Forecasting Unemployment in North Sumatra Using Double Exponential Smoothing Method

Havid Syafwan1*, Mahdhivan Syafwan2, Elvathna Syafwan3, Aulia Fitrul Hadi4, Pristiyanilicia Putri5

1Department of Informatics Management, Sekolah Tinggi Manajemen Informatika dan Komputer Royal Kisaran, Indonesia.
2Department of Mathematics, Faculty of Mathematics and Natural Sciences, Andalas University, Indonesia.
3Department of Mechanical Engineering, Politeknik TEDC Bandung, Indonesia.
4Department of Computer Science, Faculty of Computer Science, Universitas Putra Indonesia YPTK Padang, Indonesia.
5Department of Information Systems, Sekolah Tinggi Manajemen Informatika dan Komputer Royal Kisaran, Indonesia.

*havid_syafwan@yahoo.com

Abstract. The research aims to predict the unemployment in the province of North Sumatra in 2020 using the Double Exponential Smoothing (DES) method. The data used is derived from the Central Agency Statistik (BPS) of North Sumatra province where the actual data is taken within 20 years from 2000 to 2019. The accuracy method in this research uses MAD to count the number of errors, MSE to evaluate forecasting methods, and MAPE to calculate the percentage of errors. Results of this research in the form of forecasting the number of unemployment in North Sumatra in 2020 that is 381459 people in the value of alpha 0.6 with a MAD value of 77402.12, MSE value of 12524690448.31, and MAPE value of 16.35%.

1. Introduction

Many of the things we find in science are especially about the problems that one of them can be classified into forecasting. Forecasting is a means of doing effective planning to predict future uncertainty to be able to make a better decision. In its application, forecasting is used in areas such as finance and banking, social, economic, production, marketing, population, education, and many others[1]. In calculating this forecasting many methods are used, one of them using an exponential smoothing method. The exponential smoothing are categorized into time series methods that use data weighted in the past exponentially. There are three categories used in this method: Single, double, and triple exponential smoothing method[2]

Brown's double exponential method is more popular to use for economic research that contains a trend and can be used with relatively moderate and easy data processing in its data where it uses actual data from previous years. Because this method is popular used in economics, then the data used in this research is data on the number of unemployment[3].

The economic problems that directly affects the standard of living of people in developed and developing countries is the problem of unemployment that become the most important part of economic development characterized by the volume of economic growth in a country that has not been resolved by its government, especially the government in these areas[4]

North Sumatera is one of the areas in Sumatra Island, Indonesia. Judging from the area of North Sumatera and the amount of unemployment is relatively much, the forecasting to know the amount of
unemployment for the time will be very influential to the government's policy in making the right decision to strengthen the problem in the future[5].

Therefore the authors are very interested in conducting research related to forecasting the number of unemployment in North Sumatra province in 2020 using a double exponential smoothing method where by the data from the Central Statistics Agency (BPS) North Sumatera Province from 2000 to 2019. This forecasting accuracy method uses MAD to count the number of errors, MSE to evaluate forecasting methods, and MAPE to calculate the error percentage.

2. Methodology

Forecasting is an attempt to predict future conditions through past state testing. The essence of forecasting is the approximate future events based on patterns in the past and the usability of policy on projections with patterns in the past[6].

Double Exponential Smoothing Method (DES)

DES method is a linear model that uses a Smoothing process twice. The rationale of Brown's linear exponential smoothing method is similar to that of a linear moving average, as both single and double-smoothing values lag out of actual data if there are trend elements[3].

