ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 3, Issue 2, October 2023, pp. 151-162

Analyzing Potential Rice Harvest Area in Mojokerto Regency in 2021 Using Area Sample Framework (ASF)

Moh. Alfian Nugroho [a,1], Fera Yuliana [a,2], Arief Krisnah Sholikhan [a,3], Yesy Diah Rosita [b,4,*]

[a] Universitas Islam Majapahit, Jln Raya Jabon KM 0.7, Kab. Mojokerto 61364, Indonesia
[b] IT Telkom Purwokerto, Jln. D.I. Panjaitan No 128, Jawa Tengah, 53147, Indonesia
[1] alfiannugroho1402@gmail.com; [2] ferayuliana310@gmail.com; [3] ariefsholikhan412137@gmail.com; [4] yesy.diah@ittelkom-pwt.ac.id*
* Corresponding author

ARTICLE INFO

Keywords
Agriculture
Harvested area
Production data
ASF
MAPE
MAE
RMSE

ABSTRACT

The agricultural sector is crucial for achieving SDG 2, addressing hunger, ensuring food security, and promoting sustainable agriculture. This study applies the Area Sample Framework (ASF) to estimate rice harvest yields in Mojokerto Regency, emphasizing the importance of accurate agricultural data for effective policy formulation and SDG support. ASF utilizes square segment-based sampling units to provide potential rice harvest area data. However, research on the accuracy of ASF-derived data, especially for predicting the next year's rice harvest, is limited. This study evaluates ASF data accuracy for 2019, 2020, and 2021 using three key metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Results show varying accuracy each year. In 2019, MAPE was 91%, with MAE and RMSE around 2,714.75 ha and 15,463,954.79 ha, indicating high accuracy. Conversely, in 2021, MAPE rose to 107%, with MAE and RMSE near 2,680.09 ha and 14,677,241.22 ha, revealing lower prediction accuracy. This study underscores the importance of continuous monitoring and enhancing data accuracy to support sustainable agriculture and food security, especially in regions like Mojokerto Regency. Further research should investigate factors affecting harvested area efficiency and ways to improve prediction accuracy for effective SDG implementation.

1. Introduction

The agricultural sector has a very significant contribution to achieving the second goal of the Sustainable Development Goals (SDGs) program, namely no hunger, achieving food security, improving nutrition, and encouraging sustainable agricultural cultivation [1]. The agricultural sector is one of the sectors that has become the center of attention in national development, especially those related to the management and utilization of strategic results [2], especially those concerning food commodities [3]. If the planted area is directly proportional to the harvested area, intensification by increasing the planting intensity will increase the efficiency of the harvested area [4]. The efforts to increase the field is knowing the level of efficiency of the harvested area and what affects it [5]. Furthermore, the availability of timely and accurate agricultural data is the foundation to be able to realize agricultural policies that are right on target, data collection on the Harvard of both rice is still

p-ISSN 2798-253X
e-ISSN 2798-3153

using the conventional method using the Agricultural Statistics (AS) data list [1]. The Area Sample Framework (ASF) method was officially implemented as a method of collecting data on rice harvested areas throughout Indonesia and marked the start of a new era of national food data. The resulting production data is expected to be more objective, fast, and accurate using the ASF [6]. Rice yields will become more predictable, so they can simulate crop yields if the SDGs program is to be achieved. The results of the area of rice harvested cannot be predicted, so it is difficult to simulate the Sustainable Development Goals (SDGs) program [1]. When compared with ideal conditions, the current conditions are still far from expectations. By using the calculation of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) on the previous year's Rice Harvested Area data, it is possible to produce an estimate of Rice Yield for this year with an accuracy rate of around 90% [7]. The calculation of the MAE will be used to measure the estimated Rice Yield based on the 2019 and 2020 Rice Harvested Area data. and 2020 in percentage form [8]. One of the advantages of the KSA method (Water Resources Study) is its ability to provide estimates of rice harvest potential up to three months in advance based on direct observations of the rice crop's growth phases in the field. This advantage is crucial for meeting the data requirements as the basis for anticipatory and forward-looking policy formulation related to rice commodities in Indonesia. With this information, authorities can take more timely and effective actions in managing water resources and rice farming to meet the needs of the Indonesian population [16].

2. The Proposed Method

In this chapter, there is a fundamental theory that supports the research and assists in the implementation process. The Sample Area Framework (ASF) is a method that can be utilized to enhance the objectivity of data collection methods through area-based surveys conducted via direct observations of sample segments. Its purpose is to estimate the area through extrapolation from samples to the population within a relatively short period (rapid estimate). This study elaborates on predicting rice harvest yields based on the land area conditions calculated using the MAE, RMSE, and MAPE formulas. The aim of this research is to determine the accuracy level of rice harvest predictions according to the Area Sample Framework data.

