Forecasting of Raw Material Needed for Plastic Products Based in Income Data Using ARIMA Method

Baihaqi Siregar, Erna Budhiarti Nababan, Alexander Yap, Ulfi Andayani
Department of Information Technology
University of Sumatera Utara
Medan, Indonesia
baihaqi@gmail.com

Fahmi
Department of Electrical Engineering,
University of Sumatera Utara,
Medan, Indonesia
fahmimn@usu.ac.id

***Abstract*— Forecasting is a process of predicting something future by doing calculations from previous data. In this case the authors will forecast the sale of plastic production by using ARIMA Box-Jenkins method for 2015 forecasting. The data used is the sales data of plastic factory production in Bandung from 2012 to 2014. This research will use ARIMA procedure in SAS that allows for identification, Estimation and forecasting of Time Series models. The measurement of the accuracy of forecasting results is done with the MAPE (Mean Absolute Percentage Error) value. Forecasting results conducted for 2015 using ARIMA (3.0, 2) on plastic product sales data for 2012 to 2014 resulted in a prediction accuracy rate of 74% for PP Trilene and 68% for PP Tintapro products.**

Keywords: forecasting, ARIMA, SAS, and MAPE

I. INTRODUCTION

The competition between companies Today is getting tighter, supported by the rapid development of science and technology forcing the company to compete to produce various quality products for the costumers but efficient from the production cost. Therefore, it is necessary to find forecasting techniques used to predict future production results from previous production data.

Forecasting is an accurate calculation in determining something to come by using past data. The results of forecasting are used for production planning, scheduling, and control of raw material inventory. Sales forecasting will greatly affect the working capital of a company, i.e. the amount of goods produced and raw materials ordered must be in accordance with the needs. Weak forecasting will lead to exhausted stock or stocks, low service levels, inefficient resource utilization [1]. Therefore it is necessary forecasting techniques that have an accurate upcoming picture.

Several studies have been conducted to forecast sales data. One of them is the research done by Stephanie. To forecast the sale of baby milk products with the method of gray system theory and neural network [2]. The method of gray system theory and neural network is chosen because it has a higher level of accuracy than other traditional methods. Furthermore Jarrett & Kyper conducted a study on ARIMA modeling with an intention to forecast and analysis of Chinese stock prices [3]. Kandananond performs autocorrelation comparisons on different types of forecasting [4]. In order to forecast constant data required, Box and Cox therefor use transformations in data that have not been constant [5]. Further Warsini Perform comparison of exponential smoothing methods and ARIMA box-Jenkins as a method of forecasting composite stock price index (CSPI). Existing data will be analyzed and compared forecasting result with smoothing method and ARIMA method [6].

Forecasting techniques are often used is the time series method. Time series is a process of predicting past data to predict it into the future. In time series method there are several traditional models such as moving average (MA), autoregressive (AR), exponential smoothing (ESM). However, this traditional model is still considered less accurate. Therefore it will be forecast using ARIMA model to get better result in this research.

II. PROBLEM IDENTIFICATION

Preparation of long-term plans is necessary to get a picture of the company's business in the future. Whether business activity will increase or decrease. For the factory this will affect the production process is on how many products will be produced that impact on raw material inventory to be ordered. Therefore, an approach is needed that is able to forecast between the amount of production and inventory control in the future based on historical data of previous sales.

III. PREVIOUS RESEARCH

There are several studies that have been done by previous authors on ARIMA methods applied in forecasting, among others, ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices [3], Comparison of Exponential Smoothing Methods and ARIMA (Box-Jenkins) as Forecasting Method Index Composite Stock Price (IHSG) [6], Monthly Rainfall Forecast Pre casting in Medan City by Box-Jenkins Method [7]. There are several studies that have been done by

978-1-5386-0355-0/17/$31.00 ©2017 IEEE

previous authors on ARIMA methods applied in forecasting, among others, ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices [3], Comparison of Exponential Smoothing Methods and ARIMA (Box-Jenkins) as Forecasting Method Index Composite Stock Price (IHSG) [6], Monthly Rainfall Forecast Pre casting in Medan City by Box-Jenkins Method [7].