The steps used in applying Brown's DES method can be seen in the following equations[6]:

1. Calculate the first exponential smoothing value with symbol \( S'_t \)
   \[
   s'_t = \alpha x_t + (1 - \alpha) s'_{t-1}
   \] (1)

2. Calculates the second exponential smoothing value with symbol \( S''_t \)
   \[
   s''_t = \alpha s'_t + (1 - \alpha) s''_{t-1}
   \] (2)

3. calculate the value of the constant with the symbol \( \alpha_t \)
   \[
   \alpha_t = 2s'_t - s''_t
   \] (3)

4. Determining slope with symbols \( b_t \)
   \[
   b_t = \frac{\alpha}{1-\alpha} (s'_t - s''_t)
   \] (4)

5. Value of the forecasting with symbol \( F_{t+m} \)
   \[
   F_{t+m} = \alpha_t + b_t (m)
   \] (5)

Where :
- \( s'_t \) : Single Exponential Smoothing value
- \( s''_t \) : Double Exponential Smoothing value
- \( \alpha_t \) : size of \( t \) period constants
- \( b_t \) : Slope or trend value of the corresponding data
- \( F_{t+m} \) : Forecasting value for next period
- \( m \) : the period of forecasting
- \( x_t \) : Actual value
- \( \alpha \) : Exponential smoothing parameters of magnitude \( 0 < \alpha < 1 \)

Mean Absolute Deviation (MAD)

MAD method is used to evaluate forecasting using the sum of the absolute errors. This method is done by measuring the forecasted accuracy with an average suspected error. The formula for calculating the MAD value can be seen in the following equations.

\[
MAD = \frac{\sum_{t=1}^{n}|x_t - F_t|}{n}
\] (6)
### Mean Squared Error (MSE)

MSE is a method for evaluating forecasting methods. Each mistake or remnant is used. This approach regulates the wrongdoing of the great forecasting of those mistakes. That method generates moderate mistakes that are likely to be better for small mistakes but sometimes make a big difference. The Mean Squared Error is the mean of an unquoted forecast error. The formula for calculating the MSE values can be seen in the following equations:

\[
MSE = \frac{\sum_{t=1}^{n} (x_t - F_t)^2}{n}
\]  

Where eq (6), (7), and (8) is

- \(x_t\) = data actual in the \(t\) period
- \(F_t\) = Forecasting value in \(t\) period
- \(n\) = Number of data
- \(t\) = Forecasting period

### Mean Absolute Percentage Error (MAPE)

MAPE is the average percentage of absolute errors calculated by looking for an absolute value error in each period, then divided by the actual observation value and absolute percentage error. The formula for calculating the MAPE value can be seen in the following equation:

\[
MAPE = \frac{\sum_{t=1}^{n} |x_t - F_t|}{\sum_{t=1}^{n} x_t} \times 100\% 
\]  

Where MAPE is the average percentage of absolute errors calculated by looking for an absolute value error in each period, then divided by the actual observation value and absolute percentage error. The formula for calculating the MAPE value can be seen in the following equation:

The ability of forecasting is very good if it has a MAPE value of less than 10% and has good forecasting ability if MAPE value is less than 20%[3].

### Table 1. MAPE Value For Evaluation Forecasting

| MAPE Value | Accuracy |
|------------|----------|
| MAPE \(\leq 10\%\) | High |
| \(10\% < MAPE \leq 20\%\) | Good |
| \(20\% < MAPE \leq 50\%\) | Reasonable |
| MAPE \(\geq 50\%\) | Low |

Forecasting accuracy will be higher when the values of MAD, MSE, and MAPE are increasingly small. The necessity to compare the calculations that have the least MAD value, as the smaller the value of MAD, means the smaller the difference between forecasting results and actual value [6].

### 3. Result and Discussion

The data used for forecasting of unemployment in the province of North Sumatra in 2020 is to use unemployment data from 2000 to 2019 (for 20 years) in North Sumatra. Table 2 Below is the historical data on the number of unemployment in North Sumatra for 20 years at the Central Statistic Agency (BPS) of North Sumatra Province.

