PAPER • **OPEN ACCESS**

Seasonal Time Series Forecasting using SARIMA and Holt Winter's Exponential Smoothing

To cite this article: G A N Pongdatu and Y H Putra 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* **407** 012153

IOP Conf. Series: Materials Science and Engineering **407** (2018) 012153 doi:10.1088/1757-899X/407/1/012153

Seasonal Time Series Forecasting using SARIMA and Holt Winter's Exponential Smoothing

G A N Pongdatu* and Y H Putra

Magister of Information System, Faculty of Postgraduate, Universitas Komputer Indonesia, 112-116 Jalan Dipatiukur, Lebak Gede, Coblong, Bandung, Jawa Barat, 40132, Indonesia

*dionpongdatu@gmail.com

Abstract. The purpose of this study is to compare SARIMA and Holt-Winter's Exponential Smoothing methods in an attempt to generate customer transaction forecasting in Store X with high accuracy.This study will compare the results of sales forecasting with time series forecasting model of Seasonal Autoregressive Integrated Moving Average (SARIMA) and Holt Winter's Exponential Smoothing method. SARIMA model still accurate when the time series data is only in a short period, this model is accurate on short period foracasting but less accurate on long period forecasting. Meanwhile Holt Winter's Exponential Smoothing accurate on forecasting seasonal time series data, either it's pattern shows trend or not. Both models are compared with forecasting data showing seasonal patterns. The data used is the data of clothing retail store sales from 2013 to 2017. Accuracy level of each model is measured by comparing the percentage of forecasting value with the actual value. This value is called Mean Absolute Deviation (MAD). Based on the comparison result, the best model with the smallest MAD value is SARIMA model (1,1,0) $(0,1,0)^{12}$ with MAD value 5.592. From the comparison results can be concluded that the SARIMA model is feasible to be used as a model for further forecasting.

1. Introduction

In the apparel business it is generally known that consumer demands are very unstable. In fact consumer's choice on fashion generally based on price. In this case, to overcome this condition, the managerial store tried to reduce the price by reducing the cost of production or by buying goods directly from the manufacturer or the first party. Almost all manufacturers previously mentioned, located on a remote island or very far from store X. Thus the process of supply of goods must be done with the appropriate strategy and timing. It is intended that the goods are not stacked stockpiled in the warehouse, and remain sustainable in supply chain, storage and sales.

Forecasting customer transaction patterns are done so that the store can determine the strategy in choosing collection as well as the appropriate timing to prevent the accumulation of unsold item. Time series forecasting methods are probably the most frequently used techniques for prediction of sales data. Included in these statistical technique, various well-known models that have formal statistical foundations [1]: exponential smoothing [2], Holt Winters model [3], Box & Jenkins model [4], regression models [5] or ARIMA. These methods have been implemented in different areas and they provide satisfactory results [6].

Nevertheless, research on forecasting has not yet compared the seasonal time series model to forecast sales transaction. Based on previous studies about fuzzy neural network with application to sales forecasting [7], the purpose of this study is to compare SARIMA and Holt-Winter's Exponential Smoothing methods in an attempt to generate customer transaction forecasting in Store X with high accuracy.[8-10]

2. Methods

Seasonal Autoregressive Integrated Moving Average or better known as SARIMA method is Time Series forecasting method for stochastic model data with seasonal data pattern [7]. According to Box and Jenkins (1976), Palit and Popovic (2005), Shumway and Stoffer (2006), as cited in [8] the SARIMA modeling consists of four steps (1) The model identification phase identifies the variables in order to analyze and verify the stationarity of the time-series; this also determines the most relevant combination of auto-regression and moving average; (2) The model estimation phase reviews the models identified in the first step and determines the most efficient one; (3) The model validation phase tests the precision of the chosen model; possible enhancements are also established during this phase; (4) The model forecasting phase predicts the future data of the series which are delivered with a confidence interval.

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 [9].

2.1. SARIMA Model

Seasonal Method Autoregressive Integrated Moving Average or better known as SARIMA method is Time Series forecasting method for stochastic model data with seasonal data pattern [7].

