DEVELOPING PANEL DATA AND TIME SERIES APPLICATION (DELTA) : SMOOTHING MODULE

Nensi Fitria Deli[1] dan Siti Mariyah[2]

Sekolah Tinggi Ilmu Statistik, Jakarta-Indonesia
e-mail: [1]13.7775@stis.ac.id, [2]sitimariyah@stis.ac.id

Masuk tanggal : 18 Agustus 2017, diterima untuk diterbitkan tanggal : 30 Agustus 2017

Abstrak

Smoothing adalah salah satu metode yang umum digunakan untuk meramalkan data time series. Saat ini sudah banyak aplikasi pengolah data time series yang menyediakan fungsi smoothing, antara lain EViews, Minitab, Zaitun TS, dan R. Namun, aplikasi-aplikasi tersebut masih memiliki kekurangan, seperti tidak tersedianya fasilitas untuk membandingkan beberapa metode smoothing sekaligus. Oleh karena itu, dibangun sebuah aplikasi yang open source yaitu aplikasi DELTA modul smoothing yang menyediakan metode smoothing yang lengkap dan fasilitas untuk membandingkan beberapa metode sekaligus. Berdasarkan uji coba yang telah dilakukan, aplikasi yang dibangun telah sesuai dengan pengguna dan keluaran yang ditampilkan sesuai dengan teori yang dijadikan sebagai acuan.

Kata kunci: smoothing, peramalan, aplikasi time series, aplikasi data panel, eksponensial, rata-rata bergerak

Abstract

Smoothing is commonly used methods to predict time series data. There are many applications that help in the processing of time series data that provide smoothing function such as EViews, Minitab, Zaitun TS, and R. However, these applications have some shortcomings such as the difficulty in comparing several methods. In this study, we build an open source application that provides more complete smoothing method and a facility for comparing several methods, namely smoothing module in DELTA application. Based on the tests, it can be proved that this application is suitable for users and the displayed output is consistent with the theory.

***Keywords**: smoothing, forecasting, time series application, panel data application, exponential, moving average*

INTRODUCTION

Smoothing is one of common method used to predict time series data because it is easy to use. This method can eliminate randomness in the data so that the pattern can be projected into the future and used for forecasting [5]. There are two types of smoothing method, moving average and exponential. The moving average method can be used in relatively constant data. This method consists of a simple moving average, a double moving average, and weighted moving average. The exponential method is commonly used in forecasting for near future. This method consists of a single exponential smoothing, double exponential smoothing with one component (Brown's method), double exponential smoothing with two components (Holt's method), and Holt-Winter exponential smoothing.

Currently, there are many applications that help in time series data processing that provides smoothing function, for instance EViews, Minitab, Zaitun TS, and R. However, these applications still have some shortcomings.

The EViews application is the most widely used application for analysing time series data because this application has a user interface that is easy to use. However, this application is a paid application in case its use is only limited to those who can afford only. In addition, this application only provides some of exponential method and does not provide a moving average method. The Minitab app also has some flaws. This application is also paid and provides just a bit of smoothing methods. On the other hand, there is a free R app. But, the user must type a specific command to process the data. It is not user friendly. In addition, the lack of which is equally owned by the EViews, Minitab, Olives TS, and R is the unavailability of facilities to perform smoothing by several methods at once so users have difficulty in comparing several methods.

A survey about time series data applications was conducted on December 2016. The respondents are STIS students who have earned the course of time series. According to the survey, out of 270 respondents there were more than 30% of respondents stated that quite difficult to do a comparison of several methods of smoothing. Furthermore, there are 72.1% of the total respondents who stated that the need to create a new time series data processing application to overcome the deficiencies contained in existing applications.

This study aimed to build an application that open source namely smoothing module in DELTA application that provides more complete smoothing methods. They are simple moving average, double moving average, weighted moving average, single exponential smoothing, double exponential smoothing with one component (Brown's method), double exponential smoothing with two components (Holt's method) and Holt-Winter exponential smoothing. In addition, this application interface is user friendly and has a facility to perform analysis some smoothing methods at once so that users can compare them easily. This application is built using Python 3.5.2 programming language on data processing algorithm and Qt 5.6.0 on the interface display. This application is expected to be easier for users to perform time series data analysis using smoothing method. The output of this application is also expected to display graphs, forecast results and easier to analyse with available smoothing methods.

RELATED WORKS

Simple moving average

The moving average method is used to predict time series data that has a linear trend [6]. This method averages the overvaluation value in the last n period. In other words, the emergence of new data can be calculated by removing the oldest data and adding new data. The equations used in the implementation of this method are:

$$M_t = \frac{y_t + y_{t-1} + \cdots + y_{t-n+1}}{n} \qquad (1)$$

Forecasting in the next m period use following equation:

$$\hat{Y}_{t+m} = M_t \quad (2)$$

Double moving average

This method is similar to the simple moving average method. However, this method calculates the moving average of simple moving average. According to Montgomery et al [6] the equations which are used in this method are:

$$M'_t = \frac{y_t + y_{t-1} + \cdots + y_{t-n+1}}{n} \quad (3)$$

$$M''_t = \frac{M'_t + M'_{t-1} + \cdots + M'_{t-n+1}}{n} \quad (4)$$

Forecasting in the next m period use following equation:

$$\hat{Y}_{t+m} = 2M'_t - M''_t + \frac{2}{n-1}(M'_t - M''_t)m \quad (5)$$

Weighted moving average

This method is a method that uses a weight for averaging. According Makridakis et al [5], the weights used in this method can be calculated using the weights function as follows:

$$Q(j,m) = \begin{cases} (1-(j/m)^2)^2, & -m < j < m \\ 0, & lainnya \end{cases} \quad (6)$$

With m = (n - 1) / 2. The value of n is the period of moving average. Furthermore, weighing on j (a_j) is calculated using the following equation:

$$a_j = \frac{Q(j,m)}{\sum_{i=-m}^{m} Q(i,m)} \quad (7)$$

The equation that is used to calculate moving averages in this method is:

$$M_t = \sum_{j=-m}^{m} a_j Y_{t+j} \quad (8)$$

The equation which is used to predict data in the next period m is:

$$\hat{Y}_{t+m} = M_t \quad (9)$$

At the end of the data, the observations that are used to calculate the moving average become less. This is because to calculate moving averages on this method requires observation up to the period m+t. To ascertain the amount of weights that are used, an adjustment is made in the divisor used to calculate a_j [5].