IV. Methodology

A. General Architecture

In this study, general architecture in the process of forecasting the needs of raw materials for the manufacture of plastic products based on income data is shown in Fig. 1. The first stage is to collect data from sales history data and field reviews, then Preprocessing data and the process of SAS will be done. Forecasting results will be compressed into a JSON file that will be sent to the web server and processed into a web dash board.

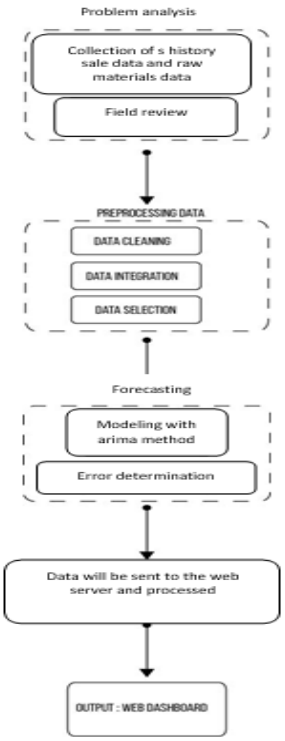

Fig. 1. General Architecture

B. Data Collection

The data used in this study is daily income data at a plastic factory in Bandung from January 2012 to December 2014. The data is grouped per month because the data will be predicted annually. The data used is data from PP type plastic income. The data used in this study amounted to 36 periods. Where 1 period is 1 month.

C. Preprocessing Data

Before the data used in the forecasting process needs to be done preprocessing data. Because the data still contains errors and incomplete. Therefore, it is necessary to prepare data preprocessing with SAS software.

1) Data Cleaning

In the data cleaning stage will be done, among others, such as eliminating noise, filling missing values, identifying outliers and improve data that is not consistent with the final goal of forecasting.

2) Data Integration

In the data integration stage to be done is to combine or combine income data from each month during the period 2012 to 2014 into one.

3) Data Selection

In the data selection stage is to select which data are relevant and necessary for forecasting.

D. Forecasting

There are several stages in the forecasting with ARIMA method that is constant data, estimation, diagnosis, and forecasting. Forecasting revenue using ARIMA method is a forecasting set of time series data that will provide forecasting results based on data in the past.

E. Level of Accuracy Calculation

To calculate the accuracy of forecasting by using the Mean Absolute Percent Error (MAPE) method.

F. Data Delivery Using Web Server

After getting the best model, the data forecasting of the model will be used to find the values of raw materials that exist in the future. After that the results of information and statistical charts obtained will be compressed into the form of JSON files (Java Script Object Notation). The JSON file will be sent to the web server.

V. Result And Discussion

A. Implementation

At this stage, ARIMA forecasting that has been obtained will be implemented into the system using a web programming language in accordance with the design that has been done.

The author uses SAS software to forecast the needs of raw materials for making plastic products based on income data using ARIMA method. Furthermore the authors use web dashboard to design the whole system to be built.

The sales data page is a page to display all sales data of plastic products from JSON files that exist on the server. On this page the user can choose the time period of the data you want to display. After selecting the option, the system will display the sales graph according to the specified option. The sales data page can be seen in Fig. 2.

Fig. 2. Sales Data Page

This page displays the total sales chart of all plastic products in a year as shown in Fig. 3.

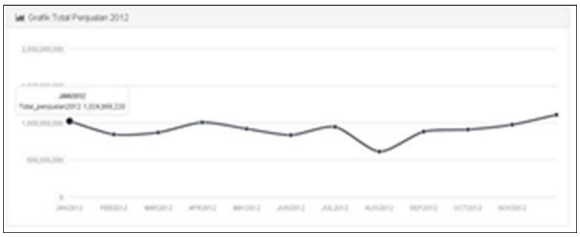

Fig. 3. Graph of Total Sales for a Year

This page also shows sales charts of each type of plastic product sales in a year as shown in Fig. 4.