Generally, SARIMA notation is:

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

With:

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

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

s : Number of periods per season

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

$$\Phi\ _{\mathrm{p}}\mathrm{B}^{\mathrm{S}}\phi_{\mathrm{p}}(\mathrm{B})(1-B)^{\mathrm{d}}(1-\mathrm{B}^{\mathrm{S}})^{\mathrm{D}}\mathrm{Z}_{\mathrm{t}} = \theta_{\mathrm{q}}(B)\Theta_{\mathrm{q}}(B^{\mathrm{S}})\mathrm{a}_{\mathrm{t}} \tag{1}$$

With :

$_{\mathrm{p}}\mathrm{B}$: AR *Non Seasonal*

$\Phi\ _{\mathrm{p}}\mathrm{B}^{\mathrm{S}}$: AR *Seasonal*

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

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

$\theta\ _q(\mathrm{B})$: MA *non seasonal*

$\Theta\ _q(\mathrm{B}^{\mathrm{S}})$: MA *seasonal*

2.2. Holt-Winter's Model

This method is used when the data show a trend and seasonal behaviour . To deal with seasonality, it have been developed a third equation parameter called the "Holt-Winters" method in accordance with the name of the inventor. The Holt-Winters method is often called the exponential smoothing method that approaches [10]. This method is divided into two parts namely Multiplicative Seasonal Method

IOP Conf. Series: Materials Science and Engineering **407** (2018) 012153 doi:10.1088/1757-899X/407/1/012153

(Multiplicative Seasonal Method) which is used for seasonal variation of data that has increased / decreased (fluctuation), and Additive Seasonal Methods used for constant seasonal variation.

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 Square Error (MSE).

2.3.1. Mean Absolute Deviation (MAD). MAD is the value of the overall forecasting error for a model. The MAD value is calculated by taking the sum of the absolute values of the forecasting error divided by the number of data periods.

$$MAD = \frac{\Sigma|actual\ value - forecasting|}{n} \quad (2)$$

Where :
n = number of data periods

2.3.2. Mean Square Error (MSE). Another popular measures of accuracy is MSE, it is widely used in study of forecasting [11]–[13]

$$MSE = \frac{\Sigma(forecast\ error)^2}{n} \quad (3)$$

Where :
n = number of data periods

3. Results and Discussion

3.1. SARIMA model

3.1.1. Model identification. The process of identifying the data model is done by looking at the plot of the actual data, as well as whether the data is stationary or not by looking at the plot of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) (Figure 1-2).

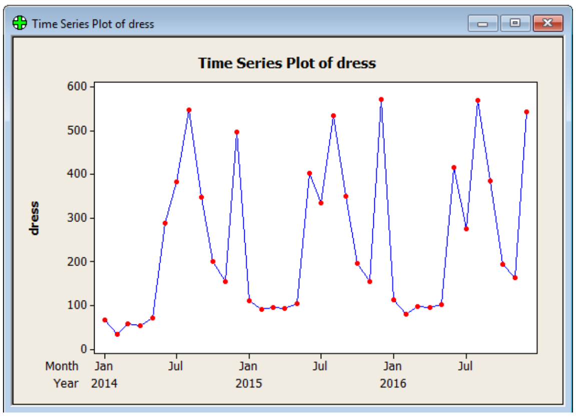

Figure 1. Time series plot.

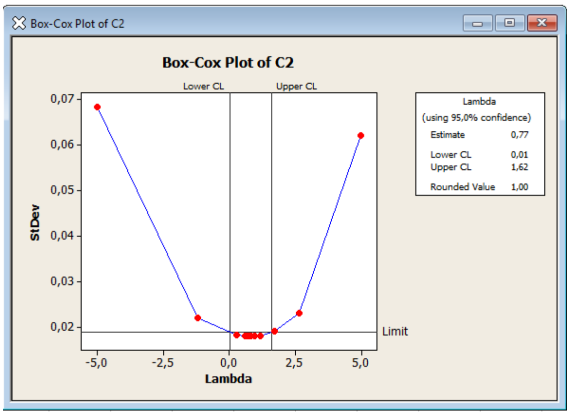

Figure 2. Box-Cox Transformation process.

Figure 1 shows that the data is influenced by seasonal pattern because there is repetition of pattern in certain period (yearly). Figure 2 shows the result of data transformation process with Box-Cox Transformation. The data is stationary on the variance because the value of rounded value on the Box-Cox plot is already worth 1 (figure 3).