Single exponential smoothing

This method is a procedure that repeats calculations continuously using the earliest data. This method is used if the data is not significantly affected by trend and seasonal factors. This method needs a parameter α or commonly called the smoothing constant. The equation which is implemented in this method is:

$$S_t = \alpha Y_t + (1-\alpha)S_{t-1} \quad (10)$$

Meanwhile, the equation used for data forecasting in the next m period is:

$$\hat{Y}_{t+m} = \alpha Y_t + (1-\alpha)\hat{Y}_t \quad (11)$$

Based on the above equation, to calculate *smoothing* the first observations required an initial smoothing value (S_0). According to Abraham and Ledolter [1], the value of S_0 can be calculated using the average of several observations. In this study, the value of S_0 is calculated using the average of the six first observations.

Double exponential smoothing with one component (Brown's method)

This method is used for data containing linear trend. The equations used in the implementation of this method are:

$$S'_n = \alpha Y_n + (1-\alpha){S_{n-1}}' \quad (12)$$

$$S''_n = \alpha {S_n}' + (1-\alpha){S_{n-1}}'' \quad (13)$$

Meanwhile, the equation used for data forecasting in the next m period is:

$$\hat{Y}_{n+m} = \beta_{0,n} + \beta_{1,n} m \quad (14)$$

This equation is based on a linear trend model with:

$$\beta_{0,n} = 2{S_n}' - {S_n}'' \quad (15)$$

$$\beta_{1,n} = \frac{\alpha}{1-\alpha}(S'_n - S''_n) \quad (16)$$

Based on those equations, to calculate *smoothing* the first observations required an initial smoothing value (S_0). The initial value for S_0 is:

$$S_0' = \hat{\beta}_{0,0} - \frac{\alpha}{1-\alpha}\hat{\beta}_{1,0} \tag{17}$$

$$S_0'' = \hat{\beta}_{0,0} - 2\frac{\alpha}{1-\alpha}\hat{\beta}_{1,0} \tag{18}$$

Coefficient in zero period ($\hat{\beta}_{0,0}$ and $\hat{\beta}_{1,0}$) obtained from the linear trend model constant from least square estimator with $\hat{\beta}_{0,0}$ is the intercept and $\hat{\beta}_{1,0}$ is the slope.

Double exponential smoothing with two component (Holt's method)

This method is principally similar with Brown's method. However, this method uses two parameters: α and β. The equations used for the implementation of this method are:

$$S_n = \alpha Y_n + (1-\alpha)(S_{n-1} + T_{n-1}) \tag{19}$$

$$T_n = \beta(S_n - S_{n-1}) + (1-\beta)T_{n-1} \tag{20}$$

Meanwhile the equation used to forecast data in the next m period is:

$$\hat{Y}_{n+m} = S_n + T_n m \tag{21}$$

Based on the above equations, to calculate predicted value of the first observation required an initial smoothing value (S_0 and T_0). These coefficient obtained from the linear trend model from least square estimator with S_0 is the intercept and T_0 is the slope.

Holt-Winter exponential smoothing

The methods mentioned above can be used for non-seasonal data. These methods are less suitable for predicting seasonal data. Holt-Winter exponential smoothing method is what can solve the problem. This method can handle seasonal factors in the data directly [5].

This method uses three smoothing equations, they are for levels, trends, and seasonal. This equation is similar to that used in Holt's method. However, in this method one more equation is added to deal with seasonal factors. This method is divided into two based on its seasonal modeling, that is additive (linear) and multiplicative (not linear).

Additive

Level : $L_t = \alpha(Y_t - S_{t-L})$ (22)

$$+(1-\alpha)(L_{t-1} + b_{t-1})$$

Trend : $b_t = \beta(L_t - L_{t-1})$ (23)

$$+(1-\beta)b_{t-1}$$

Seasonal : $S_t = \gamma(Y_t - L_t)$ (24)

$$+(1-\gamma)S_{t-L}$$

The equation used for forecast in the next m period is:

$$\hat{Y}_{t+m} = L_t + mb_t + S_{t-L+m}, \tag{25}$$

$$m = 1, 2, \dots, L$$

with L is the seasonal length (i.e. months or quarter in a year).

Multiplicative

Level : $\mathrm{L_t} = \alpha\frac{\mathrm{Y_t}}{\mathrm{S_{t\text{-}L}}}$ (26)

$$+(1-\alpha)(L_{t-1} + b_{t-1})$$

Trend : $b_t = \beta(L_t - L_{t-1})$ (27)

$$+(1-\beta)b_{t-1}$$

Seasonal : $\mathrm{S_t} = \gamma\frac{\mathrm{Y_t}}{\mathrm{L_t}} + (1\text{-}\gamma)\mathrm{S_{t\text{-}L}}$ (28)

The equation used for forecast in the next m period is:

$$\hat{Y}_{t+m} = (L_t + mb_t)S_{t-L+m}, \tag{29}$$

$$m = 1, 2, \dots, L$$

Based on above equations, it takes an initial smoothing value for level (L_t), trend (b_t), and seasonal (S_t). To calculate all these three values, it takes at least one complete seasonal. The equation can be used to calculate the initial value of L_t and b_t for additive and multiplicative are [5]:

$$L_s = \frac{1}{s}(Y_1 + Y_2 + \cdots + Y_s) \tag{30}$$

$$b_s = \frac{1}{s}\left(\frac{Y_{s+1} - Y_1}{s} + \frac{Y_{s+2} - Y_2}{s} + \cdots + \frac{Y_{s+s} - Y_s}{s}\right) \tag{31}$$

Meanwhile the equations for calculate the initial value of S_t is different for additive and multiplicative.

Additive:

$$S_1 = Y_1 - L_s, \quad S_2 = Y_2 - L_s, \quad \dots \quad S_s = Y_s - L_s \tag{32}$$

Multiplicative :

$$S_1 = \frac{Y_1}{L_s}, \quad S_2 = \frac{Y_2}{L_s}, \quad \dots \quad S_s = \frac{Y_s}{L_s} \tag{33}$$

Accuracy

User can use accuracy in comparing some smoothing methods. There are several approaches to measure the error. But this application uses mean squared error (MSE) and mean absolute percentage error (MAPE).

