The Application of Prediction of Cocoa Yields with Exponential Smoothing Method

Boldson Herdianto Situmorang
Faculty of Mathematical and Natural Sciences
Pakuan University
Bogor, Indonesia
boldson_skom@yahoo.com

Arie Qur'ania, Dian Asri Sukmawati
Faculty of Mathematical and Natural Sciences
Pakuan University
Bogor, Indonesia

Abstract*— Cocoa is the one of Indonesia superior comodities from agricultural sector, besides oil palm and rubber which contribute the foreign exchange for the country in the amount of US$ 1,05 billion from export of cocoa seed and cocoa product in 2012 [1]. Caused by cocoa plant is concidered important for Indonesia economic matters, a system of prediction of cocoa yields is needed to minimize the decrease of cocoa yields in every year. The prediction may be used of various calculation methods, among others by analysis method of Exponential Smoothing–a prediction method of moving average by weighting where data is weighted by an exponential function [2]. This research has purpose to design and develop an application which can be used to predict the cocoa yields by Exponential Smoothing method so that can help the government and farmers to predict cocoa yields per annum. The development of this application will be limited by calculation of statistics to predict the cocoa yields base on data in previous year. The data which will be used in this research is the secondary data which were from Research Institute for Industrial Plants, that are data of cocoa yields from 2008 to 2012. The development of the application is using softwares with XAMPP, PHP, and MySQL. The prediction of cocoa yields in 2013 at significance level 0,5 is 17.885 Kg and it is the prediction result with the minimum value of MSE (Mean Square Error) 4.209. After this, the results of prediction of cocoa yields per quarterly has been a little bit change in the certain seasons. For the validation in calculation of MAPD (Mean Absolute Percentage Definition) the Exponential Smoothing is exactly accurate in MAPD value at significance level 0,1 with the confidence interval 99,427 percent, at significance level 0,3 with the confidence interval 99,86 percent, and at significance level 0,5 with the confidence interval 99,57 percent.

Keywords— application; prediction; cocoa yields; exponential smoothing

I. INTRODUCTION

Cocoa is the one of Indonesia superior comodities from agricultural sector, besides oil palm and rubber which contribute the foreign exchange for the country in the amount of US$1,05 billion from export of cocoa seed and cocoa product in 2012 [1]. Caused by cocoa plant is concidered important for Indonesia economic matters, a system of prediction of cocoa yields is needed to minimize the decrease of cocoa yields in every year.

The prediction may be used of various calculation methods, among others by analysis method of Exponential Smoothing–a prediction method of moving average by weighting where data is weighted by an exponential function [2]. This research has purpose to design and develop an application which can be used to predict the cocoa yields by Exponential Smoothing method so that can help the government and farmers to predict cocoa yields per annum.

The development of this application will be limited by calculation of statistics to predict the cocoa yields base on data in previous year. The data which will be used in this research is the secondary data which were from Research Institute for Industrial Plants, that are data of cocoa yields from 2008 to 2012. The development of the application is using softwares with XAMPP, PHP, and MySQL.

II. MATERIAL AND METHODS

The development of the application is using a structured system development method by implement the steps of System Development Life Cycle (SDLC) as shown in figure 1.

A. Planning

Informations about problems, rule and regulations are collected in the planning phase. Determine criterias, restricting problems, and provide alternative solutions are also carried in this phase. The data which will be used in this research is the secondary data which were from Research Institute for Industrial Plants by:

1) Interviews

Interviews were conducted to understand the current system and the calculation process of prediction of cocoa yields in Research Institute for Industrial Plants.

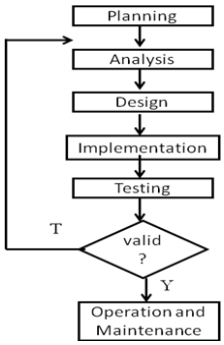

Fig. 1. Software Development Life Cycle (SDLC)