Fig. 4. Graph of Each Type Plastic Product Sales for a Year

The forecasting page is a page for displaying predictable plastic products from a JSON file on the server. On this page users can choose what plastic product forecasting results you want to display. After selecting the option, the system will display the forecast graph according to the specified option. The forecasting page results can be seen in Fig. 5.

Fig. 5. Plastic Product Forecasting Page

This page displays the graph of plastic product forecasting as shown in Fig. 6.

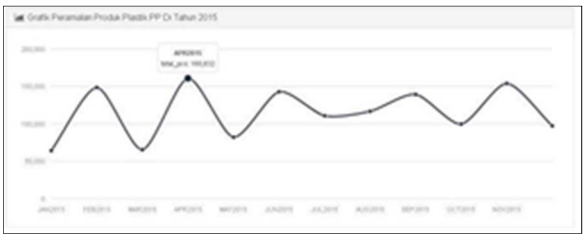

Fig. 6. Product Forecasting Chart on Plastic Products Forecasting Page

This page also displays graphs and data forecasting of raw materials needed in the future as shown in Fig. 7.

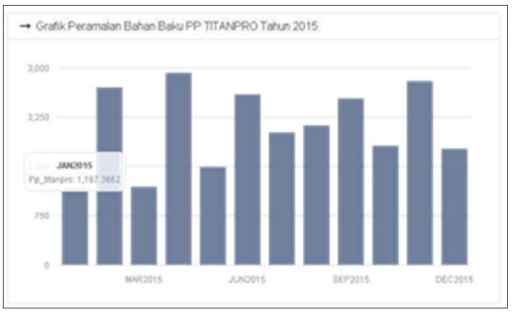

Fig. 7. Graph of Plastic Product Raw Forecasting on Plastic Products Forecasting Page

B. *System Performance Testing*

System performance test is used to know the performance of the system that has been developed. Suppose there are 36 sales data of PP plastic products grouped per month.

The first step is to find the value of autocorrelation and partial autocorrelation of PP product data. Results of autocorrelation and partial autocorrelation in Table IV.

TABLE I. AUTOCORRELATION VALUE AND PARTIAL AUTOCORRELATION OF PP PRODUCT DATA

LAG	ACF	PACF	LAG	ACF	PACF
1	0.172293	0.17229	18	0.027419	-0.0952
2	-0.0004	-0.0356	19	0.116324	0.13884
3	-0.18707	-0.1858	20	0.021105	-0.0929
4	-0.14192	-0.0831	21	-0.00971	-0.0625
5	0.116907	0.16367	22	0.054417	-0.0508
6	-0.12787	-0.2265	23	0.062469	0.06884
7	0.048208	0.07644	24	0.071478	-0.0510
8	-0.20143	-0.2103	25	-0.05374	0.00744
9	-0.13135	-0.0911	26	-0.02833	0.00566
10	-0.02478	-0.0335	27	-0.07425	-0.0990
11	-0.04582	-0.0535	28	-0.01388	-0.1054
12	0.089939	-0.0404	29	0.014525	0.061692
13	-0.17564	-0.1830	30	-0.06468	0.00239
14	-0.27060	-0.3535	31	0.005076	0.03864
15	-0.06116	0.01198	32	-0.06468	0.01423
16	0.122381	0.02523	33	-0.01958	0.03558
17	0.219755	-0.0662	34	0.004168	0.01999

To see if the data has been constant in the average or not. Can be seen in the autocorrelation coefficient is significantly different from zero. A time series data is random if its correlation coefficient is present at the interval limit (-1.96 <= z <= 1.96).