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \tag{34}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \times 100\% \tag{35}$$

APPLICATION DEVELOPMENT METHOD

The method used for application development in this research design research method. This method has several steps as follows:

1. Awareness of problem It is the stage of finding, identifying, and analysing the problems in the existing time series data application, especially in the smoothing method. The result of this search stage is the discovery of problems such as the unavailability of open source time series and data panel application provide more complete smoothing methods such as simple moving average, double moving average, weighted moving average, single exponential smoothing, double exponential smoothing with one component (Brown's method), double exponential smoothing with two components (Holt's method) and Holt-Winter exponential smoothing in one application. Another problem is the unavailability of an application that allows users to perform some smoothing methods at once.
2. Suggestion. It is the stage of getting a solution to the problems that have been identified. The solution is an algorithm or procedure in the design of the application. In this study, the result of this phase is smoothing module in DELTA application.
3. Development. It is the stage of implementing the solution by developing smoothing module in DELTA application. In this stage, unified modeling language (UML) is used for application architecture analysis and application modeling.
4. Evaluation. It is the stage of testing and evaluation of the applications that have been built to get feedback if the application is appropriate to solve the problems.
5. Conclusion. This last stage concludes the development smoothing module in DELTA application such as implementation, documentation, evaluation results and shortcomings which contains suggestions for further development.

APPLICATION DESIGN

DELTA application architecture

DELTA application is built modularly. This application consists of 6 modules. They are framework, panel data analysis, trend and decomposition analysis, stationarity, smoothing, and arima. This study takes focus on smoothing module. DELTA application architecture can be seen in Figure 1.

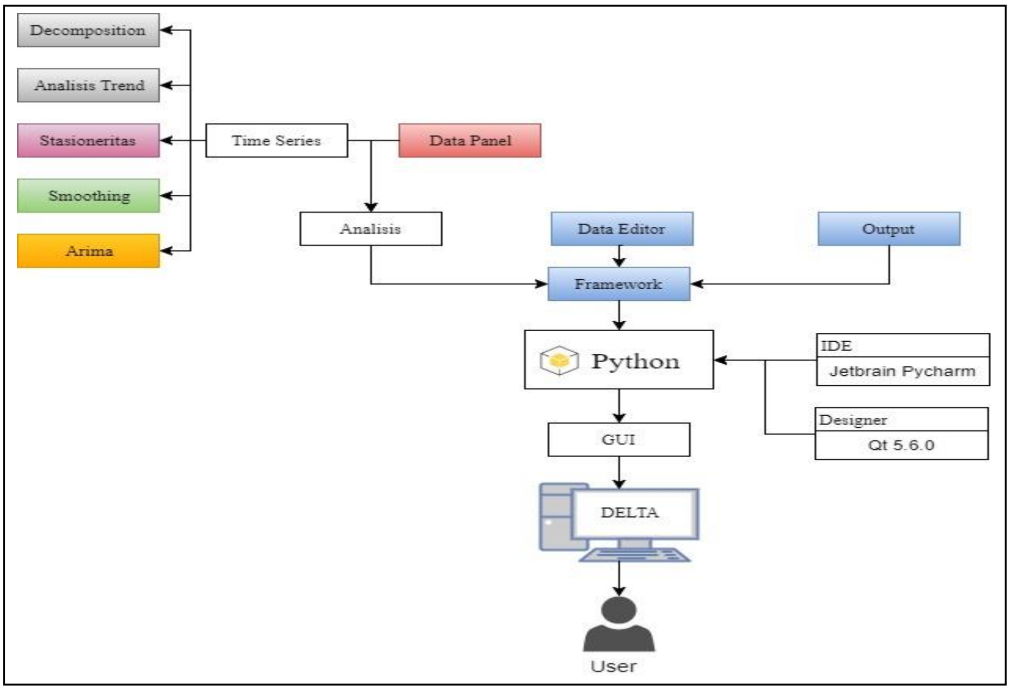

Fig. 1. DELTA application architecture

Smoothing module design

DELTA application is an object-oriented application. The design of this application uses the modeling for object-oriented applications, namely Unified Modeling Language (UML). UML visualizes, specifies, builds, and documents an object-oriented application development system. There are several diagrams in UML. Use case diagram and activity diagram are used in this study.

Use case diagram explains the interaction between the system and the user. This diagram illustrates who will use the app and in what way the user interacts with the app graphically. Use case diagram used in the design of this application can be seen in Figure 2.

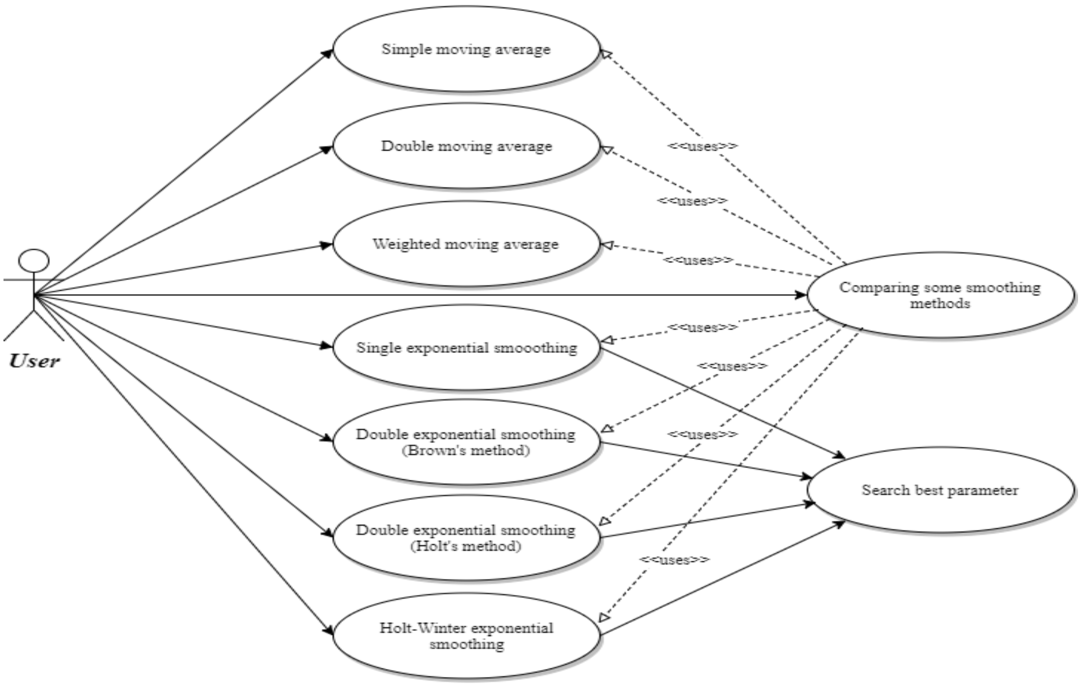

Fig. 2. Use case diagram of smoothing module in DELTA application

In Figure 1 there is one actor and nine use case. Actor interacting with this application is user. Meanwhile, use cases contained in this application are:

