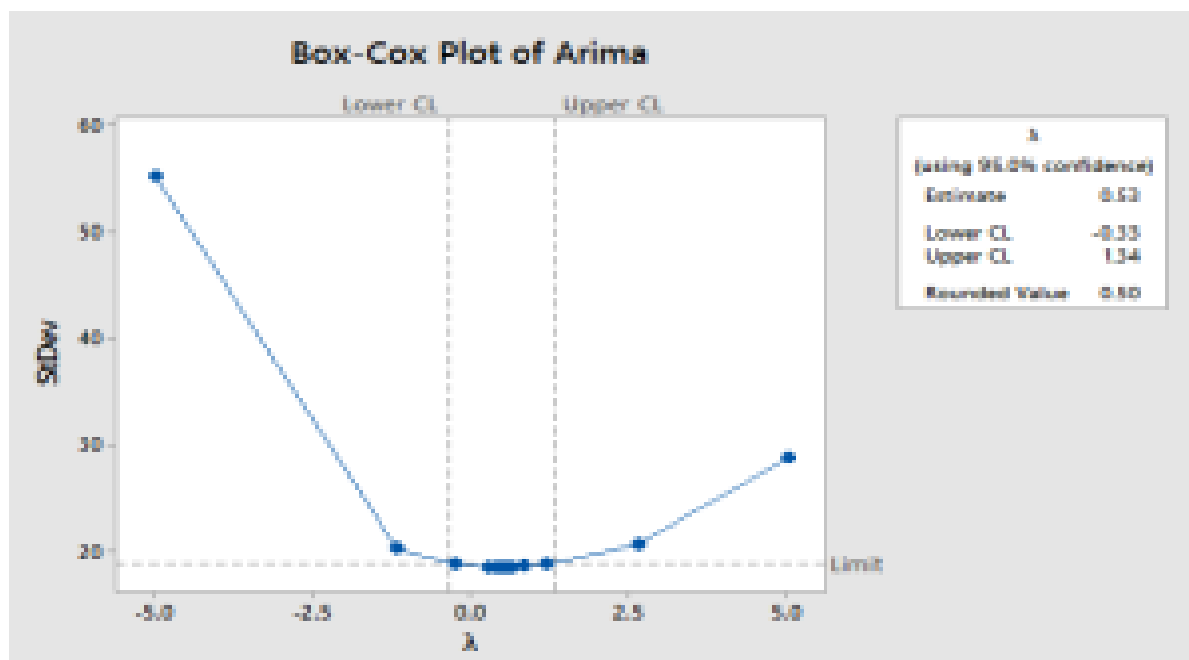


Figure 2. The stationary Test range

3.1.3. *ARIMA-ARCH model.* Determination of the best models of ARIMA ARCH can be seen from the value of the AIC and the SIC, where models have a value of AIC and SIC is the smallest model.

Table 1. Comparison model AIC and SIC ARCH model.

Model	AIC	SIC
ARIMA (0,1,3) ARCH (1)	-2.043650	-1.954967
ARIMA (2,2,0) ARCH (1)	-1.682676	-1.593350

4. Conclusions

Test results into the influential heteroskedastisitas forecasting the number of cases of tuberculosis, where when the model initially has no effect ARCH but still continued into the ARIMA-ARCH value produces a MAPE of better evidenced by the model of ARIMA (0, 1, 3) the ARCH (1) have a better accuracy

level, where has a value of MAPE 2.09 % whereas the model ARIMA (0, 1, 3) has a value of MAPE 2.31%. However, when the model of ARIMA has effect ARCH that means the model is not feasible then continued into the ARCH model produces a MAPE of less value nice compared to the initial model demonstrated by models of ARIMA (2, 2, 0) of the ARCH (1) have the MAPE 9.29 %. Whereas models ARIMA (2, 2,0) had MAPE 5.69%. But if viewed from the ability of the model in following the pattern of the actual data in the get results that model with ARIMA-ARCH shows better results in terms of following the pattern of the data actual. Therefore models of ARIMA-ARCH can be used to predict cases tuberculosis at XYZ hospital, when the models of ARIMA obtained is not feasible, a model is said to be not worth it when there diagnostic test passes as well as models that don't pass the test heteroskedastisitas. Models of ARIMA-ARCH is also used as a a comparison of models of ARIMA to determine how big an influence effect against ARCH model ARIMA, either when in the test heteroskedastisitas ARIMA models have effect ARCH or not. The results of the forecasting of the number of cases of tuberculosis in XYZ hospital by using model ARIMA-ARCH generates value MAPE 1.88% that indicates that the model has the rates great accuracy because the value of MAPE < 10%.

Acknowledgements

Praise and gratitude I pray to the omnipotent God and I am very grateful to UNIKOM and dones-lecturers who have guided me in making this journal then I say thank you also to the Balai Besar Kesehatan Paru Masyarakat Bandung which has given me the opportunity to conduct research in there.

References

- [1] Cain K P, Oeltmann J E, Kammerer J S, Moonan P K and Ricks P M 2011 Estimating the burden of tuberculosis among foreign-born persons acquired prior to entering the u.s., 2005-2009 *PLoS One*
- [2] Yan-Ling, L P Zhang, Xue-Liang Z, Kai W and Yu-Jian Z 2014 Forecast Model Analysis for the Morbidity of Tuberculosis in Xinjiang, China *Plos One*
- [3] Ajmer D 2018 An Alternative Approach of Handling the Outlier: *A Case with Indian NPS. Journal of Data Science* **18** pp.29-50
- [4] N Dube 2015 Application and Comparison of Time Series Methods on Tuberculosis Incidence Data: A case study of Zimbabwe 1990-2013
- [5] WHO 2014 *World Health Organization* Who,global tuberculosis report 2014
- [6] C Dye 2006 Global epidemiology of tuberculosis *The Lancet*
- [7] Engle R F Autoregressive conditional heteroscedasticity with estimates of vari-ance of united kingdom ination *Econometrica*.
- [8] C A Martin and S F Witt 1989 Forecasting tourism demand: A comparison of the accuracy of several quantitative methods *Int. J. Forecast.* **5** 1 7–19
- [9] S Thomassey 2014 Intelligent Fashion Forecasting Systems: Models and Applications 7(8) pp. 9–28
- [10] Siriwan W, Mullica J and Krisanadej J 2012 Development of temporal modeling for prediction of dengue infection in Northeastern Thailand. *Asian Pacific Journal of Tropical Medicine*
- [11] Bestriandita D , Yuli P and Hety U 2016 Peramalan Data Time Series Angka Penjualan Sepedamotor Honda Dengan Metode Holt – Winter Additive