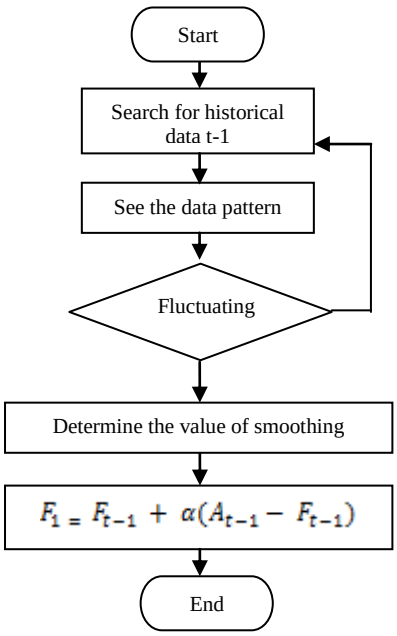

Fig. 2. Flowchart of Exponential Smoothing Method

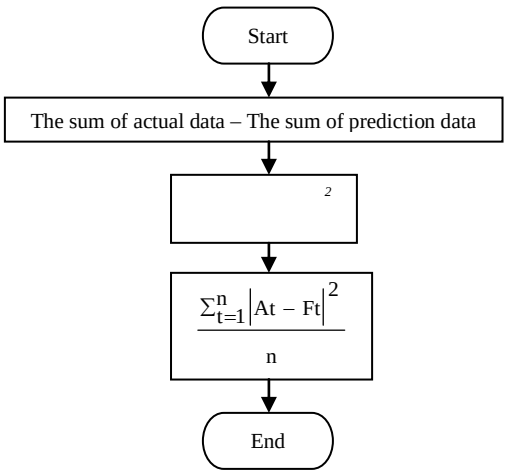

Fig. 3. Flowchart of Mean Square Error

2) Literature Study

Information required to support the implementation of the research is not only derived from Research Institute for Industrial Plants but also through reference books and journals.

B. *Analysis*

Analysis phase is the phase to identify the problems, to understand the current system, to analyze the results, and reporting. In this phase also selection and processing of statistical data related to the prediction of cocoa yields with the following steps:

1. Plot the data of cocoa yields.
2. The determination of the value of smoothing constant.
3. Prediction calculation using Exponential Smoothing with seasonal variation in the steps as shown in figure 2. Seasonal variation is fluctuations that arise every year which is usually caused by the climate, usually has a consistent pattern over time [3].

$$\text{Moving Average} = \frac{\sum x}{N} \quad (1)$$

$$\text{Total Average} = \frac{\text{Moving Average}}{nt} \quad (2)$$

Note :
x = Sum of data on the entire period
n = Sum of period
nt = Sum of data on t period

4. Validation with *MSE (Mean Square Error)* to determine the average error arising from any prediction by following calculation steps [4] as in figure 3.

C. *Design*

The design of the Application of Prediction of Cocoa Yields consists of database design, conceptual design, and detail design.

D. *Implementation*

At this phase, the Exponential Smoothing method is implemented in a computer system by using software XAMPP, PHP, and MySQL.

E. *Testing*

The purpose of testing is to ensure that the application is accordance with user requirements and to remove or to minimize program defects. Testing the level of acceptance of the application will end when the user can receive the application based on the criteria that have been established.

F. Operation and Maintenance

At this phase, The Application of Prediction of Cocoa Yields is ready to use. When the application is feasible to operate, the next phase is maintenance of application. This research has used several tools and materials as follows :

1. Software
   a. Windows 7 Ultimate
   b. Xampp v.1.7.7
   c. *Microsoft Office 2007*
   d. Notepad++ v.6.2.3
2. Hardware
   a. Processor Intel Core i3, RAM 4 GB 2.20 GHz
   b. Harddisk 500 GB
   c. Flash Disk 2 GB
   d. Printer
3. Reference books and journals

III. Results and Discussion

A. Framework of Thinking

Look at Figure 4.

B. Exponential Smoothing Method

Data per quarter from the previous period is needed to predict. Data of cocoa yields based on secondary data from Research Institute for Industrial Plants 2008-2012 are as follows:

TABLE I. Cocoa Yields in 2008-2012

#	Type of Cocoa	Cocoa Yields (Kg)				
		2008	***2009***	***2010***	***2011***	***2012***
1	Criollo	74070	74200	77280	64470	68720
2	Forestero	100780	30692	16778	18679	14562
3	Trinatario	46415	67010	63010	73810	61220