1. Simple moving average, which is a function to calculate forecasting using simple moving average method.
2. Double moving average, which is a function to calculate forecasting using double moving average method.
3. Weighted moving average, which is a function to calculate forecasting using weighted moving average method.
4. Single exponential smoothing, which is a function to calculate forecasting using single exponential smoothing method.
5. Double exponential smoothing (Brown's method), which is a function to calculate forecasting using double exponential smoothing method (Brown's method).
6. Double exponential smoothing (Holt's method), which is a function to calculate forecasting using double exponential smoothing method (Holt's method).
7. Holt-Winter exponential smoothing, which is a function to calculate forecasting using Holt-Winter exponential smoothing method.
8. Comparing of several smoothing methods is a function to calculate forecasting using several methods at once and displays the comparison of methods based on predefined comparison criteria (MSE and MAPE).
9. Search best parameter is the function to find the best parameters for all exponential smoothing methods. This function uses the grid search algorithm in search of the best parameter value.

Activity diagrams can be used to describe the steps of a use case. The activity diagram of each use case in this module can be seen in Figure 3-11.

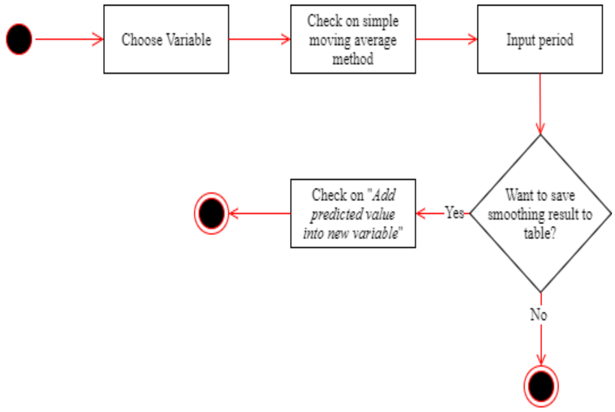

Fig. 3. Activity diagram of simple moving average use case

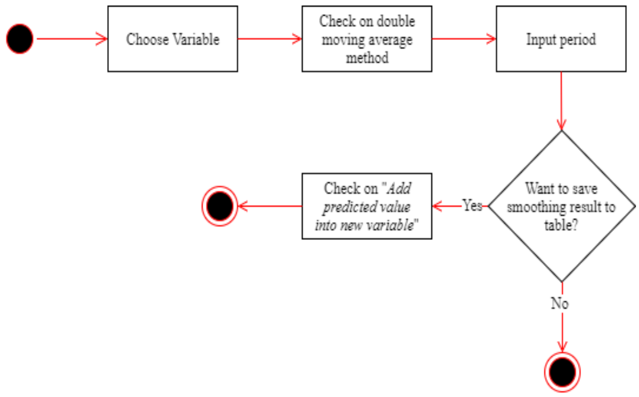

Fig. 4. Activity diagram of double moving average use case

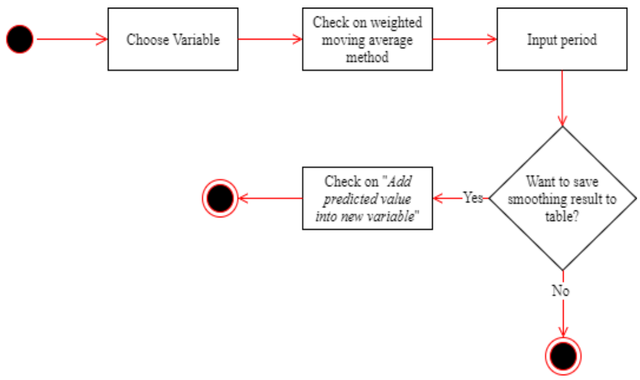

Fig. 5. Activity diagram of weighted moving average use case

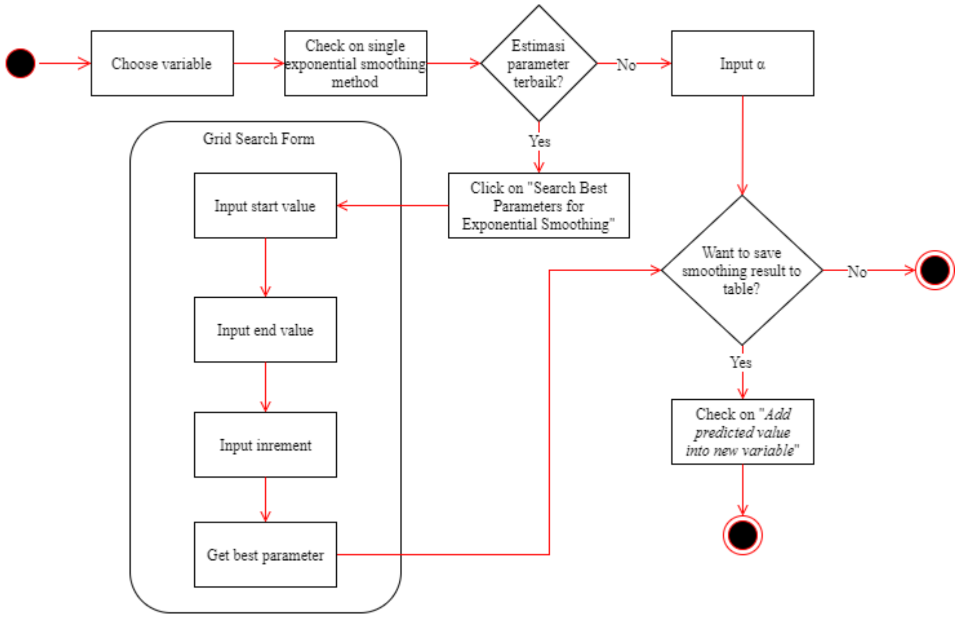

Fig. 6. Activity diagram of single exponential smoothing use case

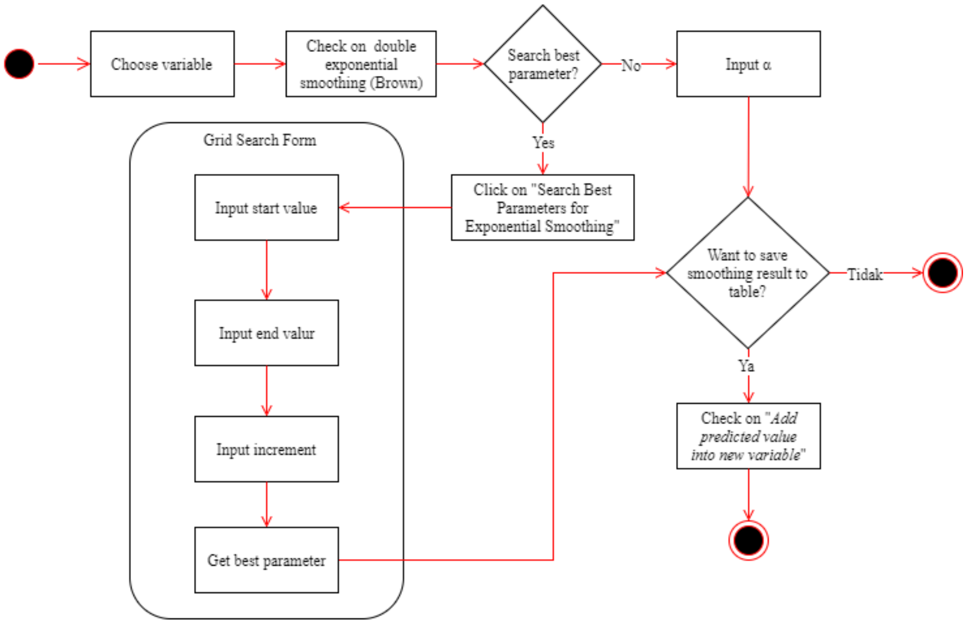

Fig. 7. Activity diagram of double exponential smoothing (brown's method) use case