2.1. Area Sample Framework (ASF)

The ASF survey uses an area-based sample frame with the sampling unit in the form of a square segment that is formed with an artificial segment shape measured at 300 m × 300 m. Observations in each segment are represented by nine observation points which are allocated systematically with fixed positions in each segment [9]. The important information produced by ASF is data on the potential area for rice harvest in the next year. Information on potential harvests for the next year is very important in determining food policy. Therefore, the issue of the accuracy of the harvest potential data becomes very crucial [10]. Unfortunately, until now, research focused on evaluating the accuracy of data on potential harvested areas obtained from ASF results has not been available. Therefore, this study aims to fill this gap by evaluating the accuracy of the data on potential rice harvest areas obtained from ASF results so that it can be used as a scientific reference in the use of ASF data to support government food policies [9]. Besides that, this study also tries to explore other alternative models that can be applied to predict rice harvested area for the next year [1].

2.2. Prediction

Prediction is a process of systematically estimating something that is most likely to happen in the future based on past and present information that is owned so that the error (difference between something that happens and the predicted result) can be minimized. Prediction of rice harvested area is important to support the national development of the agricultural sector in a country or region [9]. For the effectiveness of agricultural sector development planning, the accuracy of the prediction of rice production becomes increasingly crucial [11].

2.3. Area Condition

The agricultural sector is one of the sectors that has become the center of attention in national development, especially those related to the management and utilization of strategic results [2], especially those concerning food commodities [3]. The area as well as increasing planting intensity

p-ISSN 2798-253X
e-ISSN 2798-3153

from 1 time to 2 and 3 planting times every year [12]. If the planted area is directly proportional to the harvested area, intensification by increasing the planting intensity will increase the efficiency of the harvested area [4]. so that by knowing the level of efficiency of the harvested area and what affects it, it is also known what efforts can be taken to increase it [5].

2.4. Mean Absolute Error (MAE)

The evaluation of the rice harvest in 2021 uses several evaluation tools that are commonly used in evaluating the results of future forecasts on time series data, namely MAE [13], RMSE, and MAPE. MAE is the deviation of the prediction data in the same unit of data, by averaging the absolute error value of all prediction results. According to [14] absolute values are used to avoid positive deviation values and negative deviations which can cancel each other out [14]. The MAE formula used in this study is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i| = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (1)$$

where *MAE* is the Mean Absolute Error, *n* represents the total number of samples or data points in your dataset, y_i is the actual observed value for the ith data point, and $\hat{y}$ is the predicted or forecasted value for the i^{th}data point.

The followings are more detailed explanation of the components:

1. y_i (actual value): This refers to the actual, observed value for a specific data point. In your context, y_i represents the realized harvest for a given sample.
2. $\hat{y}_i$(predicted value): This represents the forecasted or predicted value for the i data point. In your case, $\hat{y}_i$(denotes the predicted or potential harvest.
3. $|y_i - \hat{y}_i|$ (absolute error): This is the absolute difference between the actual observed value yi and the predicted value $\hat{y}$. It captures the magnitude of the prediction error without considering its direction (overestimation or underestimation).
4. $\sum_{i=1}^{n}|y_i - \hat{y}_i|$ (sum of absolute errors): This part of the formula calculates the total sum of the absolute errors for all data points.
5. $\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$ (average absolute error): Finally, this expression computes the average absolute error by dividing the sum of absolute errors by the total number of data points (*n*).

The MAE provides a straightforward way to assess how well your forecasting model's predictions match the actual outcomes. Smaller MAE values indicate better model performance, as they suggest that the model's predictions are closer to the actual values on average.

2.5. Root Mean Square Error (RMSE)

Mean Squared error between the actual value and the forecast value. The Mean Squared Error method is generally used to check the estimation of the error value in forecasting. A low Mean Squared Error value or a mean squared error value close to zero indicates that the forecasting results are by the actual data and can be used for forecasting calculations in the coming period. Mathematically, RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|e_i|} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|} \quad (2)$$

where *RMSE* represents the Root Mean Squared Error which signifies the square root operation and denotes summation across all data points. e_i indicates the error associated with a specific sample, y_i denotes the actual harvest realization for that specific sample, $\hat{y}$ stands for the potential harvest, *n* corresponds to the total number of samples.

The followings are comprehensive breakdown of each component:

1. e_i (error): This denotes the difference between the predicted value $\hat{y}$ and the actual value yi for a specific data point. It encapsulates both the magnitude and direction (positive or negative) of the forecasting error.
2. ${e_i}^2$ (squared error): Squaring the error value e_i is essential because it removes the directional aspect of the error, ensuring that both positive and negative errors contribute equally to the calculation.

 p-ISSN 2798-253X
e-ISSN 2798-3153

3. $\sum e_i^2$ (sum of squared errors): This calculates the total sum of the squared errors across all data points, providing a measure of the overall error magnitude.
4. $\frac{\sum e_i^2}{n}$ (MSE): By dividing the sum of squared errors by the total number of data points (n), we calculate the average squared error, which provides a measure of the average magnitude of errors.
5. $\sqrt{\frac{\sum e_i^2}{n}}$ (RMSE): Taking the square root of the average squared error provides a metric that is in the same units as the original data. This gives us the RMSE value, which indicates the typical magnitude of errors in the predictions.

The RMSE is a crucial metric in assessing forecasting accuracy. A lower RMSE value signifies a better match between predictions and actual observations, indicating higher model performance.

