ISSN: 0975-766X
CODEN: IJPTFI
Available Online through *Research Article*
www.ijptonline.com

OPTIMUM MODEL OF FORECASTING DIARRHEA PATIENTS IN KOTA GORONTALO

Herlina Jusuf, Rani Hiola
State University of Gorontalo, 6, Jl. Jenderal Sudirman, Kota Gorontalo, Gorontalo, 96128 Indonesia.

Received on 15-10-2016 *Accepted on 18-11-2016*

Abstract: This study compares three forecasting model based on seasonal and trend pattern, which are Winter's Smoothing Method, Trigonometry Regression Method, and Arima Method. Those three methods will be tested to find which one is the best in order to forecast the total people suffering from Diarrhea in Kota Gorontalo. This research uses data from January 2010 until December 2015 from Health Service in Kota Gorontalo.

Keywords : Forecasting, Seasonal, ARIMA, Winter's, Trigonometry.

1. Introduction

Diarrhea patient in Kota Gorontalo increase every year in rainy season and become major problem of death rate especially infant rate in this city. Consequently, optimum model to forecast the patient number are considerable importance. To find the optimum model we compares three forecasting methods which are Winter's Smoothing Method, Trigonometry Regression Method, and Arima Method. The optimum model can be use by the stakeholders to make projections about preventive measure of diarrohea in the future.

Forecasting is an important aid in effective and eficient planning [9]. The effectivity of a desicion are based on some factors that won't be seen in the time when the decision are taken. Therefore, whether the forecasting method is used or not, the uncertainty can not be throughly eliminated [4]. To minimize minimize the uncertainty rate the optimum forecasting methods became urgently need [1,10]

Forecasting models are used in many institute such as health institute still using classic graph forecasting model that are based of element feeling so it lead to an error forecasting and wrong decision-making. To choose the exact time series method, we need to consider the various data pattern. On the decision making the, stakeholders usually used seasonal data model and trend [9]. For that kind of model there are some methods that can be used which are smoothing method, trigonometry regression method, and ARIMA method.

The basic of smoothing method is reviewing historical data over time to predict the future behaviour. Forecasting withtrigonometry time series analysis are based of trigonometric function over period of time. While ARIMA model are used for seasonal data pattern not for seasonal pattern. Those three methods will be tested to find which one is the best in order to forecast the total people suffering from Diarrhea in Kota Gorontalo

Time series data is set of evenly spaced numerical data, from monthly or yearly. Where *t* to *t+1* until *t+n*[9].

Seasonality is a periodic and recurrent pattern [4]. A simple way to decide the seasonal model data is to split the period of time to three part which are Trend (T_t), seasonal element (S_t), and error (ε_t) [3]. Then we called the distribution of $Z_t = T_t + S_t + \varepsilon_t$[2]

1.1 Forecasting Model using Smoothing Method

Smoothing methods is known as Winter Methods with three smoothing equation which are:

Trend Smoothing : $b_t = \gamma (S_t - S_{t-1}) + (1 - \gamma) b_{t-1}$

Seasonal Smoothing : $I_t = \beta \frac{X_t}{S_t} + (1 - \gamma) I_{t-l}$

Overall Smoothing : $S_t = \alpha \frac{X_t}{I_{t-L}} + (1 - \alpha)(S_{t-1} + b_{t-1})$

And forecasting with smoothing method is defined as: $F_{t+n} = (S_t + b_t m) I_{t-L+m}$

Where L = seasonal lenght,

b = *trend* component,

I = Seasonal compability,

F_{t+n} =forecasting for future m period,

α, β, γ = smoothing constant.

1.2 Forecasting Model using Trigonometric

$S_t = \sum_{i=1}^{n} A_t sin \left(\frac{2\pi i}{S} t + \emptyset_t\right)$.[7], a model of seasonal combination $S_t$and trendT_t is:

$Z_t = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{t^i}{i} + \sum_{i=1}^{n} A_i sin(f_i t - \emptyset_i) + \varepsilon_i$[3]

$= f'(t)\, \beta + \varepsilon_i$

A trigonometric regression model to explain patterns and seasonal trends are:

$Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2f\pi t) + \beta_3 \sin(2f\pi t) + \varepsilon_t$, [6]

t = time period (months) → 1,2,3,.....,72

f = frequency → (in this case more 1/12 shows) the length of seasonal

$\beta_0, \beta_1, \beta_2, \beta_3$ = the regression parameter

ε_t = the forecast error in period t.