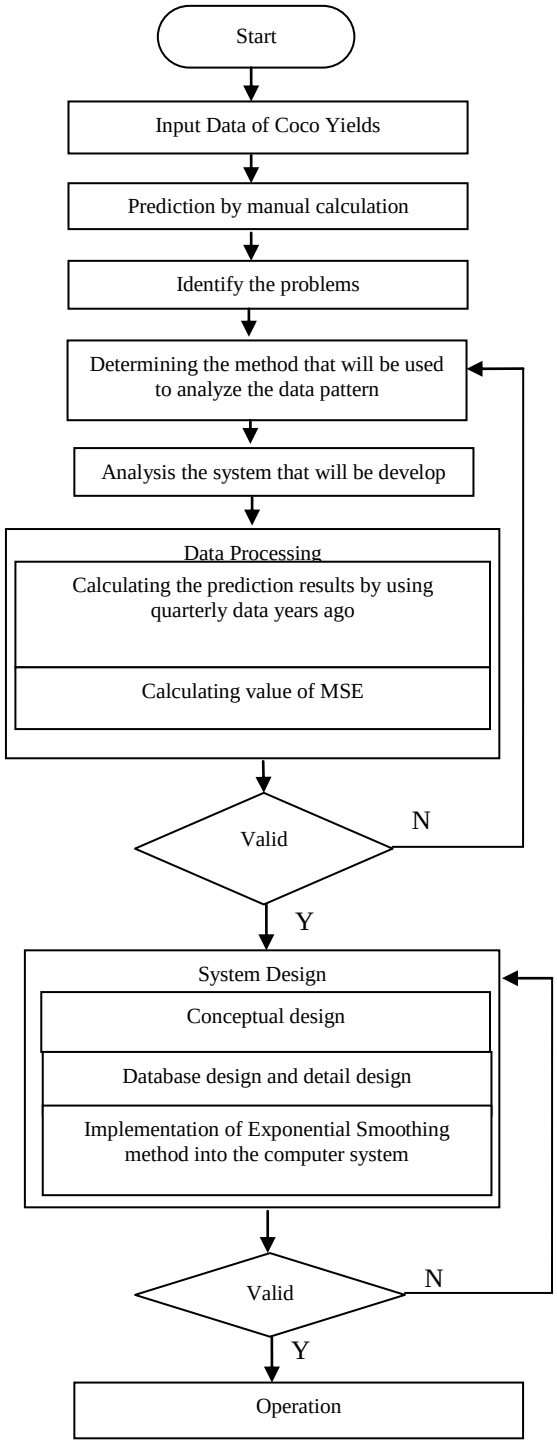

Fig. 4. Framework of Thinking

TABLE II. COCOA YIELDS IN 2010 - 2012

#	Type of Cocoa	Cocoa Yields (Kg) in 2010			
		I	*II*	*III*	*IV*
1	Criollo	17370	18420	19610	21880
2	Forestero	4185	4176	4255	4162
3	Trinatario	18720	15580	13370	15340
		Cocoa Yields (Kg) in 2011			
4	Criollo	14260	16430	17720	16060
5	Forestero	4643	4643	4672	4721
6	Trinatario	17335	18110	20155	18210
		Cocoa Yields (Kg) in 2012			
7	Criollo	15510	16730	17210	19270
8	Forestero	3000	1702	9060	800
9	Trinatario	19771	19776	11443	10230

TABLE III. CALCULATION RESULTS OF EXPONENTIAL SMOOTHING

Year	Q	$F_{t-1}+0,1(A_{t-1}-F_{t-1})$	$F_{t-1}+0,3(A_{t-1}-F_{t-1})$	$F_{t-1}+0,5(A_{t-1}-F_{t-1})$
2008	IV	12459	13217	13975
2009	I	13108	14936	16462
2010	II	16439	19798	20266
2011	III	17862	19926	19904
2012	IV	16550	16330	16140
2013	I	16616	16594	16675

Based on data of cocoa yields in table 1 and 2 obtained motion pattern of secondary data at a time sequentially. So that the appropriate model are to be used in this forecasting is a time series model based on data collected or observed at all times in a row.

Time series model was tested by smoothing based on the average data from previous years by summing previous predicted values, which is between the actual value and prediction value, which is a regular movement up and down and happen at the same time [5].

After a test on time series model, the model is applied in predicting cocoa yields in 2013 for 1st qurter is Exponential Smoothing with significance level (α):0,1; α:0,3; α: 0,5.

Based on the analysis of time series is known that the cocoa yields have increased and decreased every year, so it will need to apply seasonal variation using a simple average method to determine the index of seasonal and quarterly target.