2.6. Mean Absolute Precentage Error (MAPE)

We also use another alternative evaluation measure, namely MAPE. MAPE is used to determine the percentage deviation of the forecast/forecast value with its realization, MAPE is a statistical measurement of the accuracy of the forecast (prediction) in the forecasting method. Measurements using MAPE can be used by the wider community because MAPE is easy to understand and apply in predicting forecasting accuracy. and MAPE also includes calculation methods that are more often used by statisticians to calculate the level of accuracy, MAPE method provides information on how big the forecasting error is. Mathematically, MAPE can be expressed as follows:

$$MAPE = \frac{\sum\left(\left|\frac{e_i}{y_i}\right|\right) \times 100}{n} = \frac{\sum\left(\left|\frac{y_i - \hat{y}_i}{y_i}\right|\right) x\ 100}{n} \tag{3}$$

where *MAPE* stands for Mean Absolute Percentage Error which represents the summation across all data points, e_i signifies the error associated with a specific sample, yi denotes the actual harvest realization for that specific sample, $\hat{y}$ indicates the potential harvest, *n* corresponds to the total number of samples.

The followings are detailed explanation of each component:

1. e_i (error): This represents the discrepancy between the predicted value $\hat{y}$ and the actual value yi for a specific data point. Essentially, it captures the difference between the forecasted outcome and the observed reality.
2. $|e_i|$ (absolute error): The absolute value of the error $|e_i|$ is taken to disregard the direction of the error (whether it's an overestimation or underestimation) and focus solely on its magnitude.
3. $\sum|e_i|$ (Sum of Absolute Errors): This part of the formula calculates the total sum of the absolute errors across all data points.
4. $\sum\frac{|e_i|}{n}$ (average absolute error): This expression calculates the average of the absolute errors by dividing the sum of absolute errors by the total number of data points *(n)*.
5. $\sum\frac{|e_i|}{n} \times 100$ (MAPE): To express the error as a percentage, the average absolute error is multiplied by 100. This gives us the MAPE value, which represents the average percentage deviation between forecasts and actual outcomes.

The MAPE metric provides a percentage-based measure of forecasting accuracy, making it more interpretable and applicable across different contexts. A lower MAPE indicates higher accuracy, as it suggests that the model's predictions are closer to the actual values on average.

2.7. Analysis Method

The framework of our research is briefly presented in Fig. 1. In this study, we evaluate the estimated harvest potential of rice plants in 2021 from ASF results. Evaluation is basically done by comparing the estimated potential harvest with its realization using the scheme in Fig. 1. After we get this data, we try to evaluate the accuracy of the data on the potential harvest area of KSA because no one has ever evaluated the data, the next step is to determine what methods we want to use in measuring the level of accuracy of this KSA result data, namely by using the MAE, RMSE, and MAPE. The application of these three methods aims to obtain more accurate results so that they can support further research. This data consists of data for 2019, 2020, and 2021 from January to December, where each month the number varies from hundreds of hectares to thousands of hectares.

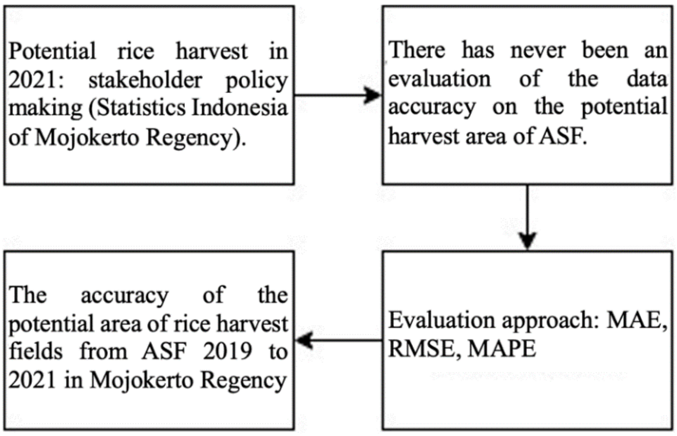

Fig. 1 Research flow chart.

Fig. 1 depicts the stages of analysis in a study concerning the accuracy level of potential rice paddy cultivation areas. This research utilizes data from the Central Statistics Agency of Mojokerto Regency during the period of 2019–2021. The purpose of this data collection is to compute the accuracy level of information regarding the potential rice paddy cultivation areas. The methodology employed in this study is the ASF. This method is used to collect data and subsequently calculate the accuracy level of information regarding the potential rice paddy cultivation areas.

Furthermore, to measure the accuracy of the data pertaining to potential rice paddy cultivation areas, the MAE, RMSE, and MAPE calculation methods are utilized. The data used in this analysis were the result of evaluating the potential rice paddy cultivation areas in Mojokerto Regency from 2019 to 2021, based on the outcomes obtained from the sampling area framework.

3. Method

3.1. Data

Analysis of Potential Rice Harvest Area Data in Mojokerto Regency 2021 Using ASF explanation are presented in Table 1 and Fig. The study utilized rice harvest area data collected through ASF for 2019-2021 in Mojokerto Regency. The data was sourced from the Central Statistics Agency of Mojokerto Regency. It covered 12 periods from January to December, showing changing results in each period. The highest recorded yield was 54,993 ha in 2019, particularly in April with 11,322 ha. The lowest yield was in December 2019 at 1,062 ha. In 2021, the lowest yield was in February at 271 ha [1]. Table 1 highlights similar trends, with the lowest value in 2019's December (1,062 ha) and the highest in April (11,322 ha). For 2020, April had the highest value (17,400 ha) and December had the lowest value (702 ha). In 2021, February had the lowest value (271 ha) and April had the highest value (12,727 ha). Overall, the highest value occurred in April 2020 (17,400 ha) and the lowest was in February 2021 (271 ha) [1]. Through the ASF method, the study assessed data accuracy using MAE, RMSE, and MAPE. This analysis supports the second SDGs, contributing to hunger eradication, food security, improved nutrition, and sustainable agriculture [1].