1.3 Forecasting Model using ARIMA

$\emptyset_P(B)\emptyset_P(B^S)\nabla^d\nabla_S^D Z_t = \theta_q(B^i)\theta_Q(B^S)_{at}$[8,9]

The seasonal model of ARIMA is (p,d,q) $(P,D,Q)^S$

1.4 Measurement of the Forecast Error

Mean Absolute Deviation : MAD = $\frac{\sum_{t=1}^{n}\|\gamma_t - \hat{\gamma}_t\|}{n}$

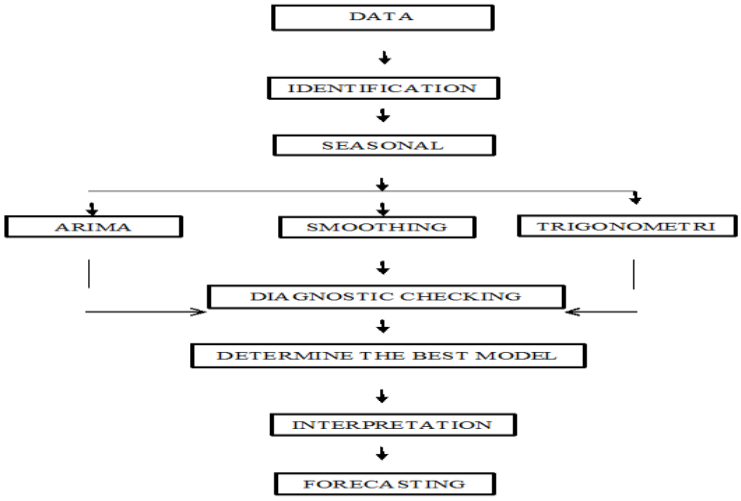
Mean Standard Error (MSE) = $\frac{\sum(\gamma_t - \hat{\gamma}_t)^Z}{n}$

Mean Absolute Percentage Error (MPE) = $\frac{\sum\frac{(\gamma_t - \hat{\gamma}_t)}{\gamma_t}}{n}$

2. Research Methods

This study is an applied research using secondary data to compare three smoothing methods, such as Winter's method, Trigonometric Regression method and ARIMA method, to get which one is the best forecasting model.We used a monthly data from the City Health Office Gorontalo, which is patients who have diarrhea from January of 2010 until the month of December 2015. To see the pattern of the existing data, we then made a plot of the data, which is supported by the calculation of autocorrelation r_k that describes the relationship between the time series with itself at the slowness of time (lag) k periods and the coefficient of autocorrelation partially that used to assist in establishing fit model [11]. We used minitab, one of statistical programming, to estimate the parameter iteratively.

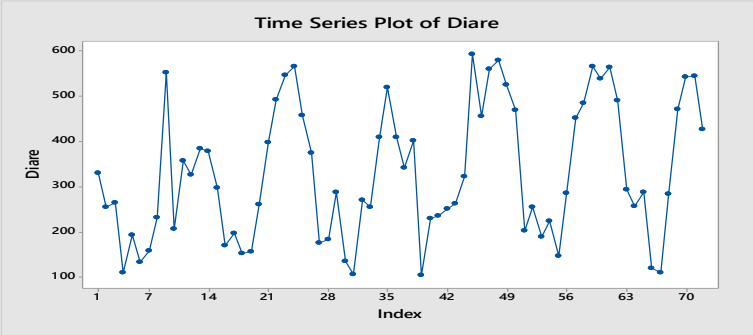
The flowchart 1 of the study showed in this picture below..



3. Result

3.1 Data pattern Identification

The first step in forecasting model is to understand the pattern of that data, so we plot the diarrhea data to get the identification of the data pattern. Figure 2 provides the results of Diarrhea's time series plot:



From figure 2 we can see that the pattern data of the diarrhea patients in the city of Gorantalo showed that there are trend and seasonal. This trend indicates that there is an increasing trend in Gorontalo diarrhea patients in each month so that it can be used to predict the data with diarrhea.