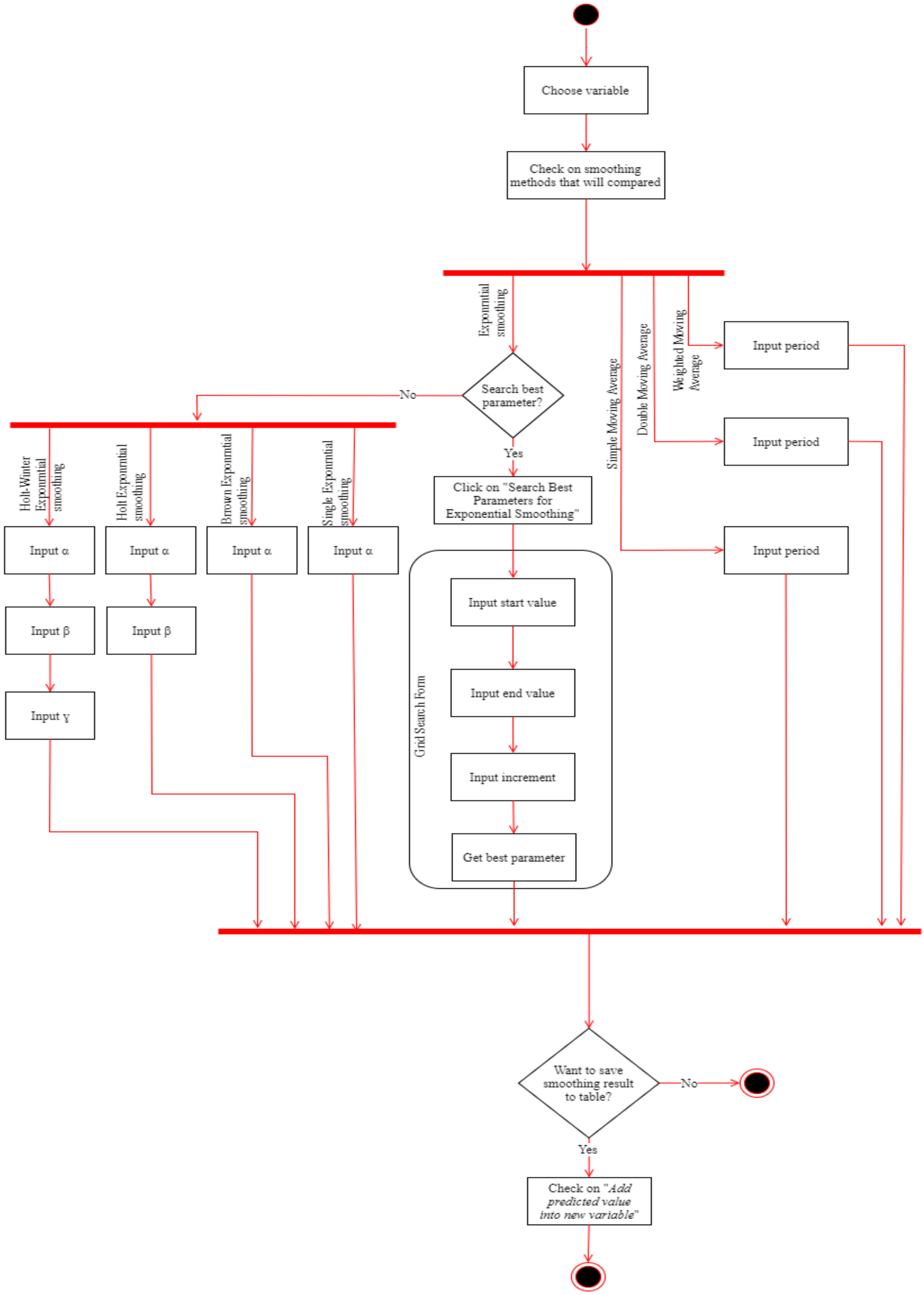

Fig. 8. Activity diagram of comparing some smoothing methods use case

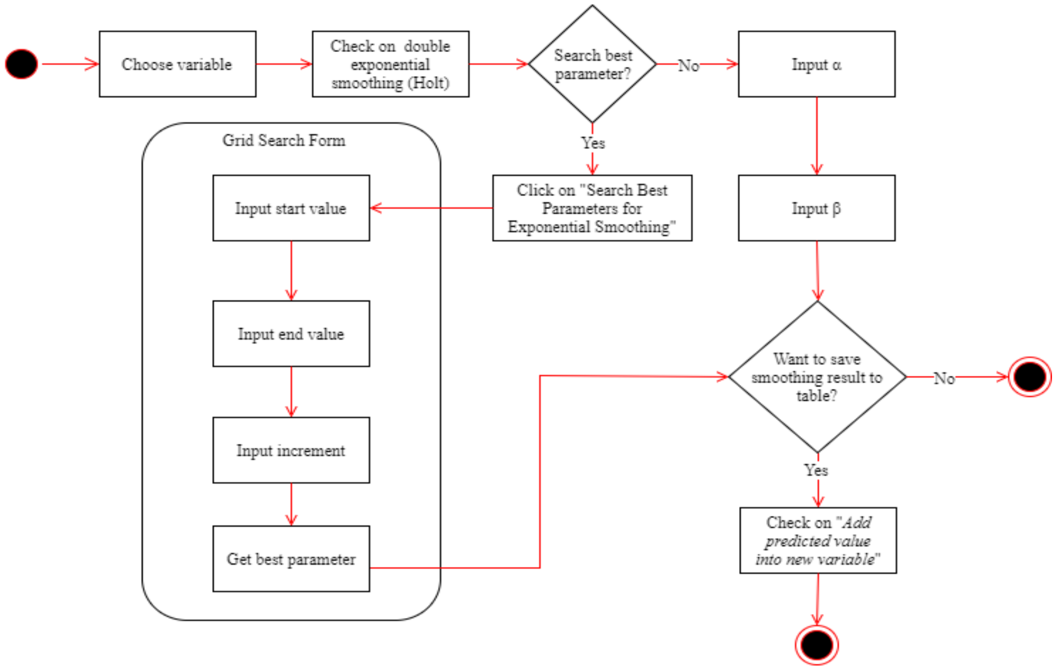

Fig. 9. Activity diagram of double exponential smoothing (holt's method) use case

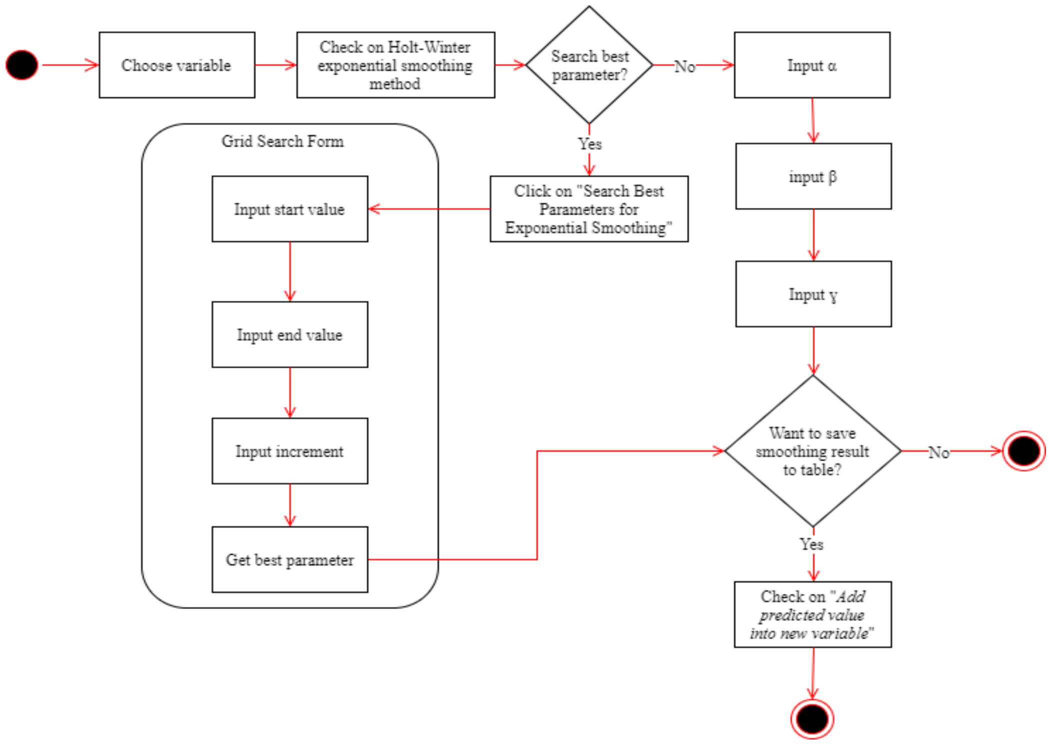

Fig. 10. Activity diagram of Holt-Winter exponential smoothing use case

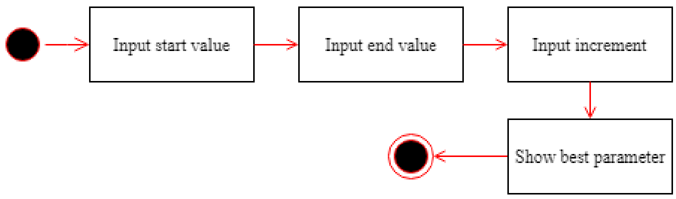

Fig. 11. Activity diagram of search best parameter use case

Application interface design

DELTA application interface design, smoothing analysis interface design, grid search interface design, and output design can be seen in Figure 12-15.