Table 1. Harvested Area in 2019 (ha)

Period	2019	2020	2021
January	1,762	1,994	1,580
February	1,587	2,090	271
March	11,175	2,160	7,737
April	11,322	17,400	12,727
May	4,743	5,323	5,852
June	3,644	3,022	2,887
July	6,098	6,888	6,697
August	6,575	7,060	4,994
September	4,074	3,865	2,642
October	1,680	2,225	3,085

p-ISSN 2798-253X
e-ISSN 2798-3153

Period	2019	2020	2021
November	1,271	1,778	1,748
December	1,062	702	1,039

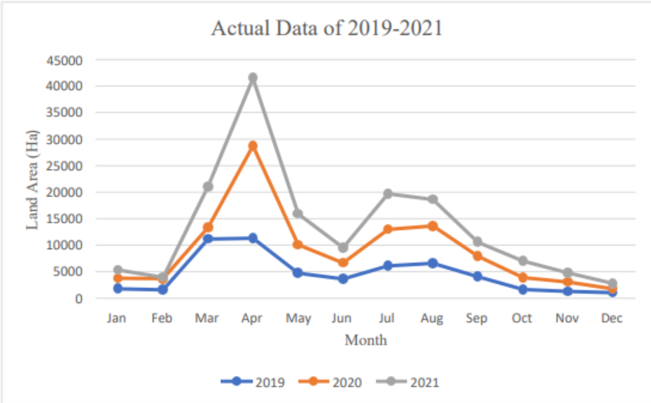

Fig. 2 Actual data line chart.

3.2. Method

The initial stage of this research involves the collection of data regarding the rice paddy cultivation area used during the period from 2019 to 2021. Once the data were gathered, the subsequent step was to conduct data evaluations. The purpose of this evaluation is to assess the quality of the collected data and also to prevent any gaps or missing data. In order to perform this evaluation, the available data were grouped into three clusters based on the respective years: 2019, 2020, and 2021. This grouping aids in better organizing the data and facilitating further analysis.

Following the completion of the clustering phase, the next step involves calculations using three methods: MAE, RMSE, and MAPE. These methods were employed to measure the extent of accuracy in the potential rice paddy cultivation area data that had been collected. Thus, conclusions can be drawn regarding how representative and accurate this data are in reflecting the actual potential rice paddy cultivation area. The following is the processing sequence that begins with MAE, followed by RMSE, and finally MAPE using the data that had been clustered into three categories.

1. Calculation of Rice Harvest Area in 2019 Estimation

Table 2 shows data processing table from January 2019 to December 2019. The actual value is the value of the KSA data, the forecast value is the forecast value obtained from calculations with the benchmark value of the actual value, the error value is the result of calculating the actual value minus the forecast value, the MAE value is the absolute number of the error values, the RMSE value is the squared value of the MAE value, and the MAPE value is the percentage result of the forecast and MAE values. As seen from Table 2, the MAPE value is 91.50%. Meanwhile, MAE and RMSE resulted in 2,714.5 ha and 15,463,954.79 ha respectively.

Table 2. Data Processing in 2019 (ha)

2019 Period	Actual	Prediction	Error	MAE	RMSE	MAPE
January	1,762	1,762	0	0	0	0
February	1,587	1,762	-175.00	175.00	30,625.00	9.93
March	11,175	1,727	9,448.00	9,448.00	89,264,704.00	547.08
April	11,322	3,617	7,705.40	7,705.40	59,373,189.16	213.06
May	4,743	5,158	-414.68	414.68	171,959.50	8.04
June	3,644	5,075	-1,430.74	1,430.74	2,047,028.39	28.19
July	6,098	4,789	1,309.40	1,309.40	1,714,540.93	27.34
August	6,575	5,050	1,524.52	1,524.52	2,324,172.94	30.19
September	4,074	5,355	-1,281.38	1,281.38	1,641,937.08	23.93
October	1,680	5,099	-3,419.10	3,419.10	11,690,277.24	67.05
November	1,271	4,415	-3,144.28	3,144.28	9,886,520.58	71.21

p-ISSN 2798-253X
e-ISSN 2798-3153

2019 Period	Actual	Prediction	Error	MAE	RMSE	MAPE
December	1,062	3,786	-2,724.43	2,724.43	7,422,502.67	71.95
SUM				32,576.95	185,567,457.49	1,097.97
				2,714.75	15,463,954.79	91.50
				MAE	RMSE	MAPE

2. Calculation of Rice Harvest Area in 2020

Table 3 presents data processing table from January 2020 to December 2020. The actual value is the value of the KSA data, the forecast value is the forecast value obtained from the calculation with the benchmark value of the actual value, the error value is the result of calculating the actual value minus the forecast value, the MAE value is the absolute number of the error values, the RMSE value is the squared value of the MAE value, finally the MAPE value is the percentage result of the forecast and MAE values. As seen from Table 3, the MAPE value is 93.62%. At the same time, MAE and RMSE resulted in 2,735.83 ha and 23,301,525.03 ha, respectively.

