IOP Conf. Series: Materials Science and Engineering **166** (2017) 012034 doi:10.1088/1757-899X/166/1/012034

Application of Holt exponential smoothing and ARIMA method for data population in West Java

A. Supriatna, D. Susanti, E. Hertini

Department of Mathematics, Faculty of Mathematics and Natural Science, Padjadjaran University, Bandung, West Java

E-mail: Asupriatna55@ymail.com

Abstract. One method of time series that is often used to predict data that contains trend is Holt. Holt method using different parameters used in the original data which aims to smooth the trend value. In addition to Holt, ARIMA method can be used on a wide variety of data including data pattern containing a pattern trend. Data actual of population from 1998-2015 contains the trends so can be solved by Holt and ARIMA method to obtain the prediction value of some periods. The best method is measured by looking at the smallest MAPE and MAE error. The result using Holt method is 47.205.749 populations in 2016, 47.535.324 populations in 2017, and 48.041.672 populations in 2018, with MAPE error is 0,469744 and MAE error is 189.731. While the result using ARIMA method is 46.964.682 populations in 2016, 47.342.189 in 2017, and 47.899.696 in 2018, with MAPE error is 0,4380 and MAE is 176.626.

1. Introduction

The population has a data pattern trend up because the population has increased. Increased happens will have an impact on many things, including the increasing need for food, house, or jobs. In addressing improvements, required a prediction of the number of people in order to define strategies. Prediction of total population can found using Holt and Autoregressive Integrated Moving Average (ARIMA) method. Holt is an exponential smoothing method by performing weighting on past observations and estimate trend in data. While ARIMA method is studied in depth by George Box and Gwilym Jenkins and consists of four stages: identification, estimation, diagnostic checking, and forecasting [4].

2. Methods and Materials

The data used is the population of West Java for 1998-2015 periods (Table 1, Figure 1).

2.1. Holt Exponential Smoothing

Holt is a double exponential smoothing method has two parameters. In the weighting parameters, the new value has a greater than previous observations. In addition, there is weighted to estimate the trend of data. Generally, Holt has two smoothing constants (the values between 0 and 1) and three equations [5]

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

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

Forecast : $F_{t+m} = S_t + b_t(m)$

Table 1. Data population for 1998-2015 [1]

Period	Data	Period	Data
1998	33.261.409	2008	42.194.809
1999	34.555.622	2009	42.693.951
2000	35.723.473	2010	43.053.732
2001	36.795.565	2011	43.826.775
2002	37.291.946	2012	44.548.431
2003	37.980.422	2013	45.340.799
2004	39.140.812	2014	46.029.668
2005	39.960.869	2015	46.497.175
2006	40.737.594		
2007	41.483.729		

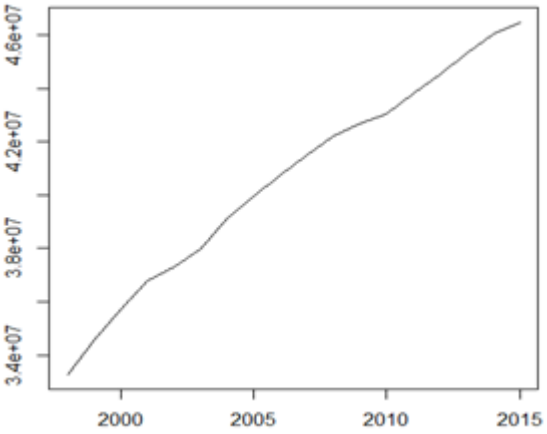

Figure 1. Plot of data population

2.2. Autoregressive Integrated Moving Average (ARIMA)

ARIMA is the short-term forecasting [2]. According Makridakis et al (1999), the stages of ARIMA are Identification, parameter estimation, diagnostic checking, and forecast. ARIMA models are generally expressed with (p,d,q) .

ARIMA Model with $(p,2,q)$ [4]:

$$(1-\mathrm{B})^{d} X_{t} = e_{t}$$

e_t : error,

X_t : Observation value for t ($t = 1, 2, ..., n$),

$(1-\mathrm{B})^{d}$: mathematics operation for differencing.

3. Results

3.1. Holt Exponential Smoothing

Holt method generally has two smoothing constants (with value between 0 and 1). To obtain the optimal parameters used by trial and error. Table 2 is the process of finding the value of α and γ. The value will selected with smallest MAPE error. Then the parameters to be used are $\alpha = 0.9$ $\gamma = 0.9$.

Table 2. Selection Process for parameter α and γ

No	α	γ	MAPE
1	0,1	0,1	3,889468
2	0,1	0,5	2,131118
3	0,1	0,9	1,425721
4	0,2	0,1	2,556987
5	0,2	0,5	1,104799
6	0,2	0,9	0,714864
7	0,3	0,1	1,836226
8	0,3	0,5	0,74986
9	0,3	0,9	0,56141
10	0,4	0,1	1,41355
11	0,4	0,5	0,598116
12	0,4	0,9	0,591477
13	0,5	0,1	1,144885
14	0,5	0,5	0,51336
15	0,5	0,9	0,619235
16	0,6	0,1	0,961856
17	0,6	0,5	0,523453
18	0,6	0,9	0,592657
19	0,7	0,1	0,829935
20	0,7	0,5	0,520288
21	0,7	0,9	0,547798
22	0,8	0,1	0,730374
23	0,8	0,5	0,496619
24	0,8	0,9	0,496724
25	0,9	0,1	0,667488
26	0,9	0,5	0,484451
27	0,9	0,9	0,469744

Based on Holt method, with alpha 0.9 and gamma 0.9 is obtained prediction of the total population in 2016 is 47.028.975, populations in 2017 is 47.535.324, and populations in 2018 is 48.041.672. The value of MAPE error is 0,469744 and MAE error is 189.731.

