Double Exponential Smoothing Forecasting Food Crop Yields Using Geographic Information Systems

Dovel Pirmanto

Islamic Library and Information Science Study Program, Faculty of Adab and Dakwah

Institut Agama Islam Negeri Kerinci

Kerinci, Indonesia

dovelpirmanto16@gmail.com

**Abstract— Food is a source of basic needs for every living creature, so food security is an interesting issue for every country. This causes problems regarding food and land use, especially in the City of Sungai Penuh. Food problems arise as a result of the lack of information regarding the appropriate use of land and the productivity of the land itself. In the current industrial era 4.0, forecasting can be done using information technology tools that provide convenience and efficiency in forecasting time and can be integrated with geographic information systems. Forecasts made by the community based on past experience without considering factors that influence crop yields, so that they can cause losses in terms of time and costs. Forecasting is part of a data science process that predicts future events based on past data mathematically. There are many methods for forecasting future events, one of which is Double Exponential Smoothing. Results of data analysis using the Double Exponential Smoothing method by considering error values with α = 0.1 and 0.5 and β = 0.1 and 0.5. calculation of the smallest error value, namely: ME = 80.92, MAD = 5.58, MAPE = 11%, MSE = 52.69 by combining the value of α= 0.1 and the value of β = 0.1 to produce a prediction of the corn harvest in Kumun Debai District in 2024 is 45 tons and in 2025 amounting to 40 tons.**

***Keywords—*** ***Double Exponential Smoothing, Forecasting, Crops, Geographic Information Systems.***

#  INTRODUCTION

Food is a source of basic needs for humans which continues to increase from year to year. So this becomes an important challenge issue for every country, because the human population reaches nine billion people. The increasing need for food in Sungai Penuh City has given rise to problems related to land use. This problem arises as a result of the lack of information relating to the level of land productivity and the suitability of land as food land. Currently, forecasting can use information technology for easy data processing with the help of automatic systems. Before the development of technology related to forecasting, forecasting was still done manually which did not look at previous data or was more qualitative but not quantitative. [1].

Forecasting is a branch of science that deals with possible future events using mathematical calculations involving historical data, various forecasting methods provide a structured way of analyzing data so that data analysis becomes better. The use of these data analysis techniques is expected to provide analysis results that are more accurate by testing deviations from the stored analysis data [3]. One method for making forecasts is double exponential smoothing (DSE).

Technological developments really help people in obtaining information, one of the information needed is geographic information. The public can use this geographic information for various fields such as agricultural distribution, regional planning, research, mapping and other purposes. Geographic information is able to present data processing results that are integrated with spatial and non-spatial data. Various previous studies have developed and utilized Geographic Information Systems to support decision making in agricultural land mapping which shows that the results are good, this is supported by integration between GIS and decision support systems [9]. Sustainable use of Geographic Information Systems in the analysis of agricultural land use data that combines biophysical aspects, economic impacts and social impacts that arise in the northern Sinai region [10].

Based on the problems described above, it is very necessary to predict crop yields in order to minimize the impact of losses in terms of time and costs if the wrong prediction of crop yields is detrimental to farmers. So it is necessary to forecast agricultural yields to determine the productivity of land use in Sungai Penuh City, especially food crops. The method used in this research is double exponential smoothing, namely the application of the double exponential smoothing method to predict food crop harvests that are integrated with a geographic information system. Based on previous research, the use of the double exponential smoothing method is very appropriate for predicting annual data and this method takes into account various factors that influence the prediction results. So the results of calculations using double exponential smoothing can be mapped using GIS.

# RESEARCH METHODOLOGY

## Forecasting

Forecasting possible future events, based on science and consideration of data obtained in the past which has been stored as a data base. Utilization of previous data as a reference source to enable future events or data mathematically [4]. Forecasting is a statement regarding the future value of a variable such as demand. This means that a forecast is a prediction about the future [5]. One branch of science in projecting possible future events is forecasting. The use of historical data as a reference tool for forecasting next year's sales based on historical sales data from the previous year using mathematical science [3].

## Forecast Model

Forecast models provide a more orderly and focused way of working, so that better analysis techniques are obtained, which can be expected to be able to provide greater trust or confidence, because they can be tested and scientifically proven deviations or deviations that occur. Based on their nature, forecasts are divided into two, namely:

1. *Quantitative Approach*

Quantitative approach is a prediction that is based on a party's opinion and the data cannot be represented strictly as a number or value. A person who compiles the results of forecasting data analysis using one of the mathematical methods is one of the determining factors for the results of forecasting analysis.