Table 3. Data Processing in 2020 (ha)

2020 Period	Actual	Prediction	Error	MAE	RMSE	MAPE
January	1,994	1,994	0	0	0	0
February	2,090	1,994	96.00	96,00	9,216.00	4.81
March	2,160	2,013	146.80	146,80	21,550.24	7.29
April	17,400	2,043	15,357.44	15,357.44	235,850,963.35	751.87
May	5,323	5,114	208.95	208,95	43,660.94	4.09
June	3,022	5,156	-2,133.84	2,133.84	4,553,266.32	41.39
July	6,888	4,729	2,158.93	2,158.93	4,660,975.64	45.65
August	7,060	5,161	1,899.14	1,899.14	3,606,745.74	36.80
September	3,865	5,541	-1,675.69	1,675.69	2,807,921.09	30.24
October	2,225	5,206	-2,980.55	2,980.55	8,883,667.62	57.26
November	1,778	4,609	-2,831.44	2,831.44	8,017,044.36	61.43
December	702	4,043	-3,341.15	3,341.15	11,163,289.03	82.64
SUM				32,829.93	279,618,300.33	1,123.47
				2,735.83	23,301,525.03	93.62
				MAE	RMSE	MAPE

3. Calculation of Rice Harvest Area in 2021

Table 4 depicts data processing table from January 2021 to December 2021. The error value is the result of calculating the actual value minus the forecast value, the MAE value is the absolute number of the error values, the RMSE value is the squared value of the MAE value, and the MAPE value is the percentage result of the forecast and MAE values. As seen from Table 4, the MAPE value is 107.10%. Meanwhile, MAE and RMSE resulted in 2,680.09 ha and 14,677,241.22 ha, respectively. In comparison to other years' data, the data from 2021 is the furthest away from being accurate because it produced the highest percentage result of 107.10%.

Table 4. Data Processing in 2021 (ha)

2020 Period	Actual	Prediction	Error	MAE	RMSE	MAPE
January	1,994	1,994	0	0	0	0
February	2,090	1,994	96.00	96,00	9,216.00	4.81
March	2,160	2,013	146.80	146,80	21,550.24	7.29
April	17,400	2,043	15,357.44	15,357.44	235,850,963.35	751.87
May	5,323	5,114	208.95	208,95	43,660.94	4.09
June	3,022	5,156	-2,133.84	2,133.84	4,553,266.32	41.39
July	6,888	4,729	2,158.93	2,158.93	4,660,975.64	45.65
August	7,060	5,161	1,899.14	1,899.14	3,606,745.74	36.80
September	3,865	5,541	-1,675.69	1,675.69	2,807,921.09	30.24
October	2,225	5,206	-2,980.55	2,980.55	8,883,667.62	57.26
November	1,778	4,609	-2,831.44	2,831.44	8,017,044.36	61.43
December	702	4,043	-3,341.15	3,341.15	11,163,289.03	82.64
SUM				32,829.93	279,618,300.33	1,123.47
				2,735.83	23,301,525.03	93.62

p-ISSN 2798-253X
e-ISSN 2798-3153

2020 Period	Actual	Prediction	Error	MAE	RMSE	MAPE
				MAE	RMSE	MAPE

4. Results and Discussion

In this chapter, the comprehensive results of the conducted research are presented. Through the analysis performed, variations in the form of Absolute Percentage Error (APE) differences in each period are observed, as outlined in the research results report. From these findings, a deeper exploration of this matter can be undertaken within a subchapter that emphasizes explanation and discussion. Consequently, this approach can provide a more comprehensive and in-depth perspective, allowing for the formulation of conclusions that are in line with the research findings.

4.1. Results

The analysis of Fig. 3 reveals distinct patterns in the relationship between actual and forecasted data for the year 2019. In January, both actual and forecasted values were 1,762, indicating a close match. However, in February, while the forecast remained at 1,762, the actual data dropped significantly to 1,587. March witnessed a notable increase in actual data to 11,175, while the forecast decreased slightly to 1,727. April showed the lowest actual data point of the year at 11,322, accompanied by a considerable increase in the forecast to 3,617. May brought a substantial decline in actual data to 4,743, contrasting with a rise in the forecast to 5,158. June saw a further decrease in actual data to 3,644, with a corresponding drop in the forecast to 5,075. July marked an increase in actual data to 6,098, while the forecast decreased to 4,789. August saw a decrease in actual data to 6,575, but the forecast increased to 5,050. September showed the highest actual data point of the year at 4,074, along with a peak in the forecast at 5,355. October experienced a decrease in both actual and forecasted data, with actual data falling to 1,680 and forecast data to 5,099. November continued the downward trend, with actual data at 1,271 and forecast data at 4,415. Lastly, December saw the lowest actual data point of the year at 1,062, accompanied by a decrease in the forecast to 3,738.