3.2Applied for Forecasting Method

3.2.1. Winter's Method

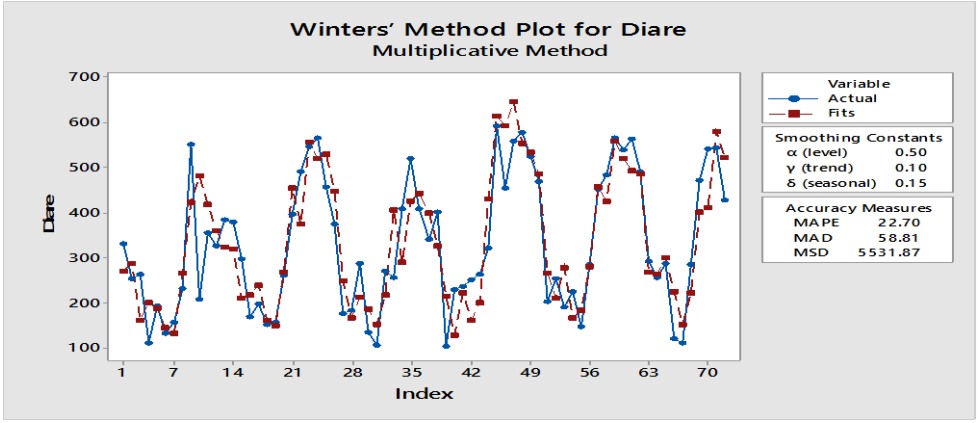
Smoothing method can be used to forecast data series that have a seasonal pattern and the trend. This method is a multiplicative method winters with the parameter of Alpha (level) is 12:50, while Gamma (trend) is 0:10 and Delta (seasonal)is 12:15. Forecasting Model with Winter's method using alpha (level) = 0.50, Gamma (trend) = 0.10, and Delta (seasonal) = 0.15 is:

Exponential Smoothing : $S_t = 0{,}50(X_t/I_{t-12}) + 0{,}50(S_{t-1} + b_{t-1})$

Trend Smoothing : $b_t = 0{,}10\ (S_t - S_{t-1}) + 0{,}90\ b_{t-1}$

Seasonal Smoothing : $I_t = 0{,}15\ (X_t/S_t) + 0{,}85\ I_{t-L}$

Plot of the original data and the results of forecasting using Winter's can be seen in figure 3.



3.2.2. Trigonometric Method

The pattern of trend and seasonal in forecasting model using trigonometric regression model expressed in the form below.

$Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2f\pi t) + \beta_3 \sin(2f\pi t) + \varepsilon_t$

Therefore the forecasting model for the number of diarrhea patient in Gorontalo expressed in this form below.

$Y_t = 276 + 1.50t - 51.9\ X_1 + 161\ X_2$

where: $X_1 = \text{Sin}\ (\pi t/12)$

$X_2 = \text{Cos}\ (\pi t/12)$

S = 76.88 R-Sq = 73.90 R- Sq (adj) = 72.70

3.2.3. ARIMA

The procedures of Box-Jenkins [5] to create the ARIMA model are identification, parameter estimation, diagnosic checking, and forecasting. From the result of the best ARIMA model for forecasting in the number of diarrhea patientbased on the lower MSE is ARIMA $(0,1,1)(0,1,1)^{12}$. Therefore that model is a best model for forecasting the number of diarrhea patient in Gorontalo. The forecasting result expressed in this form below.

$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0{,}8173e_{t-1} - 0{,}8513e_{t-12} + 0{,}6958e_{t-13}$

4. Disscussion

Increasing the number of diarrhea patient that always happen in almost the same time made the data have a seasonal and trend pattern. The seasonal pattern describing the seasons change from dry to wet season or otherwise, so thisaffect the incidence of diarrhea Gorontalo.