1. *Qualitative approach*

Qualitative approach is a forecast that is based on past quantitative data and can be made in the form of numbers which can be called time series data which has a relationship with one or more variables that influence it.

To make a forecast, first determine the forecast period that will be carried out. Based on the forecast time, it can be grouped into three categories [3] namely:

* + 1. *Short Term Forecast*

Short term forecasts can be classified as forecasts for a period of less than three months. These forecasts are usually used to plan purchases, work scheduling, sales, number of workers, and production levels.

* + 1. *Medium Term Forecast*

Medium term forecasts can be classified into forecasts for a period of three months to three years. Medium-term forecasts are used for sales planning, production planning and budgeting, cash budgets and analyzing various operating plans.

* + 1. *Long Term Forecast*

Long term forecasts can be classified as forecasts for a period of more than three years. Medium-term forecasts are used to plan new products, capital expenditures, location, or facility development and research and development.

## Exponential Smoothing

Exponential Smoothing is a procedure for continuous improvement in forecasting the latest observation objects. He emphasizes the exponential decrease in priority for older observation objects. In other words, recent observations will be given higher priority for forecasting than older observations. There are several methods that are grouped into the Exponential Smoothing method, namely:

* + 1. *Single Exponential Smoothing*

Simple exponential smoothing is another term for single exponential smoothing which is used for short-term forecasting analysis, namely forecasting for monthly data. Single exponential smoothing assumes that data will remain at the fluctuation of its mean value. Here's the formula:

Ft + 1 = α Xt + (1-α) Ft  (1)

Theory the value of α = 0.1 or 0.5 or 0.9 (depending on the problem and achieving the desired results)

Contoh : Ft (feb + 1)=0,1 . 9265 + (1-0,1) 9325 Ft Maret = 9319

Ft = Forecast value for the t-th time period

Ft-1 = Forecast value for the past time period, t-1 Xt + (1-α) = Actual time series value

* + 1. *Double Exponential Smoothing*

The use of the double exponential smoothing method looks at aspects of historical data sets that have a trend. In this method, there are two attributes that are always updated, namely level and trend. So this method adds two attributes at the analysis stage, namely level (α) and trend (β).

At = αYt + (1- α) (At-1 +Tt-1) (2)

Tt = β (At – At-1) + (1 – β) Tt-1 (3)

Ft+m = At + Ttm (4)

At = exponential smoothing value

= smoothing constant for data (0 < <1)

= smoothing constant for trend estimation (0 < < 1)

Yt = actual value in period t

Tt = trend estimation

F(t+m) = Forecast value

* + 1. *Triple Exponential Smoothing*

This method is a development of the two previous methods which only involve level and trend factors, but this method adds a seasonal factor (γ) so that this method has two derivatives that take into account the type of season, namely:

* Multiplicative Seasonal Model
* Additive Seasonal Model

The difference between the Multiplication Seasonal Model and Additive Seasonal is as follows:

The Multiplicative Seasonal Model is multiplying the results of level and trend calculations with Seasonal calculations. Meanwhile, the Additive Seasonal Model is adding the results of level and trend calculations with Seasonal calculations. But what will be discussed here is only Triple Exponential Smoothing with the Multiplicative Seasonal Model.

At = α$\frac{Yt}{St-L}$ + (1- α)(At-1 +Tt-1) (5)

Tt = β (At- At-1) + (1- β) Tt-1 (6)

St = γ$\frac{Yt}{At}$ + (1- γ) St-l (7)

Yt+p = (At + Ttp) St-L-p (8)

At = exponential smoothing value

a = smoothing constant for data (0 < a < 1)

b = smoothing constant for trend estimation (0 < b < 1)

m = smoothing constant for seasonal estimation (0 < m < 1)

Yt = actual value in period t

Tt = trend estimation

St = seasonal estimates

L = the length of the season

p = the number of future periods that will be forecast

## Measurement Error

These measures can be used to compare different forecasting models, along with monitoring forecasts to ensure that they are working properly. The three most well-known measures are mean absolute deviation (MAD), mean squared error (MSE), and mean absolute percent error (MAPE). But here the average error (ME) will be added as the first reference for finding the error value.