### Table 2. Unemployment data in North Sumatra from 2000-2019

| Periode | Years | Number of Unemployment | Periode | Years | Number of Unemployment |
|---------|-------|------------------------|---------|-------|------------------------|
| 1       | 2000  | 335729                 | 11      | 2010  | 491809                 |
| 2       | 2001  | 229212                 | 12      | 2011  | 402125                 |
| 3       | 2002  | 355504                 | 13      | 2012  | 379870                 |
| 4       | 2003  | 404117                 | 14      | 2013  | 412202                 |
From table 2, the unemployment data in North Sumatra Province from 2000 to 2019 can be made in the form of charts below.

The problem when using the DES method is determining the size of $\alpha$. The rule is to choose the value of $\alpha$ in such a way that MAD, MSE, and MAPE are of the least value. However, in this research, the calculation process using the DES method uses the help of facilities available in Microsoft Excel, a very complicated calculation becomes simple when looking for the best $\alpha$ value[5].

In the calculation process using DES method, the magnitude of the alpha value used is 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8 and 0.9. On the DES method when $t = 1$ (first year), forecasting is unavailable so the $s'_1$ value (the first forecasting) equals the $x_1$ value (the actual data of 2000). The following manual calculations of unemployment forecasting in North Sumatra with an alpha value ($\alpha$) of 0.6.

1. $s'_1 = 335729$
   $s''_2 = (0.6).229212 + (1 - 0.6).335729 = 271818,80$
   $s'_3 = (0.6).355504 + (1 - 0.6).271818,80 = 322029,92$
   ...
   $s'_{20} = (0.6).382438 + (1 - 0.6).390511,61 = 385667,44$

2. $s''_{1} = 335729$
   $s''_{2} = (0.6).271818,80 + (1 - 0.6).335729 = 297382,88$
   $s''_{3} = (0.6).322029,92 + (1 - 0.6).297382,88 = 312171,10$
   ...
   $s''_{20} = (0.6).385667,44 + (1 - 0.6).389876,33 = 387351,00$

3. $\alpha_1 = 2s'_{1} - s''_{1} = 2.(335729) - 335729 = 335729$
   $\alpha_2 = 2s'_{2} - s''_{2} = 2.(271818,80) - 297382,88 = 246254,72$
   $\alpha_3 = 2s'_{3} - s''_{3} = 2.(322029,92) - 312171,10 = 331888,74$
   ...

Figure 1. Graph Unemployment in North Sumatra from 2000-2019
\[ a_{20} = 2s'_{20} - s''_{20} = 2 \cdot 385667.44 - 387351.00 = 383983.89 \]

\[ (4) \ b_1 = \frac{0.6}{1 - 0.6} (s'_{1} - s''_{1}) = \frac{0.6}{1 - 0.6} (335729 - 335729) = 0 \]

\[ b_2 = \frac{0.6}{1 - 0.6} (s'_{2} - s''_{2}) = \frac{0.6}{1 - 0.6} (271818.80 - 297382.88) = -38346.12 \]

\[ b_3 = \frac{0.6}{1 - 0.6} (s'_{3} - s''_{3}) = \frac{0.6}{1 - 0.6} (322029.92 - 312171.10) = 14788.22 \]

\[ ... \]

\[ b_{20} = \frac{0.6}{1 - 0.6} (s'_{20} - s''_{20}) = \frac{0.6}{1 - 0.6} (385667.44 - 387351.00) = -2525.33 \]

\[ (5) F_2 = a_1 + b_1 (1) = 335729 + 0 = 335729 \]

\[ F_3 = a_2 + b_2 (1) = 246254.72 + (-38346.12) = 207908.60 \]

\[ F_4 = a_3 + b_3 (1) = 331888.74 + 14788.22 = 346676.96 \]

\[ ... \]

\[ F_{21} = a_{20} + b_{20} (1) = 383983.89 + (-2525.33) = 381458.55 \]

The calculation process using the DES method is repeated for each of the alpha values done in the unemployment data in North Sumatra. The results of unemployment forecasting with \( \alpha \) value = 0.6 using the DES method seen in the following table