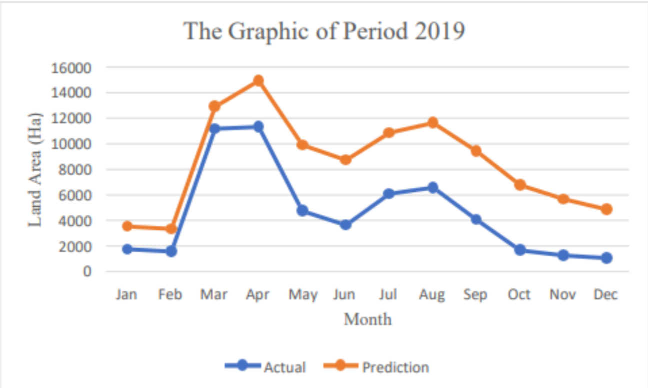

Fig. 3 Data processing flowchart in 2019.

The Fig. 4 shows that the data for the year 2020 displayed various trends and fluctuations in both actual and forecasted values. In February, the forecast was 1,994, but the actual data exceeded expectations, increasing to 2,090. March saw another increase in actual data, reaching 2,160, while the forecast also increased but to a lower value, standing at 2,013. April marked the highest actual data point for the year at 17,400, with the forecast at 2,043. May witnessed a significant decrease in actual data to 5,323, while the forecast increased to 5,114. June saw actual data decrease to 3,022, while the forecast increased to 5,156. July brought an increase in actual data to 6,888, while the forecast decreased to 4,729. In August, actual data increased to 7,060, and the forecast also rose to 5,040. September showed a decrease in actual data to 3,865, but the forecast reached its highest point for the year at 5,541. October saw both actual and forecasted data decrease, with actual data at 2,225 and forecast data at 5,206. November continued the downward trend, with actual data at 1,778 and forecast data at 4,609. Finally, in December, both actual and forecasted data decreased, with actual data hitting its lowest point for the year at 702 and forecast data at 404.

p-ISSN 2798-253X
e-ISSN 2798-3153

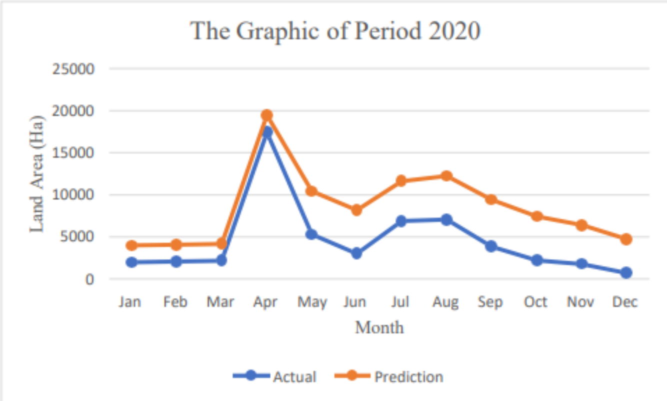

Fig. 4 Data processing flowchart in 2020.

It can be seen from Fig. 5 that the data for the year 2021 showcased various trends and fluctuations in both actual and forecasted values. January started with both actual and forecasted data at 1,580, showing an initial match. February witnessed a sharp decline in actual data to 271, while the forecast remained unchanged at 1,580. March saw a notable surge in actual data to 7,737, while the forecast dropped to 1,318, marking the lowest point in 2021. April stood out with the highest actual data of the year at 12,727, and the forecasted data increased to 2,604. May showed a dip in actual data to 5,852, while the forecast increased to 4,627. June continued the downward trend with actual data at 2,877, and the forecast rose to 4,872. July witnessed another rise in actual data, reaching 6,697, while the forecast decreased to 4,475. August saw a decrease in actual data to 4,994, while the forecast increased to 4,919. September displayed a drop in actual data to 2,642, making the forecast the highest in 2021 at 4,934. October recorded an increase in actual data to 3,085, while the forecast decreased to 4,476. November saw another drop in actual data to 1,748, with the forecast also declining to 4,198. Finally, December marked a significant decrease in actual data to 1,039, and the forecast decreased further to 3,708.

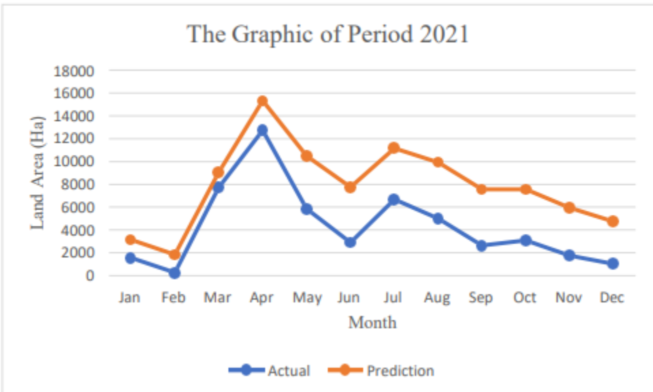

Fig. 5 Data processing flowchart in 2021.