1. ME (Mean error) or Average Error Value

ME = $\frac{∑ Actual-Forecasting}{n}$ (9)

1. MAD (Mean Absolute Deviation) or Absolute Error Average Deviation Value

ME = $\frac{∑|Actual-Forecasting}{n}$ (10)

1. MSE (Mean Square Error) or Mean Squared Error Value

MSE = $\frac{∑|Actual-Forecasting|2}{n}$ (11)

1. MAPE (Mean Absolute Percent Error or Average Value of Percentage Error )

MAPE = $\frac{\sum\_{i-1}^{n}\left|Actual-Actual\right|/Actual}{n}$ (12)

## Geographic Information System

GIS is a system that is capable of analyzing data from data input, data processing, to producing output in the form of spatial data which is generally displayed in geographic form. Any data that refers to a location on the earth's surface can be called spatial data, for example population density data in an area, road network data in a city, data on the distribution of sampling locations, and so on [6]. GIS consists of a combination of database management in collecting and storing large amounts of geospatial data, to determine the relationship between the entities of each data used, plus map layers which function to describe geospatial data relationships in two and three dimensions in the form of maps. [8].

GIS basically consists of 3 data processing parts related to an input process, process, report of analysis results which are the result of management and processing of the data [15]. The reason GIS is needed is because it is very difficult to handle spatial data, especially because maps and statistical data quickly become outdated so that there is no service to provide more effective data and information. [6]. Database Management System (DBMS) is a sub-system used to organize spatial and attribute data into a database so that it is easy to call, update and delete for data processing needs [7].

In general, a Geographic Information System consists of four main components, namely [10]:

1. Data, basic materials that can be processed or processed into information that has meaning so that it is useful to use. The data used in GIS is divided into two, namely spatial data and non-spatial data.
2. Hardware, is a physical device in the form of a computer and its supporting instruments.
3. Software is a module system that functions to operate GIS.
4. Use includes human resources or human intelligence (brainware).

# RESULTS AND DISCUSSION

## Research Stages

The first stage of this research is data collection, which is the initial stage to be able to predict harvest results. data collection from various sources and literature and secondary data used is previous year's harvest data. After collecting data and analyzing it using the double exponential smoothing method to predict agricultural yields by paying attention to the standardization of each component that influences the prediction results. The second stage is a literature study which collects information from various relevant and scientific sources to support the application of prediction methods. The third stage in this research is system requirements analysis. System requirements analysis is carried out so that the forecasting system that is the output of this research meets needs.

The fourth stage in this research is system design which refers to the results of system requirements analysis. This design is carried out to create a system according to the results of the needs analysis. The fifth stage in this research is system testing to find out whether the system has deficiencies or not, so that users will find it easy to use the system. The sixth stage in this research is the implementation of the system that has been created. This implementation is carried out to find out whether this system can predict or not.

## Design

At the application design stage there are several processes that will be carried out starting from the system workflow to the system database, namely as follows:

1. Flowchart



Figure 1. Flowchart of the Double Exponential Smoothing Forecast System

Based on the flowchart above it can be explained as follows:

* Enter harvest data to look for the alpha value
* Calculate forecasting values
* Displays forecasting results
* Calculating the Mean Error (ME) value
* Calculate the Mean Absolute Deviation (ME) value
* Calculate the Mean Percentage Error (MAPE) value
* Calculate the Mean Square Error (MSE) value
* Displays the output of estimated harvest results for the following year. Next save the forecast, otherwise it's done.
* geografis untuk setiap kecamatan. Kemudian selesai Displays the estimated harvest results in the geographic information system for each sub-district. Then it's done.
1. Use Case



Figure 2. *Use Case* Diagram

The use case diagram describes the activities that admins can carry out in a forecasting system with a geographic information system, namely:

* Admin can add data for each sub-district in Sungai Full City, as well as edit data and delete data.
* Admin can input, delete and edit plant type data.
* Admin can input, delete and edit harvest data for each sub-district.
* Admin can manage geographic information systems.
* Admin can see forecast results that have been processed using the double exponential smoothing method and can print forecast results reports. Admin can see forecast results on the geographic information system.

## Double Exponential Smoothing Calculation

This research was conducted in Sungai Penuh City, Jambi Province with secondary data from 2018 to 2023.

1. RICE FIELD PRODUCTS /(TON)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Subdistrict** | **2018** | **2019** | **2020** | **2021** | **2022** | **2023** |
| Tanah Kampung | 13823 | 14211 | 10082 | 12432 | 16790 | 16299 |
| Kumun Debai | 4905 | 4548 | 7400 | 5524 | 10197 | 9874 |
| Sungai Penuh | 17842 | 12270 | 10727 | 13205 | 4272 | 4953 |
| Pondok Tinggi | 0 | 6140 | 9181 | 7780 | 10863 | 8958 |
| Sungai Bungkal | 0 | 5821 | 8372 | 7949 | 5494 | 9889 |
| Hamparan Rawang | 1952 | 0 | 0 | 0 | 1584 | 11866 |
| Pesisir Bukit | 4767 | 0 | 0 | 0 | 6208 | 7669 |
| Koto Baru | 0 | 0 | 0 | 0 | 2871 | 5984 |

Source : BPS Kota Sungai Penuh

TABLE II. CORN HARVEST RESULTS /(TON)



The initial step of this forecasting system is to input data on harvest results each year based on plant code, plant name and number of harvests (actual value). The next process is calculating the forecast value using the forecast method (double exponential smoothing). The following are the results of the forecasting calculations.