To see if the data has been constant in the variance or not can be done with the box-cox test in Fig. 8.

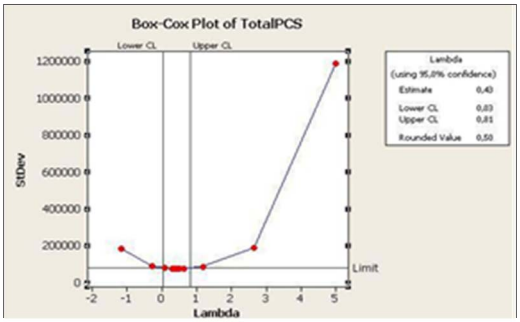

Fig. 8. Box-Cox Constant Variance Test

Lambda = 0.5 is found so that the root transformation is done. Then got the value of transformation result in Table II.

TABLE II. PP PRODUCT DATA TRANSFORMATION VALUE

Lag	Transformation	Lag	Transformation
1	94.809	19	117.703
2	90.723	20	79.806
3	100.101	21	87.981
4	78.079	22	103.467
5	101.075	23	87.415
6	87.881	24	92.608
7	92.687	25	69.284
8	87.240	26	85.415
9	96.574	27	98.893
10	107.599	28	105.021
11	107.951	29	111.749
12	112.631	30	94.962
13	84.262	31	102.577
14	106.514	32	89.870
15	87.714	33	95.077
16	94.767	34	117.542
17	109.343	35	100.673
18	119.076	36	93.375

Then select the appropriate ARIMA model based on autocorrelation and partial autocorrelation and got the suitable model is ARIMA (3.0, 2).

After finding the suitable model, then the AR and MA coefficients are estimated and the t-value value can be seen in Table III.

TABLE III. AR & MA COEFFICIENTS AND TRANSFORMATION VALUE FOR PP PRODUCT

Parameter	Estimation	T-value
MA 1,1	-1.72557	-2.29
MA 1,2	-0.96848	-1.15
AR 1,1	-1.41829	-9.34
AR 1,2	-0.42228	-4.31
AR 1,3	0.32890	22.83

To find t-table, firstly we have to find the value of degrees of freedom (DF) = amount of data - the number of parameters of ARIMA model (3.0, 2) and obtained DF=31.

Next look for the residual autocorrelation value of ARIMA (3.0, 2). Then get the value of autocorrelation in Table IV.

TABLE IV. ARIMA RESIDUAL AUTOCORRELATION VALUE (3.0, 2)

Lag	ACF	Lag	ACF
1	0.059	3	-0.240
2	-0.131	4	-0.018
5	0.085	15	-0.069
6	-0.026	16	0.226
7	-0.010	17	0.135
8	-0.153	18	0.024
9	-0.108	19	0.001
10	-0.080	20	0.032
11	0.113	21	-0.017
12	0.059	22	0.005
13	-0.001	23	0.098
14	-0.361	24	0.018

After obtaining the result of the residual autocorrelation, then look for the value of chi square with (2.13) based on lag 6, 12, 18, and 24. And search for degrees of freedom (DF) = number of lag - number of ARIMA parameters (3.0, 2). The results can be seen in Table V.

TABLE V. CHI-SQUARE RESULTS WITH LJUNG-BOX TESTING

Lag	DF	Chi-Square
6	1	3.58
12	7	6.57
18	13	19.82
24	19	20.99

Next is to find the value of raw material of PP Trilene with 80% calculation of product forecasting of PP and PP Tintapro with 20% of result of PP product forecasting result can be seen in Table VI.