4.2. Statement of Results

According to Table 4, the results derived from applying three different accuracy calculation methods to the data obtained through the ASF are the following: in 2019, the MAE stood at 2,714.75 ha, RMSE amounted to 15,463,954.79 ha, and the MAPE reached 91.50%. In 2020, the MAE was 2,735.83 ha, the RMSE was 23,301,525.03 ha, and the MAPE was 93.62%. Lastly, in 2021, the MAE was 2,680.09 ha, the RMSE was 14,677,241.22 ha, and the MAPE was notably higher at 107.10%. The accuracy levels can be assessed by concluding these results. Analyzing the application of MAE, RMSE, and MAPE on rice harvest area data collected through the ASF in Mojokerto Regency over three years (2019, 2020, and 2021) shows that the most significant outcome was in 2019, where the MAPE value reached 91.50%. The corresponding MAE and RMSE values were 2,714.75 ha and 15,463,954.79 ha, respectively. On the other hand, the year with the most significant deviation from

p-ISSN 2798-253X
e-ISSN 2798-3153

accurate figures was 2021, exhibiting a MAPE value of 107,10%, alongside an MAE of 2,680.90 ha and an RMSE of 14,677,241.22 ha. These outcomes provide valuable insights into the accuracy levels across these three years, which can contribute to further research to achieve more precise results.

Table 4. Data Processing Results

Year	**Accuracy Method**		
	MAE (ha)	***RMSE (ha)***	***MAPE (%)***
2019	2,714.75	15,463,954.79	91.50
2020	2,735.83	23,301,525.03	93.62
2021	2,680.09	14,677,241.22	107.10

4.3. Explanation and Discussion

In general, each phase in 2019, 2020, and 2021 has a varied pattern in the accuracy of the rice harvested area in Mojokerto Regency as a result of ASF. Forecasting results are also very good at capturing annual patterns from data shown by graph patterns of forecast results that are identical to their realization throughout 2019 to 2021. The more identical a prediction pattern is with the actual, the more accurate the prediction. It indicates that harvest predictions in 2019 are pretty accurate. The big gap difference only occurred in March and April. Meanwhile, forecasting in 2020 and 2021 was not as good as in 2019. The gap between forecasting and actual results was quite large, which occurred from March to April every year. However, forecasting results can capture seasonal patterns very well as forecasting in the three years.

The pattern presented above corresponds to the results of the evaluation using measurements on Figure 3, Figure 4, and Figure 5. The results showed that the prediction of the rice harvest in 2019 in the Mojokerto Regency had lower MAE, RMSE, and MAPE values when compared to the two phases. It is in line with the previous description, where the line graph plot of the prediction of the rice harvest based on 2019 was almost identical to that of the realization of the harvest area compared to the other two phases. The MAE and RMSE values for the 2019 phase were approximately 2,714.75 ha and 15,463,954.79 ha, respectively. In addition, the MAPE of the harvest prediction based on the 2019 phase was 91%, indicating that the prediction of rice harvested area based on 2019 had a very good level of accuracy. Meanwhile, the predicted harvest in 2021 had the highest MAE and RMSE values compared to the predictions for the other two years, which were around 2680.09 ha and 14,677,241.22 ha, respectively. Similarly, 2021 had the highest MAPE value in comparison to other years, which was 107%. Given these results, a reexamination is still required, considering the high value of MAE and RMSE. In other words, in 2021, it tends to give less accurate predictions of harvested area [15].

5. Conclusion

The accuracy of applying the ASF method on rice-harvested areas in Mojokerto Regency from 2019 to 2021 using the MAE, RMSE, and MAPE methods had varying results yearly. With the most accurate accuracy in 2019, the MAE and RMSE values were around 2,714.75 ha and 15,463,954.79 ha, respectively, and the MAPE value was 91%. However, reviewing the data collection annually is better to get more accurate results. When viewed more specifically from the results of the KSA, the highest gap rate occurred from March to April every year. Further research is needed to evaluate the accuracy of the potential rice harvest area based on ASF results in more depth. Subsequently, after further study, it is expected to provide a good level of accuracy. The longer the period of the rice harvest realization data from the ASF results, the better the accuracy. Based on the results of the data that were observed during testing the accuracy of forecasting using the method in calculating harvested area in the Mojokerto Regency in 2021, it can be concluded that the results of the analysis of potential harvested area data in Mojokerto Regency in 2021 applied the KSA method using three methods to calculate the accuracy of the land area, regardless of whether the land grew or shrunk each year due to land development. The efficiency of the harvested area is one of the determinants of rice production. The size of rice production is very dependent on the value of the efficiency of a harvested area on the land in the area. Mojokerto Regency, which has a large rice field area, is one

p-ISSN 2798-253X
e-ISSN 2798-3153

of the pillars in meeting food needs, which are increasing along with the increase in the population of Mojokerto Regency. This potential for rice production in Mojokerto Regency can participate in SDGs program. Then if the other regencies can have the same or more potential for rice production, that will make SDGs program work very well and can do namely no hunger, achieve food security, improve nutrition, and encourage sustainable agricultural cultivation. In the future, further and in-depth research is needed on the causes of the low efficiency of the harvested area or land use in order to find ways or efforts to increase efficiency considering the existing potential and the contribution that can be made in meeting food needs in Mojokerto Regency for the next one year. This study was conducted by taking the area of Mojokerto Regency which in general does not necessarily represent the characteristics of the harvested area in Indonesia as a whole and observations made only at one time. Therefore, a study is needed that can take the potential for harvested areas with different characteristics and through more than one times observations to obtain more reliable results.