1. *Estimated Rice Harvest Results for Tanah Kampung District*

TABLE III. RESULTS OF CALCULATION OF DOUBLE EXPONENTIAL SMOOTHING RICE DISTRICT TANAH KAMPUNG

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast /ton |
| 2018 | 13823 | - |
| 2019 | 14211 | 13823 |
| 2020 | 10082 | 14075 |
| 2021 | 12432 | 11537 |
| 2022 | 16790 | 11578 |
| 2023 | 16299 | 14559 |

TABLE IV. CALCULATION RESULTS OF THE FORECAST ERROR VALUE OF RICE HARVEST IN TANAH KAMPUNG DISTRICT



Based on Table IV, the calculation results show that the smallest error value is: ME = 25288.68, MAD = 2265.80, MAPE = 17%, MSE = 6935856.07. This is used to predict the next year's harvest, namely:

TABLE V. FORECAST RESULTS OF RICE HARVEST IN TANAH KAMPUNG DISTRICT

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast/ton |
| 2018 | 13823 | - |
| 2019 | 14211 | 13823 |
| 2020 | 10082 | 14075 |
| 2021 | 12432 | 11537 |
| 2022 | 16790 | 11578 |
| 2023 | 16299 | 14559 |
| 2024 |  | 16065 |

Figure 3. Comparison of Actual and Estimated Paddy Rice Harvest

1. *Forecast of Corn Harvest Results for Kumun Debai District*

TABLE VI. RESULTS OF CALCULATION OF DOUBLE EXPONENTIAL SMOOTHING CORN HARVEST IN KUMUN DEBAI DISTRICT

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast/ton |
| 2018 | 43 | - |
| 2019 | 47 | 43 |
| 2020 | 58 | 43 |
| 2021 | 44 | 45 |
| 2022 | 41 | 45 |
| 2023 | 49 | 44 |

TABLE VII. CALCULATION RESULTS OF CORN HARVEST FORECAST ERROR VALUE FOR KUMUN DEBAI DISTRICT



Based on Table VII, the calculation results show that the smallest error value is: ME = 25288.68, MAD = 2265.80, MAPE = 17%, MSE = 6935856.07. This is used to predict the next year's harvest, namely:

TABLE VIII. FORECAST RESULTS OF CORN HARVEST IN KUMUN DEBAI DISTRICT

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast/ton |
| 2018 | 43 | - |
| 2019 | 47 | 43 |
| 2020 | 58 | 43 |
| 2021 | 44 | 45 |
| 2022 | 41 | 45 |
| 2023 | 49 | 44 |
| 2024 |  | 45 |
| 2025 |  | 40 |
| 2026 |  | 35 |

Figure 4. Comparison of Actual and Forecast Corn Harvest

1. *Forecast of Sweet Potato Harvest Results for Pesisir Bukit District*

TABLE IX. RESULTS OF CALCULATION OF DOUBLE EXPONENTIAL SMOOTHING SWEET POTATOES, PESIR BUKIT DISTRICT

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast/ton |
| 2018 | 0 | - |
| 2019 | 51 | - |
| 2020 | 129 | 28,05 |
| 2021 | 203 | 86,12 |
| 2022 | 136 | 158 |
| 2023 | 120 | 159,34 |

TABLE X. RESULTS OF CALCULATION OF THE ERROR VALUE OF THE FORECAST OF SWEET POTATO HARVEST IN PESIR BUKIT DISTRICT



Based on Table From the calculation results, the smallest error value is obtained, namely: ME = 137.85, MAD = 66.03, MAPE = 0.57%, MSE = 5696.84, this is used to predict the next year's harvest, namely:

TABLE XI. FORECAST RESULTS OF SWEET POTATO HARVEST IN PESIR BUKIT DISTRICT

|  |  |  |
| --- | --- | --- |
| Year | Actual/ton | Forecast/ton |
| 2018 | 0 | - |
| 2019 | 51 | - |
| 2020 | 129 | 28,05 |
| 2021 | 203 | 86,12 |
| 2022 | 136 | 158 |
| 2023 | 120 | 159,34 |
| 2024 |  | 150,05 |
| 2025 |  | 77,89 |

Figure 5. Comparison of Actual and Forecasts of Sweet Potato Harvest

So by using the value α = 0.1 and the value β = 0.1, it can be estimated that the yield of sweet potatoes in Pesisir Bukit District in 2018 was 150.05 tons and in 2019 it was 77.89 tons.