TABLE VI. FORECASTING RESULTS OF PP TINTAPRO AND PP TRILENE PRODUCT

Month	PP Tintapro Product Material	PP Trilene Product Material
Jan 2015	6300	1575
Feb 2015	7974.4	1993.6
Mar 2015	7224	1806
Apr 2015	7313.6	1828.4
Mei 2015	8072.2	2018.8
Jun 2015	6736.8	1684.2
Jul 2015	8366.4	2091.6
Aug 2015	6865.6	1716.4
Sep 2015	7834.4	1958.6
Oct 2015	7604.8	1901.2
Nov 2015	7033.6	1758.4

The next step is to test the accuracy of the forecasting against the original data by calculating the error value. The value of this error is obtained by comparing the value of the forecasting results with the actual data value already owned. The result of error value from PP Tintapro product can be seen in Table VII and error value for PP Trilene in Table VIII.

TABLE VII. FORECASTING RESULTS OF PP TINTAPRO

Month	Original Data	Forecasting	Error
Jan 2015	10650	6300	1125
Feb 2015	9000	7974.4	256.4
Mar 2015	5000	7224	1006
Apr 2015	6000	7313.6	328
Mei 2015	10500	8072.2	681.2
Jun 2015	6500	6736.8	15.8
Jul 2015	11000	8366.4	658.4
Aug 2015	6000	6865.6	216.4
Sep 2015	5000	7834.4	708.4
Oct 2015	7200	7604.8	101.2
Nov 2015	5000	7033.6	2458.4

TABLE VIII. FORECASTING RESULTS OF PP TRILENE

Month	Original Data	Forecasting	Error
Jan 2015	2700	1575	4350
Feb 2015	2250	1993.6	1025.6
Mar 2015	800	1806	2224
Apr 2015	1500	1828.4	1313.6
Mei 2015	2700	2018.8	2427
Jun 2015	1700	1684.2	236.8
Jul 2015	2750	2091.6	2633.6
Aug 2015	1500	1716.4	865.8
Sep 2015	1250	1958.6	2834.4
Oct 2015	1800	1901.2	404.8
Nov 2015	1300	1758.4	2033.6

After getting the error value every month, then next is calculate the value of MAPE from product of PP Trilene and PP Tintapro that is by summing result of division of error value every month with original data then divide by number of month. The result of MAPE for PP Trilene product is 26% means the accuracy rate of 74% and for PP Tintapro product of 32% means the accuracy of 68%.

VI. CONCLUSION

Based on the testing of raw material forecasting by using ARIMA method obtained some conclusions:

- Forecasting results by using ARIMA method has a fairly good accuracy on short-term forecasting, while for long-term forecasting will usually tend to be flat/constant. Because ARIMA method completely ignores the independent variables and uses only the past values of the independent variables to produce short-term forecasting.
- Forecasting by using ARIMA method on PP product has accuracy level for PP Trilene product equal to 74% and for PP Tintapro product equal to 68%.

REFERENCES

[1] M. E. Nenni, L. Giustiniano and L. Pirolo, "Demand Forecasting in the Fashion Industry: A Review," International Journal of Engineering Business Management 5(37), 2013.

[2] L. Stepvhanie, "Peramalan Penjualan Produk Susu Bayi Dengan Metode Grey System Theory dan Neural Network," Universitas Indonesia, 2012.

[3] J. E. Jarrett and E. Kyper, "ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices," International Journal of Engineering Business Management 3(53), 2011.

[4] K. Kandananond, "A Comparison of Various Forecasting Methods for Autocorrelated Time Series," International Journal of Engineering Business Management 4(4), 2012.

[5] G. E. Box and D. R. Cox, "An Analysis of Transformations," Journal of The Royal Statistical 26(2), pp. 211-243, 1964.

[6] Warsini, "Perbandingan Metode Pemulusan (Smoothing) Eksponensial dan ARIMA (Box-Jenkins) Sebagai Metode Peramalan Indeks Harga Saham Gabungan (IHSG)," Universitas Sumatera Utara, 2011.

[7] J. H. Sembiring, "Peramalan Curah Hujan Bulanan Di Kota Medan Dengan Metode Box-Jenkins," Universitas Sumatera Utara, 2010.