Prediction calculation of cocoa yields for 1st quarter of 2013 by using exponential smoothing with α : 0,1 and α:0,3 to give more weight to smaller and α:0,5 to give equal weight among the previous results prediction compared with previous actual data. Criollo cocoa prediction calculation results can be seen in Table 4.

Q4 2008 = 12080 + 0,1 (15870 – 12080) = 12459 Kg
Q1 2009 = 12459 + 0,1 (18950 – 12459) = 13108 Kg
Q2 2010 = 16050 + 0,1 (19940 – 16050) = 16439 Kg
Q3 2011 = 17734 + 0,1 (19015 – 17734) = 17862 Kg
Q4 2012 = 16513 + 0,1 (16890 –16513) = 16550 Kg

MSE calculation results with data per annum (from 2008 until 2012) based on the amount of data invoved in the prediction of cocoa yields in 2013:

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{167992}{22} = 7.636 \text{ Exponential Smoothing 0,1}$$

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{10054}{22} = 4.957 \text{ Exponential Smoothing 0,3}$$

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{92598}{22} = 4.209 \text{ Exponential Smoothing 0,5}$$

C. Seasonal Variation

Seasonal variation calculation is by calculating the average of cacao yields per quarter and the total average to determine the seasonal index as a parameter of Criollo yields targeted search per quarter.

Cocoa Yields Average =

$$\frac{74070 + 74200 + 77280 + 64470 + 68720}{5} = 71748 \text{ Kg}$$

$$\text{Q1 Average} = \frac{12080 + 16230 + 17370 + 14260 + 15510}{5} = 15090 \text{ Kg}$$

$$\text{Q2 Average} = \frac{15870 + 18180 + 18420 + 16430 + 16730}{5} = 17126 \text{ Kg}$$

$$\text{Q3 Average} = \frac{18950 + 19850 + 19610 + 17720 + 17210}{5} = 18668 \text{ Kg}$$

$$\text{Q4 Average} = \frac{27170 + 19940 + 21880 + 16060 + 19270}{5} = 20864 \text{ Kg}$$

Total Average = 71748 / 4 = 17937 Kg

After the results of the total average and average per quarter is obtained, then the seasonal index of cocoa yields can be calculated.

$$\text{Seasonal Index} = \frac{\text{Average per quarter x 100}}{\text{Total average}} \quad (6)$$

$$Q1 = \frac{15090 \times 100}{17937} = 84.1278 \text{ Kg}$$

$$Q2 = \frac{17126 \times 100}{17937} = 95.4786 \text{ Kg}$$

$$Q3 = \frac{18668 \times 100}{17937} = 104.0754 \text{ Kg}$$

$$Q4 = \frac{20864 \times 100}{17937} = 116.3182 \text{ Kg}$$

Fig. 5. Main Page

Fig. 6. Page of Cocoa Types

D. *The Application of Prediction of Cocoa Yields*

The main page contains menus Home, Cocoa Types, Production, Prediction of Cocoa Yields, and Setting. The view of the main page is shown in Figure 5.

Cocoa Types page displays a table of the types of cocoa with a description of each type of cocoa. The view of Cocoa Types page is shown in Figure 6.

Production page is designed for administrator to be able to manipulate data of cocoa yields which used as a reference for predicting the cocoa yields for next years. The view of Cocoa Types page is shown in Figure 7.

Prediction of Cocoa Yields page is designed to predict the cocoa yield with exponential smoothing method. Administrator performs input data of cocoa yield by selecting the type of cocoa and the year that will be the parameters of cocoa yield in the next years. The view of Cocoa Types page is shown in Figure 8.