4.1 Model Identification using Smoothing Method

Implementation winters exponential smoothing method for forecasting the number of diarrhea patients we used Alpha (level): 0.50, Gamma (trend): 0.10, Delta (musiman) : 0.15. To determine whether the Winter's method is good method to forecast the number ofdiarrhea patients, we need to check whether the forecasting error already meet the IIDN assumption$(0,\sigma^2)$.

Therefore we foundthat the forecasting error for Winter's methodhas met the assumption.

By cumulating the forecastingof the data, we foundthe estimation for forecasting the number of diarrhea patient in Gorontalo are MAPE : 22.70, MAD : 58.81 , MSD/MSE : 5531,87. One of the problem when we used Winter's Method is to decide the parameter. One of the problems when we usedWinter's method is how to determine the

parameter values that will minimize the MSE/MSD. Usually approaches todetermine this valueis by trial and error. But because such approaches will take a long time then this method is rarely used.

4.2 Model Identification using Trigonometric Method

Trigonometric regression model is in accordance with the pattern of existing data and could explain 73.9% of the variation number of diarrhea patients. This means that the provisions of this regression model in predicting the number of diarrhea patients in Gorontalo amounted to 73.9% for error 5911. Test of conformity and significance of the model parameters shows that simultaneously and partially, the model is appropriate. All parameters are significant at alpha 1%, which means that the trend and seasonal patterns are appropriate. Then we need to check the assumption for error of the model to determine whether the IIDN assumptions $(0, \sigma^2)$ has been fulfilled. Results showed that the model has an error which meets the assumptions and it's indeed good to forecast the number of diarrhea patients Gorontalo.

4.3 Model Identification using ARIMA

Box-Jenkins steps to create the ARIMA model areidentification, parameter estimation, and diagnostic checking andforecasting.

4.3.1 Identification steps

The procedures of these steps are checking the pattern of the data, ACF and PACF to know whether the data is stationer or not. Also using these procedureswe would like to get the first model. If the model show the data is not stationer, then we need to perform differencing in the data. The next step is to check the pattern ofACF and PACF to get the appropriate decision relate to this stationery of the data. Using a computer program, we got the pattern of ACF and PACF.

The ACF showed that the data is not stationer in its seasonal mean. It showed that the lag 12.24 and 36 are higher and not going down exponentially. Based on this result, our first step is to perform a seasonal differencing (D=1). The result of this step showed that the data is still not stationer in non seasonal mean, so the next step is to perform a differencing in non seasonal from the data that we got after the seasonal differencing.The result of the ACF and PACF pattern showed that the data has been stationary, so we can used the first model ofthe ARIMA. Based on the pattern, ACF were significant only in lag 12 and PACF relatively down exponentially, so we got the first model for ARIMA is ARIMA $(1,1,0)(0,1,1)^{12}$.Furthermore, we need to perform an estimation parameter to chechk the first model of ARIMA and do an evaluation to choose the best model.

4.3.2 Estimation parameter and diagnostic checking steps

Estimated parameters of the first model provides an estimated parameter are all significant at α=5% and the error is white noise. To check whether the parameter is significant or not, we conducted a statistical t-test, while to check whether the error is white noise or not we conducted a statisticalLjung-Box Chisuarestest. The next step is to determine the best model between the two models before, that will be used for forecasting. The criteria used to measure the goodness of a model is the mean of square error (MSE). MSE is a selection tool models based on error estimation results. Error showss how much difference the estimation results with values to be estimated. This value will be used to determine which is the best model[12]. The comparison showed that ARIMA $(0,1,1)(0,1,1)^{12}$ gives the smaller MSE, which means that this model produces an average error is smaller than the ARIMA$(1,1,0)(0,1,1)^{12}$. We then used this ARIMA $(0,1,1)(0,1,1)^{12}$model to get the forecast number of diarrheapatients in Gorontalo in the future.

5 Forecasting Steps

In this step we perform a forecasting using the best model, ARIMA $(0,1,1)(0,1,1)^{12}$, so for estimate with 95% interval we got

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0{,}8173e_{t-1} - 0{,}8513e_{t-12} + 0{,}6958e_{t-1}$$

6 Choosing The Best Method

Comparison of three forecasting methods to determine which method is the best method done by comparing the error criteria of forecasting methods, such as MSD/MSE, MAD and MAPE, where the best method is a method that having smaller value for those three chriteria if comparing to the other methods.