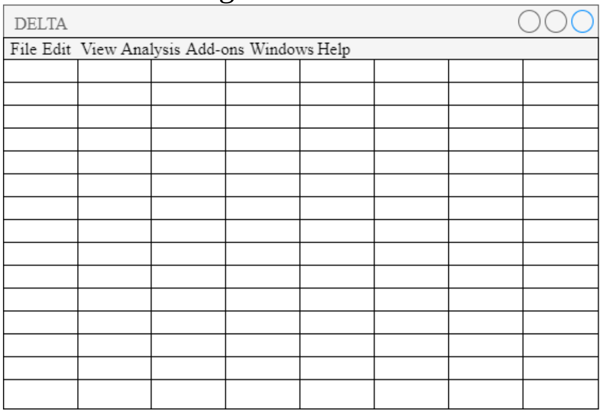

Fig. 12. DELTA application interface design

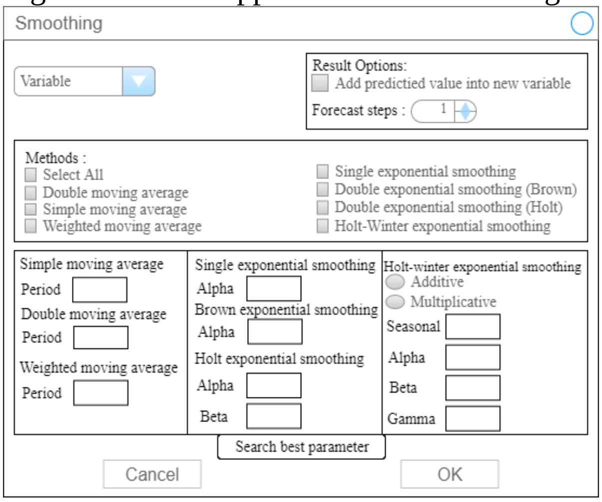

Fig. 13. Smoothing analysis interface design

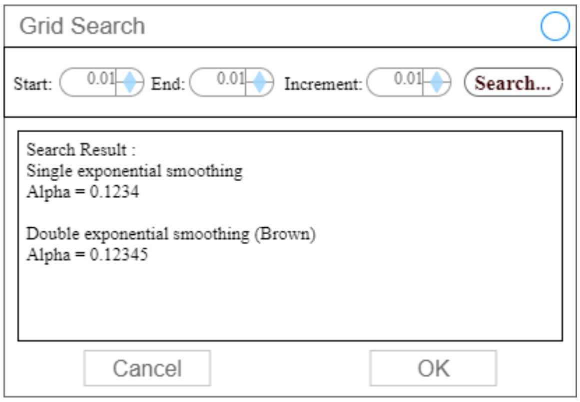

Fig. 14. Grid search interface design

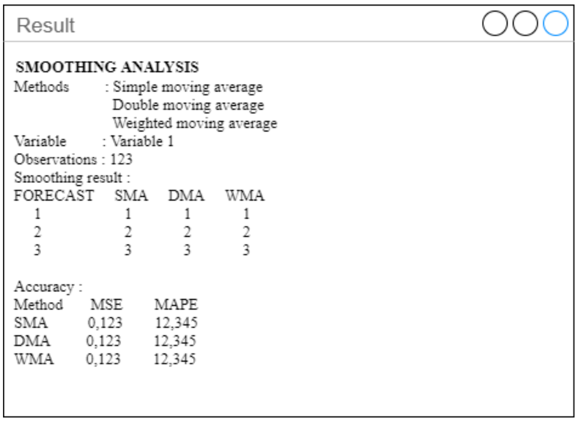

Fig. 15. Output design

IMPLEMENTATION

Implementation of DELTA application's main page interface

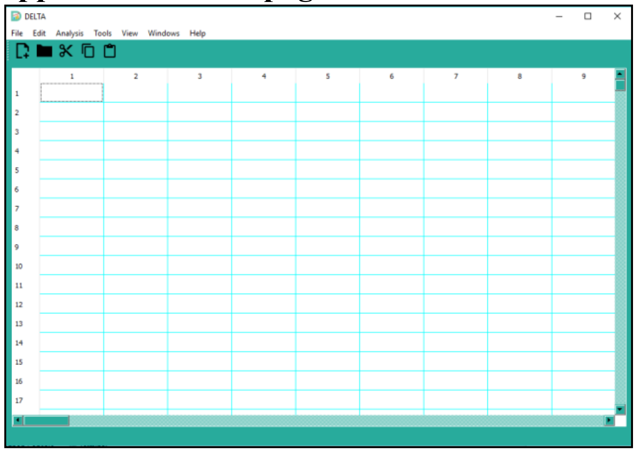

Fig. 16. DELTA application's main page interface

The main page of the application user interface is can seen in Figure 16. Users can input the data from a file with Excel (.xls or .xlsx) type and comma delimited (.csv) type. Another way to input the data is input data in the table by typing directly in the application.

Implementation of smoothing window's interface

Implementation of smoothing window's interface can be seen in Figure 17.

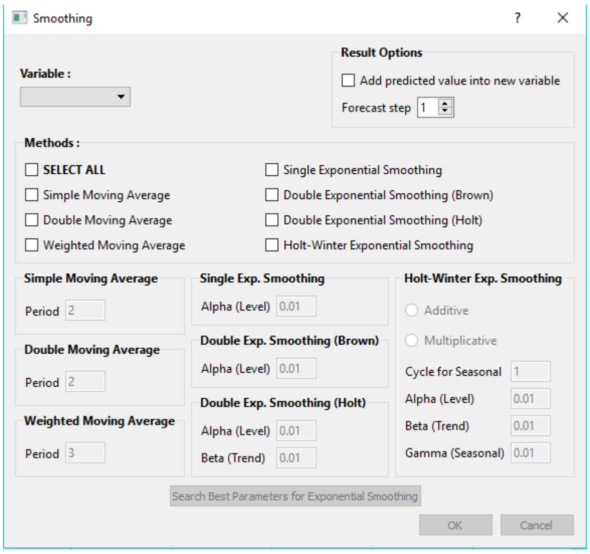

Fig. 17. Smoothing window interface

Implementation of grid search's interface

Grid search window's interface can be seen in Figure 18. This window appears when the user presses the "best search parameter for exponential smoothing" button in smoothing window.

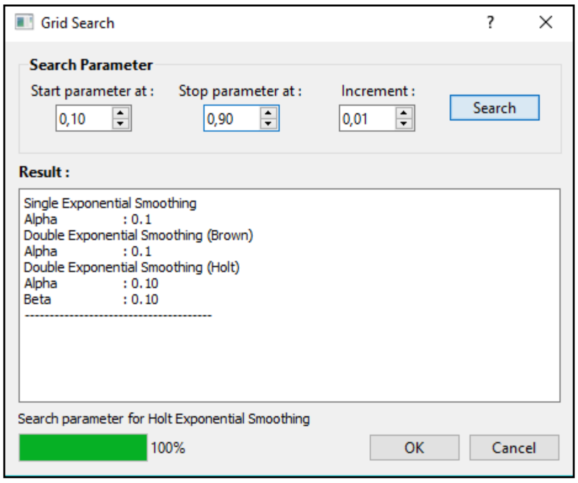

Fig. 18. Grid search window interface

Implementation of the output page's interface

The output interface can be seen in Figure 19. Once the user presses the "OK" button on the smoothing window, the application processing forecast calculation using smoothing methods that have been selected with the parameters entered by the user and then the result is displayed on the output window.

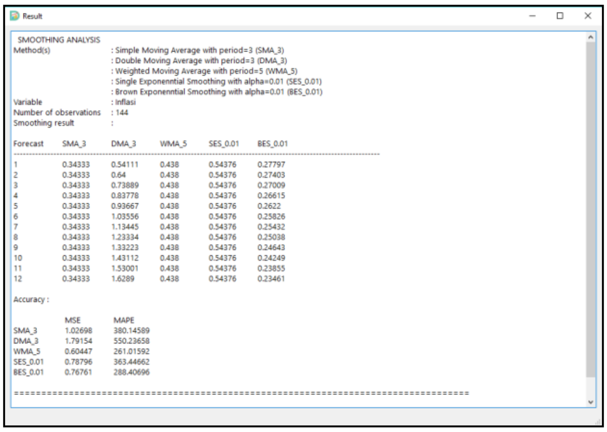

Fig. 19. Output window interface

TESTING

There are four approaches in testing this module. They are white-box testing, black-box testing, validation testing, and system usability scale (SUS).