Fig. 7. Page of Production

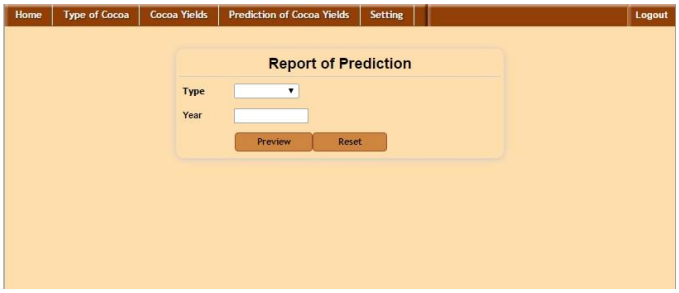

Fig. 8. Prediction of Cocoa Yields page

E. *Validation Test*

Validation test is a test to determine the accuracy of the prediction calculation result of the application system. This test is performed by comparing the result of forecasting using single exponential smoothing method manually with the result of the application system. Below are manually prediction calculation results of Criollo in 2013:

TABLE IV. PREDICTION CALCULATION RESULTS USING EXPONENTIAL SMOOTHING METHOD (A : 0,1)

Year	Quarter	$Ft_{-1}+0,1(At_{-1}- Ft_{-1})$	Ft (Kg)
2013	I	16550 + 0,1 (17210 – 16550)	16616
	II	16616 + 0.1 (19270 – 16616)	16881
	III	16881 + 0.1 (16970 – 16881)	16889
	IV	16889 + 0.1 (18240 – 16889)	17024

TABLE V. PREDICTION CALCULATION RESULTS USING EXPONENTIAL SMOOTHING METHOD (A : 0,3)

Year	Quarter	$Ft_{-1}+0,3(At_{-1} - Ft_{-1})$	Ft (Kg)
2013	I	16330 + 0,3 (17210 –16330)	16594
	II	16594 + 0.3 (19270 – 16594)	17396
	III	17396 + 0.3 (16970 – 17396)	17268
	IV	17268 + 0.3 (18240 – 17268)	17559

TABLE VI. PREDICTION CALCULATION RESULTS USING EXPONENTIAL SMOOTHING METHOD (A : 0,5)

Year	Quarter	$Ft_{-1}+0,5(At_{-1} - Ft_{-1})$	Ft (Kg)
2013	I	16140 + 0,5 (17210 – 16140)	16675
	II	16675 + 0.5 (19270 – 16675)	17972
	III	17972 + 0.5 (16961 – 17972)	17471
	IV	17466 + 0.5 (18239 – 17466)	17855

TABLE VII. THE SMALLEST VALUE OF MSE

Year	The Smallest Mean Square Error		
	Exponential Smoothing		
	α : 0,1	*α : 0,3*	*α : 0,5*
2013	7.636	4.957	4.209
2014	5.495	3.553	3.018
2015	4.318	2.669	2.267
2016	3.225	2.078	1.765
2017	2.583	1.664	1.413
2018	2.115	1.362	1.157
2019	1.763	1.135	0.964
2020	1.492	0.961	0.816

Manually calculation of MSE (Mean Square Error)

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{167992}{22} = 7.636 \text{ Exp. Smoothing 0,1}$$

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{10054}{22} = 4.957 \text{ Exp. Smoothing 0,3}$$

$$MSE = \frac{\sum (At-Ft)^2}{n} = \frac{92598}{22} = 4.209 \text{ Exp. Smoothing 0,5}$$

Manually calculation of seasonal variation

$$\text{Yields Avg} = \frac{74070+74200+77280+64470+68720}{5} = 71748 \text{ Kg}$$

$$\text{Q1 Average} = \frac{12080+16230+17370+14260+15510}{5} = 15090 \text{ Kg}$$

$$\text{Q2 Average} = \frac{15870+18180+18420+16430+16730}{5} = 17126 \text{ Kg}$$

$$\text{Q3 Average} = \frac{18950+19850+19610+17720+17210}{5} = 18668 \text{ Kg}$$

$$\text{Q4 Average} = \frac{27170+19940+21880+16060+19270}{5} = 20864 \text{ Kg}$$

Total Average = 71748 / 4 = 17937 Kg

Manually calculation of seasonal index

$$\text{Seasonal Index} = \frac{\text{Average per quarter x 100}}{\text{Total Average}}$$

$$Q1 = \frac{15090 \times 100}{17937} = 84.1278 \text{ Kg}$$

$$Q2 = \frac{17126 \times 100}{17937} = 95.4786 \text{ Kg}$$

$$Q3 = \frac{18668 \times 100}{17937} = 104.0754 \text{ Kg}$$

$$Q4 = \frac{20864 \times 100}{17937} = 116.3182 \text{ Kg}$$

Prediction of Criollo yields in 2013 on the application (as shown in Figure 9) looks for a change of predictions that have a bit of a complex pattern so that there data decompositioning into several components. Each component will be studied and searched one by one, once found will be merged again into the value of assessment or prediction with a simple average which is a regular movements that have meaning ups and downs at the same time or at a the adjacent time. Whether or not the movement of the seasons on the application can be seen in Figure 10.