**Table 3. Unemployment Forecasting Results with Alpha Value of 0.6**

| Years   | Number of Unemployment | Years   | Number of Unemployment |
|---------|------------------------|---------|------------------------|
| 2001    | 335729,00              | 2011    | 467511,12              |
| 2002    | 207908,60              | 2012    | 361441,47              |
| 2003    | 346676,96              | 2013    | 332410,40              |
| 2004    | 430393,23              | 2014    | 383649,28              |
| 2005    | 859098,39              | 2015    | 376338,48              |
| 2006    | 745994,52              | 2016    | 426041,62              |
| 2007    | 682735,47              | 2017    | 366448,18              |
| 2008    | 581508,89              | 2018    | 365526,28              |
| 2009    | 533814,48              | 2019    | 392099,80              |
| 2010    | 512486,64              | 2020    | 381458,55              |

After all the results of the forecasting on each alpha is obtained, the process of evaluation calculation is carried out using three accuracy methods, MAD, MSE, and MAPE. This evaluation was performed at each of the alpha value of 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8 and 0.9. Thus, from the calculation results of the three accuracy methods obtained the error value in the following table 4

**Table 4. MAD, MSE, and MAPE Value**

| Alpha value | MAD        | MSE        | MAPE       |
|-------------|------------|------------|------------|
| 0.1         | 103786,47  | 18923326969,77 | 22,35%     |
| 0.2         | 88865,37   | 15623309260,76 | 19,03%     |
| 0.3         | 80022,83   | 13616785348,01 | 17,06%     |
| 0.4         | 80789,18   | 12565206389,20 | 17,07%     |
| 0.5         | 78870,36   | 12525274580,74 | 16,54%     |
| 0.6         | 77402,12   | 12524690448,31 | 16,35%     |
In table 4 can be seen that the best forecasting for the unemployment data is at an alpha value of 0.6 which has a MAD value of 77402.12, the MSE value of 12524690448.31, and the MAPE value of 16.35% which is the forecasting of unemployment in North Sumatra in 2020 is 381458.55 \approx 381459 people.

4. Conclusion
From the discussion, it can be concluded that this unemployment forecasting uses the Double Exponential Smoothing method with forecasting accuracy method namely Mean Absolute Deviation (MAD) as the method of calculating error number, Mean Squared Error (MSE) to evaluate the forecasting method, and Mean Absolute Percentage The results of this research in the form of forecasting the number of unemployment in North Sumatra in 2020 by 381459 people at the alpha value of 0.6 with a MAD value of 77402.12, MSE value of 12524690448.31, and MAPE value of 16.35%.

References
[1] E. Munarsih, “Peramalan Jumlah Pengangguran di Provinsi Sumatera Selatan dengan Metode Autoregressive Integreted Moving Average (ARIMA),” J. Penelit. Sains, vol. 19, no. 1, p. 168439, 2017.
[2] S. Makridakis et al., “The accuracy of extrapolation (time series) methods: Results of a forecasting competition,” J. Forecast., vol. 1, no. 2, pp. 111–153, 1982, doi: 10.1002/for.3980010202.
[3] Z. Arifin, “Peramalan Pengangguran Menggunakan Metode Double Exponential Smoothing Di Provinsi Kalimantan Timur,” Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf., vol. 4, no. 1, pp. 24–29, 2019.
[4] K. Dumi, Č. Č. Anita, and B. Žmuk, “Forecasting Unemployment Rate in Selected European Countries Using Smoothing Methods,” Int. J. Soc. Behav. Educ. Econ. Bus. Ind. Eng., vol. 9, no. 4, pp. 1073–1078, 2015.
[5] “An Initial Study on the Forecast Model for Unemployment Rate Mohd Nadzri Mohd Nasir, Kon Mee Hwa and Huzaifah Mohammad 1.”
[6] I. F. Talia, “Peramalan Tingkat Kemiskinan Penduduk Provinsi Kalimantan Timur Menggunakan Metode Double Exponential Smoothing,” vol. 4, no. 2, pp. 121–127, 2019.