White-Box Testing

White-box testing tested the internal performance of the application. Testing on this application is done by running the application in PyCharm Community Edition 2017.1 for all functions that have been made. In this test is done unit testing and integration testing. Based on test results, all classes (units) contained in this smoothing module can run well and well integrated in DELTA application project.

Black-box Testing

Black-box testing is a test that focuses on testing the application interface. This test is done by checking and showing the function of the application can be operated and check the error on the interface display. Black-box testing aims to test a specific function in the application without the need to know the internal performance of the application. These test results prove that all the functions on this smoothing module has been running well.

Validation

What is tested in this test is the input and output of the application. With two different inputs given, the application is tested to see if the application has provided output according to the applied theory.

In testing the results of each method analysis, comparative applications used are Zaitun TS and Microsoft Excel. Testing results can be seen in Table 1.

Table 1. Validation Test Result For All Of Analysis Methods

Smoothing Method	Result
Simple moving average	√
Double moving average	√
Weighted moving average	√
Single exponential smoothing	√
Double exponential smoothing with one component (*Brown's method)*	√
Double exponential smoothing with two component (*Holt's method*)	√
Holt-Winter exponential smoothing (additive)	√
Holt-Winter exponential smoothing (multiplicative)	√

Based on Table 1, it can be seen that the output of all of the methods provided in the application DELTA smoothing module in accordance with the theory implemented. The best parameter search results are also validated with the Zaitun TS application. Testing result can be seen in Table 2.

Table 2. Validation Result For Best Parameter Search

Smoothing Method	Result
Simple moving average	√
Double moving average	√
Weighted moving average	√
Single exponential smoothing	√
Double exponential smoothing with one component (*Brown's method)*	√
Double exponential smoothing with two component (*Holt's method*)	√
Holt-Winter exponential smoothing (additive)	X
Holt-Winter exponential smoothing (multiplicative)	X

Based on Table 2, it can be seen that the best result from smoothing module in DELTA application is accordance with Zaitun TS. However, there are differences in the results of the method of Holt-Winter exponential smoothing in both additive and multiplicative approach. This is due to the initialization smoothing approach applied to this module is different from the Zaitun TS application thus smoothing results given were different.

System Usability Scale

System Usability Scale (SUS) is used for testing the application on the user side. Testing was conducted on July 11, 2017 with ten respondents.

Table 3. SUS Score For Smoothing Module In Delta Application

Respondent	Score
1	77,5
2	90
3	85
4	75
5	87,5
6	95
7	77,5
8	62,5
9	75
10	70
Average	79,5

According Usability.gov [12], minimum score to determine that a application is good from the user side is 68. Based on Table 3, this module obtained a score of 79.5. Therefore, it can be proved that smoothing module in DELTA application is good module from the user side aspect.

RESULTS

Based on the experiments that have been done, the application output built in this study has the same results with compared application on each method. The final SUS score earned is well above the minimum limit. This results proved that the built application has been in accordance with the user and the output is displayed also in accordance with the theory and formulas used as a reference. However, there are differences in search best parameter function for Holt-Winter Exponential Smoothing. The result of the search for this smoothing method in this module is different with Zaitun TS application. This difference is caused by the approach used for smoothing initialization for this method in this application is different from the Zaitun TS application.

CONCLUSIONS

A smoothing module in an open source application for processing time series data has developed. This module provides more complete smoothing methods, such as simple moving average, double moving average, weighted moving average, single exponential smoothing, double exponential smoothing with one component (Brown's method), double exponential smoothing with two components (Holt's method), and Holt-Winter exponential smoothing. This module also provides the facility that allows users to process some smoothing methods at once and search the best parameters for exponential smoothing. In addition, the results of this application module also been compared with comparative applications and have similar results with the compared application.

REFERENCES

[1] J. L. Whitten and L. D. Bentley, System Analysis & Design Methods, New York: McGraw-Hill, 2007.

[2] M. A. Weisfeld, The Object-Oriented Thought Process, Third Edition, USA: Addison-Wesley Professional, 2009.

[3] R. S. Pressman, Software Engineering : A Practicioner's Approach, Seventh Edition, New York: McGraw-Hill, 2010.

[4] J. P. Mueller, Beginning Programming with Python For Dummies, New Jersey: John Wiley and Sons, 2014.

[5] D. C. Montgomery, L. A. Johnson and J. S. Gardiner, Forecasting & Time Series Analysis 2nd Edition, Singapore: McGraw-Hill, 1990.

[6] S. G. Makridakis, S. C. Wheelwright and R. J. Hyndman, Forecasting : Methods and Applications, Chichester: John Wiley and Sons, 1997.

[7] B. Abraham and J. Ledolter, Statistical Methods for Forecasting, New Jersey: Wiley-Interscience, 1983.

[8] Usability.gov, "System Usability Scale (SUS)," [Online]. Available: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html. [Accessed 6 7 2017].

[9] Python Software Foundation, "Python," [Online]. Available: https://www.python.org. [Accessed 10 Januari 2017].

[10] P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, New York: Springer, 2002.

[11] G. E. Bo, G. M. Jenkins and G. C. Reinsel, Time Series Analysis Forecasting and Control 3rd Edition, USA: Prentice-Hall, 1994.

[12] B. Juanda and J. , Ekonometrika Deret Waktu, Bogor: IPB Press, 2012.

[13] A. M. Sholihah, "Perancangan dan Implementasi Program Aplikasi Time Series dengan Metode Pemulusan dan Dekomposisi Untuk Peramalan," Jakarta, 2007.

[14] N. E. Oktafiani and F. Andriyani, "Aplikasi Pemulusan Eksponensial dari Brown dan dari Holt Untuk Data Yang Memuat Trend," in *Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III*, 2012.

[15] Yuniastari, N. L. A. Kartika and I. W. W. Wirawan, "Peramalan Permintaan Produk Perak Menggunakan Metode Simple Moving Average dan Exponential Smoothing," *Jurnal Sistem dan Informatika Vol.9,* pp. 97-106, 2014.