Table 1. Result of comparing value for MSD/MSE, MAD and MAPE.

Error criteria of forecasting	Forecasting Method		
	Winter's	Trigonometric Method	ARIMA $(0,1,1)(0,1,1)^{12}$
MSD/MSE	5531.87	5582	4421.00
MAD	58.81	58.86	51.27
MAPE	22.70	23.41	20.88

Based on table 1, we concluded that the ARIMA $(0,1,1)(0,1,1)^{12}$ is the best method to describe and predict patterns or fluctuations in the number of diarrhea patients. These results provide a fact that the more complex methods, in this case the Box-Jenkins ARIMA have identification, estimation and forecasting steps, it tends to give better results.

The forecasting forthe number of diarrhea patients for 2016 in Gorontalo can be seen in table 2.

Tabel 2. Forecasting for the number of diarrhea patientsin Gorontalofor 2016.

No	Year_Month	Model		
		Winter's	Trigono	ARIMA
1	2016 January	436	499	476
2	2016February	389	423	435
3	2016 March	210	337	240
4	2016 April	192	265	236
5	2016 May	217	227	248
6	2016 June	155	232	203
7	2016 July	141	281	178
8	2016 Agustus	255	361	302
9	2016 September	407	449	476
10	2016 Oktober	404	525	481
11	2016 November	455	566	550
12	2016 December	413	563	520

7. Conclusion

From those three methods used to forecast the number of diarrhea patients with diarrhea in Gorontalo, we found that the ARIMA $(0,1,1)(0,1,1)^{12}$ is the best method due to itssmallest value, which is 4576 compared with the error value of other methodssuch as Winter's method andTrigonometric method,which are 5911 and 5121.82.

ABOUT THE AUTHORS

First Author Herlina Jusuf, Dr.-M.Kes., a lecturer and researcher in the Department of Public Health, State University of Gorontalo. The author's major is biostatistics.

Second Author Rani Hiola, M.Kes., a lecturer and researcher in the Department of Public Health, State University of Gorontalo. The author's major is biostatistics

Acknowledgements

This work was supported by Department of Health, Gorontalo, Indonesia.

References

1. Arsyad, L. *Peramalan Bisnis.* BPFE (2001). [in Yogyakarta]
2. Assauri, S. *Tehnik dan Metode Peramalan.* FE : Unair (1990)[in Surabaya].

3. Bovas, A and Johanes Ledalter. *Statistical Methods For Forcasting.* John Wiley &Sons, Inc Publicafion(2005) [in New Jersey].
4. Bowerman, B and Richard O'Connell. *Forcasting Time Series and Regression AnApplied Approach 4 th Edition*. Thomson Brooks/Cole(2005)[in California].
5. Box, G.E.P., Jenkins, G.M., and Reinsel. G.C., *Time Series Analysis Forcastingand Control*, 3d edition, Englewood Cliffs : Prentice Hall(1994).
6. Cryer, J.D. *Time Series Analysis With Application In R, Second Edition.* Springer Science & Busines Media(2008)
7. Djauhari, M. *Metode Peramalan.* Karunika UT(1986) [in Jakarta]
8. Koehler, A.B.*Forecasting Model's and Prediction Intervals For the Multivariate, Holt Winter's Method International.* (2001)Journal Of *Forecasting 17, 269- 286.*
9. Makridakis, Whellrigt, dan McGee.*Metode dan Aplikasi Peramalan edisi revisi.* Bina Rupa Aksara Publiser(2001) [in Tangerang].
10. Mulyono, S. *Peramalan Bisnis Dan Ekonometrika.* BPFE (1999) [in Yogyakarta]
11. Wei, W.S William. *Time series Analysis* Univariate and Multivariate Methods 2nd Edition Pearson Addison Wesley Publishing Company (2006) [inCalifornia].
12. Wahba, G. *"Spline Models for Observational Data, Society for Industrial and Applied Mathematics".*(1990)[in Pennsylvania].