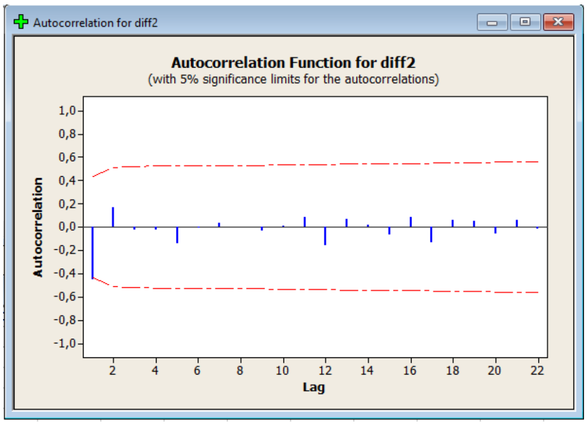

Figure 3. ACF Plot.

Figure 3 shows the plot of the ACF, where the data is stationary to the range because there is no lag coming out of the dashed line (confidence interval).

3.1.2. Parameter Estimation. The model estimation phase reviews the models identified in the first step and determines the most efficient one. The candidate is shown in table 1.

Table 1. SARIMA candidate model.

MODEL	MAD	EXPLANATION
SARIMA $(1,1,0)(1,1,0)^{12}$	9,430	AR(1) dan SAR(12) significant
SARIMA $(0,1,1)(0,1,0)^{12}$	6,678	MA(1) significant
SARIMA $(1,1,0)(0,1,0)^{12}$	5,592	AR(1) significant

From table 1 it can be concluded that the best SARIMA model for dress sales data is SARIMA $(1,1,0)$ $(0,1,0)^{12}$ because it has the smallest MAD value that is 5,592. This shows that the model can be used for forecasting.

3.1.3. Forecasting with SARIMA. After determining the model, selected model can now be used for forecasting, table 2 below shows the forecasting phase (table 2).

Table 2. Forecasting 3 months ahead.

Period	Actual Value	Forecast	MAD
January 2017	110	107,832	2,168
February 2017	74	72,408	1,592
March 2017	98	84,984	13,016
		MAD =	**5,592**

Table 2 shows the forecasting result with SARIMA model $(1,1,0)(0,1,0)^{12}$ for January to March 2017.

3.2. Holt-Winter's Exponential Smoothing

The steps of forecasting with Holt-Winter's Exponential Smoothing method is to analyze the data, whether it contains trend and seasonal elements by looking at the pattern formed, then the data is predicted using Holt-Winter's Exponential Smoothing with seasonal multiplication method or seasonal addition method, then compare where the smallest error value between two methods. Based on the comparison of both multiplicative and additive method apparently that the multiplicative method shows the smallest error (Table 3).

Table 3. Comparison of different smoothing weight with seasonal multiplicative method.

PERIOD	ACTUAL	SEASONAL MULTIPLICATIVE METHOD									
		SMOOTHING WEIGHT									
		0,1	**ERROR**	**0,2**	**ERROR**	**0,3**	**ERROR**	**0,4**	**ERROR**	**0,5**	**ERROR**
Jan-17	110	139,68	29,68	126,95	16,95	128,42	18,42	121,99	11,99	119,22	9,22
Feb-17	74	84,30	10,30	76,66	2,66	78,40	4,40	74,00	0,00	72,93	1,07
Mar-17	98	100,05	2,05	90,98	7,02	90,80	7,20	82,39	15,61	80,77	17,23
		MAD =	**14,01**	**MAD=**	**8,87**	**MAD=**	**10,01**	**MAD=**	**9,20**	**MAD=**	**9,18**

PERIOD	ACTUAL	SEASONAL MULTIPLICATIVE METHOD									
		SMOOTHING WEIGHT									
		0,6	**ERROR**	**0,7**	**ERROR**	**0,9**	**ERROR**	**0,4**	**ERROR**	**1**	**ERROR**
Jan-17	110	121,61	11,61	117,24	7,24	107,21	2,79	121,99	11,99	97,31	12,69
Feb-17	74	74,49	0,49	68,95	5,05	63,98	10,02	74,00	0,00	67,20	6,80
Mar-17	98	82,91	15,09	76,25	21,75	79,50	18,50	82,39	15,61	80,26	17,74
		MAD =	**9,06**	**MAD=**	**11,34**	**MAD=**	**10,44**	**MAD=**	**9,20**	**MAD=**	**12,41**

Table 3 shows the Comparison of different smoothing weight with seasonal multiplication method. It appears in Table 3 that the smallest MAD value in the seasonal multiplication model is the result of forecasting with a smoothing weight of 0.2 with MAD 8.87.