## Screen Display

In the initial display, the admin can see a graph of land area based on sub-district. There are several menu bars, namely master data consisting of sub-district data, plant types, annual harvest results. In the analysis results menu there are forecast results for food crop harvests in each sub-district according to the double exponential smoothing calculation. The geographic information system menu will display harvest results based on the analysis results and you can see the distribution of plants in the geographic information system.



Figure 7. Harvest Forecast Results Graph



Figure 8. Forecast Results Menu Display



Figure 9. Geographic Information System Menu Display

# CONCLUSION

Based on the research results, it can be concluded that the double exponential smoothing method can be applied to predict food crop yields. This research was carried out in Sungai Penuh City to predict several types of food crops such as the yield of lowland rice, corn and sweet potatoes. In this study, to estimate corn harvest yields in Kumun Debai District using the smallest α and β values, ME = 80.92, MAD = 5.58, MAPE = 11%, MSE = 52.69 were obtained. So it is estimated that the corn harvest in Kumun Debai District in 2024 will be 45 tons, in 2025 it will be 40 tons and in 2026 it will be 35 tons. The Double Exponential Smoothing method has good performance in forecasting annual data.

##### References

1. Anggodo, Yusuf Ptiyo dan Wayan Firdaus Mahmudy. Peramalan Butuhan Hidup Minimum Menggunakan Automatic Clustering Dan Fuzzy Logical Relationship. *Jurnal Teknologi Informasi Dan Komputer* (JTIIK), 3(2), p.94-120.
2. CC, Holt. 2004 Forecasting Seasonals And Trends By Exponentially Weighted Moving Averages. *Int. J. Forecast*. 20. P.5-13.
3. Heizer, Jay dan Barry, Render. 2015. *Manajemen Operasi, Edisi 11,* Jakarta: Salemba Empat.
4. Assauri, Sofjan. 2016. *Manajemen Operasi Produksi Edisi 3*. Jakarta : Raja Grafindo Persada.
5. Stevenson, William J Dan Sum Chee Choung. 2014. *Operation Menagement: An Asian Prespective*. Jakarta : Salemba Empat.
6. Budiyanto, E., 2016. *Sistem Informasi Geografis dengan Quatum GIS*. Yogyakarta : Andi Offset.
7. Nyerges, T., 2009*. GIS and Society*, University of Washington, Seatle, WA, USA.
8. Vatsavai, R., S. Shekhar, T. E. Burk, Lime, S., 2014. Mapserver : A high performance, interoperable, and open source web mapping and geo-spatial analysis system, *Geographic Information Science*, pp.400-4017.
9. Chen Tian-En, L.P., Chen, Yunbin, G.Y., WANG.2009. Spasial Decision Support System For Percision Farming Based On GIS Web Service, *Information Technology And Application*. 2 p.372-376.
10. Mohammed, E.S. Saleh, A.M., Belal, A.A. 2014. Sustainability Indicators For Agriculture Land Use Based GIS Spacial Modeling In Notrth On Sinai-Engypt, *The Engyptian Journal Of Remote Sensing And Space Science*. 17 p.1-15.
11. E. Vercher, A. Corberan-Vallet, J.V. Segura, Et Al. 2012. Initial Conditions Estima-Tion For Improving Forecast Accuracy In Exponential Smoothing, 20. p. 517–533.
12. R.R. Yager. 2013. Exponential smoothing with credibility weighted observations, *Inf.Sci*. 252. p. 96–105.
13. A.L. Santiago Maia, F.D.A.T. De Carvalho, Et Al. 2011. Holt’s Exponential Smoothingand Neural Network Models For Forecasting Interval-Valued Time Series, *Int. J.Forecast*. 27. p.740–759.
14. A.B. Koehler, R.D. Snyder, J.K. Ord, Et Al. 2016, A Study Of Outliers In The Expo-Nential Smoothing Approach To Forecasting, *Int. J. Forecast*. 47(28). P.7–484.
15. Nyerges, T., 2009. Gis And Society. University of washinton, staatle, WA,USA.