3.2. Autoregressive Integrated Moving Average (ARIMA)

Time series analysis assumptions that must be fulfilled is stationary data. Based on the plot seen data contains trend patterns and not stationary in mean. In addition to viewing the plot, not stationary can viewable by ACF plot. If the ACF down slowly, then data is not stationary [3]. First, data must differencing because ACF plot down slowly (Figure 2). The result since first differencing (Figure 3), data still not stationary. Second differencing done to create data stationary (Figure 4).

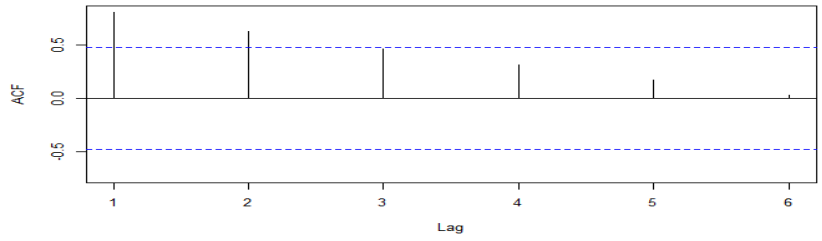

Figure 2. Plot ACF of data populations

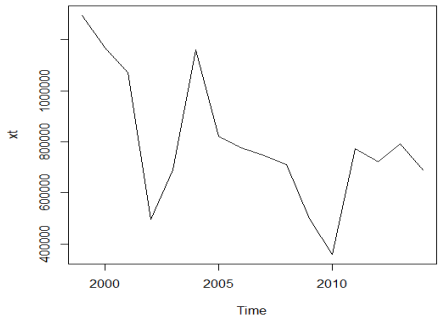

Figure 3. Result of first differencing

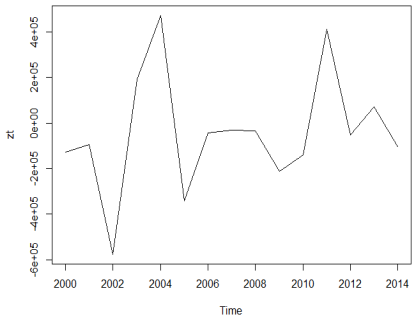

Figure 4. Result of second differencing

Since data is stationary, plot ACF and PACF used to choose the model of ARIMA(p, d, q). ACF will determine the orde of Moving Average (q), order (d) is order differencing, and PACF will determine the orde of Autoregressive (p).

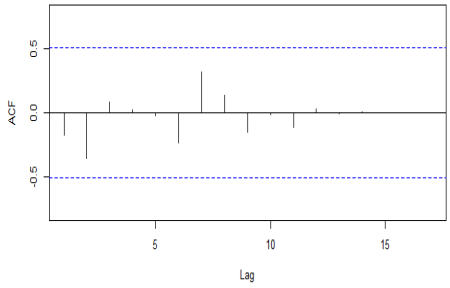

Figure 5. ACF after differencing

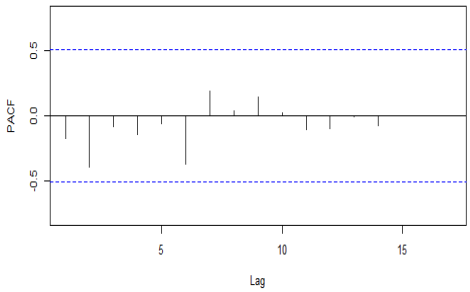

Figure 6. PACF after differencing

Based on Figure 5 and Figure 6, ACF and PACF plot show no lag was out from Bartlett so that is probably is ARIMA $(0,2,0)$. Furthermore, the suitable model used for forecast. The result of prediction are 46.964.682 populations in 2016, 47.432.189 populations in 2017, and 47.899.696 populations in 2018 with MAPE error is 0,4380 and MAE is 176.626.

4. Conclusions

The results of Holt method has prediction in 2016 is 47.028.975 populations, 2017 is 47.535.324 populations, and in 2018 is 48,041,672 population, with a MAPE error is 0.469744 and MAE is 189.731. While ARIMA method predictions in 2016 is 46.964.682 populations, 2017 is 47.432.189 populations, and in 2018 is 47.899.696 populations, with MAPE error is 0.4380 and MAE is 176,626. ARIMA method forecasting is better than Holt method because error values is smaller.

References

[1] Badan Pusat Statistik Provinsi Jawa Barat 2016 *Statistik Jumlah Penduduk Jawa Barat*.
[2] Ekananda, M 2014 *Analisis Time Series: Untuk Penelitian Ekonomi, Manajemen, dan Akutansi*. (Jakarta: PT. Mitra Wacana Media)
[3] Hyndman, J.R 2014 *Forecasting: Principles & Practice*. (Australia)
[4] Makridakis, S, et al 1999 *Metode dan Aplikasi Peramalan, Jilid 1.* (Jakarta: Binapura Aksara)
[5] Noeryanti, et al 2012 *Aplikasi Pemulusan Eksponensial Dari Brown dan Holt Untuk Data yang Memuat Trend*. Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III. Yogyakarta