3.3. Comparison of forecasting method

Some previous research in the field of forecasting with seasonal patterned data shows various results. In Rana [11] SARIMA model shows better performance compared to Single Exponential Smoothing, Holt's, SNAIVE, and Holt Winter's Exponential Smoothing. In Sartono [14] SARIMA and Holt Winter's shows more accurate forecast, each on different region of research object. In Hu [15] comparison between ARIMA and SARIMA to forecast wind speed indicates that SARIMA is more accurate because it is able to identify seasonal data..

4. Conclusions

This paper compares two forecasting method: SARIMA and Holt-Winter's Exponential Smoothing applied forecast sales transaction at store X. It is known that the best model to forecast sales transactions in store X is the SARIMA model (1,1,0) $(0,1,0)^{12}$. This conclusion is concluded based on the comparison of the smallest error value with the MAD value of 5.592 on the SARIMA model (1,1,0) $(0,1,0)^{12}$. With accurate forecasting results or estimates, the store management can determine the right strategy in determining the timing and selection of collections of goods to be displayed or to be stored first in the warehouse.

Acknowledgements
We acknowledged UNIKOM for supporting this study.

References

[1] C W Chu and G P Zhang 2003 A comparative study of linear and nonlinear models for aggregate retail sales forecasting *Int. J. Prod. Econ.* **86** 3 pp. 217–231
[2] R G Brown 2015 Smoothing, forecasting and prediction of discrete time series *Quant. methods*

Ser. Prentice-Hall Int. Ser. Manag. pp. 1–468 1963
[3] H Vaagen and S W Wallace 2009 Product variety arising from hedging in the fashion supply chains *Int. J. Prod. Econ.* **114** 2 pp. 431–455
[4] G E P Box, G M Jenkins and G C Reinsel 2000 Time Series Analysis - Forecasting and Control *Prentice Hall New Jersey 1994* **373** 5 837–900
[5] H Nair, J P Dubé and P Chintagunta 2001 Accounting for Primary and Secondary Demand Effects with Aggregate Data *Mark. Sci.* **24**(3) pp. 444–460
[6] R J Kuo and K C Xue 1999 Fuzzy neural networks with application to sales forecasting *Fuzzy Sets Syst.* **108** 2 pp. 123–143
[7] W W S Wei 2006 *Time Series Analysis Univariate and Multivariate Methods SECOND EDITION*
[8] M Bouzerdoum, A Mellit and A Massi Pavan 2011 A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant *Sol. Energy* **98** pp. 226–235
[9] P Kalekar Time series forecasting using Holt-Winters exponential smoothing *Kanwal Rekhi Sch. Inf. Technol.* 4329008 pp. 1–13
[10] S G Makridakis, S C Wheelwright and R J Hyndman 1998 Forecasting: Methods and Applications *J. Forecast.* p. 1
[11] S Rana Performance Assessment of SARIMA Model with Holt– Winter's Trend and Additive Seasonality Smoothing Method **16** 2 pp.34-40
[12] X Zhang, Y Liu, M Yang T Zhang A A Young and X Li 2013 Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China *PLoS One* **8**(5)
[13] C Christodoulos, C Michalakelis and D Varoutas 2011 On the combination of exponential smoothing and diffusion forecasts: An application to broadband diffusion in the OECD area *Technol. Forecast. Soc. Change* **78** 1 pp. 163–170
[14] Sartono 2017 Models for Train Passenger Forecasting of Java and Sumatra *J. Phys. Conf. Ser.*, **824** 1, p. 12032
[15] J Hu, J Wang and G Zeng 2012 A hybrid forecasting approach applied to wind speed time series *Renew. Energy* **60** 185–194