Method	Prediction of Cocoa Yields (Kg) for Criollo in 2013				MSE
	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter	
Significance level=0.1	16,616	16,881	16,889	17,024	7,636
Significance level=0.3	16,594	17,396	17,268	17,559	4,957
Significance level=0.5	16,675	17,972	17,471	17,855	4,209

Fig. 9. Prediction Results of Criollo

Method	Prediction of Cocoa Yileds for Criollo in 2013 With Seasonal Variation				MSE
	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter	
Significance level=0.1	14,178	16,091	18,456	19,603	7,636
Significance level=0.3	14,474	16,426	17,905	20,012	4,957
Significance level=0.5	14,717	16,702	18,206	20,348	4,209

Fig. 10. Seasonal Variation of Criollo

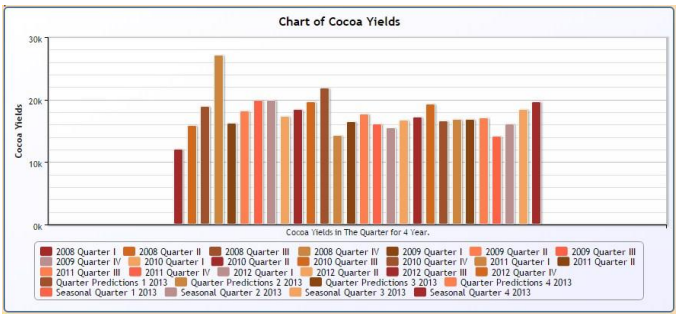

Fig. 11. Graph of Criollo Prediction Results in 2013 from 1st Quarter to 4th Quarter in The Application

The results of Criollo predictive testing from 2013 to 2020 by using time series data can be seen in the table below.

Based on the table 7 is known that the value of Mean Square Error on each year due to decreased since using secondary data from 2008 to 2012.

After testing the Mean Square Error prediction from 2013 to 2020 were calculated using exponential smoothing method with α=0.5, it will get the smallest Mean Square Error making, thus corresponding to predict cocoa yields in the next period.

Based on the graph of Criollo prediction results in 2013 from 1st quarter to 4th quarter on the application, we can see the results rapprochement between per quaterly. So do seasonal variations related to changes or fluctuation in certain seasons. By adding the index of the season it will show the difference between the predicted quarterly and seasonal variations which have the same pattern with the secondary data in the next prediction.

F. Validation Test with Mean Absolut Percentage Definition (MAPD)

MAPD test is a test to measure the accuracy of the estimated value of the prediction results through the application system or manually.

a) $MAPD = \frac{409.868 \times 100\ \%}{71544} = 0.573\%$

100% - 0.573% = 99.427 % is the prediction accuracy value with Exponential Smoothing α = 0.1

b) $MAPD = \frac{100.27 \times 100\ \%}{73558} = 0.136\ \%$

100% - 0.136%= 99.86% is the prediction accuracy value with Exponential Smoothing α = 0.3

c) $MAPD = \frac{304.299 \times 100\ \%}{70761} = 0.43\%$

100% - 0.43% = 99.57 % is the prediction accuracy value with Exponential Smoothing α = 0.5

ACKNOWLEDGMENT

Research by independent fund that aims to design and implement The Application of Prediction of Cocoa Yields with Exponential Smoothing Method can not be done properly without any moral, effort, and material support. We would like to thank Dean of The Faculty of Mathematical and Natural Sciences Dr. Prasetyorini, M.Si and Head of The Computer Science Program Prihastuti Harsani, M.Si who have given moral support, as well as the team members who have been involved so that this research can be resolved.

REFERENCES

[1] F. Ahmad. "Kakao komoditi andalan Indonesia", 2013.
[2] M. Spyros."Metode dan Aplikasi Peramalan", Erlangga, 2007.
[3] P. Subagyo."Forecasting : Konsep dan Aplikasi", BPFE Yogyakarta, 2007.
[4] S. Rahma. "Analisa trend dan forecasting", Modul 8, 2013.
[5] D. Martisunu."Statistik time series", Modul 5, 2012.