Acknowledgment

Authors express their gratitude to Universitas Islam Majapahit and Statistics Indonesia of Mojokerto Regency.

References

[1] BPS Kabupaten Mojokerto, "Luas Panen dan Produksi Padi Kabupaten Mojokerto 2021," BPS Kabupaten Mojokerto, https://mojokertokab.bps.go.id/publication/2022/03/31/456f54befab1c817cb17617d/luas-panen-dan-produksi-padi-kabupaten-mojokerto-2021.html (accessed Apr. 27, 2022).

[2] Isnaini, Sirojuzilam, and A. Purwoko, "Peranan Sistem Informasi Padi Better Rice Initiative Asia (BRIA) dalam Mendukung Ketersediaan Pangan Daerah," *Jurnal Serambi Engineering*, vol. 3, no. 2, pp. 306–320, Feb. 2018.

[3] I.G.N.Y. Putra, M. Antara, and I.D.P.O. Suardi, "Efisiensi Penggunaan Faktor-Faktor Produksi pada Usaha Tani Padi Subak Carik Tangis Wongaya Gede Tanaban–Bali," *Jurnal Manajemen Agribisnis*, vol. 6, no. 1, pp. 70–77, May 2018, doi: 0.24843/JMA.2018.v06.i01.p10.

[4] M.W.K. Rini, W. Budiasa, and Widhianthini, "Studi Kelayakan Investasi Pabrik Penggilingan Padi Terintegrasi," *Jurnal Manajemen Agribisnis*, vol. 9, no. 1, pp. 235–248, May 2021, doi: 10.24843/JMA.2021.v09.i01.p02.

[5] D.W. Triscowati, B. Sartono, A. Kurnia, D.D. Domiri, and A.W. Wijayanto, "Klasifikasi Fase Tanam Padi Menggunakan Supervised Random Forest Pada Data Multitemporal Citra Landsat-8," in *Prosiding Seminar Nasional Penginderaan Jauh 2019*, 2019, pp. 1–11.

[6] H. Putra and N.U. Walmi, "Penerapan Prediksi Produksi Padi Menggunakan Artificial Neural Network Algoritma Backpropagation," *Jurnas Nasional Teknologi dan Sistem* Informasi, vol. 6, no. 2, pp. 100–107, Aug. 2020, doi: 10.25077/TEKNOSI.v6i2.2020.100-107.

[7] Kadir and O.R. Prasetyo, "Peramalan Luas Panen Padi Indonesia dengan Model ETS (Error, Trend, Seasonal)," *Jurnal Matematika Dan Statistika Serta Aplikasinya*, vol. 9, no. 1, pp. 7–15, Jan.–Jun. 2021, doi: 10.24252/msa.v9i1.19666.

[8] F.T. Ahsani and D. Ardian, "Analisis Faktor yang Memengaruhi Swembada Beras di Indonesia Tahun 2018," in *Prosiding Seminar Nasional Official Statistics,* 2019, pp. 196–201, doi: 10.34123/semnasoffstat.v2019i1.256.

[9] K. Ruslan, "Improving Indonesia's Food Statistics through the Area Sampling Frame Method," Center for Indonesian Policy Studies, Jakarta, Indonesia, Jul. 2019, doi: 10.35497/287781.

[10] D. Febrimeli, A.Z. Siregar, and M.B. Triyoga, "Analisa Perubahan Sosial dalam Medernisasi Budidaya Tanaman Padi di Kecamatan Secanggang Kabupaten Langkat Provinsi Sumatera Utara," *Jurnal Sosial Ekonomi Pertanian*, vol. 16, no. 3, pp. 272–286, Oct. 2020, doi: 10.20956/jsep.v16i3.11941.

[11] N. Paramitasari and B. Magdalena, "Penerapan Merek dan Digital Marketing dalam Produksi Pangan Beras Desa Banjarsari," *Jurnal Abdi Masyarakat Saburai (JAMS)*, vol. 2, no. 1, pp. 42–51, Apr. 2021, doi: 10.24967/jams.v2i01.1226.

[12] K. Khudori, "Kaji Ulang Kebijakan Perberasan Rice Policy Review," *Jurnal* Pangan, vol. 28, no. 1, pp. 57–72, Apr. 2019, doi: 10.33964/jp.v28i1.421.

[13] M.N. Fawaiq, A. Jazuli, and M. M. Hakim, “Prediksi Hasil Pertanian Padi Di Kabupaten Kudusdegan Metode Brown’s Double Exponential Smoothing,” *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 4, no. 2, pp. 78–87, Dec. 2019, doi: 10.29100/jipi.v4i2.1421.

[14] A.A. Suryanto and A. Muqtadir, “Penerapan Metode Mean Absolute Error dalam Algoritma Regresi Linier untuk Prediksi Produjsu Padi,” *Saintekbu*, vol. 11, no. 1, Feb. 2019, doi: 10.32764/saintekbu.v11i1.298.

[15] U. Ramdhani, “Efisiensi Luas Panen Padi di Provinsi Sulawesi Selatan,” *Ziraa’ah Majalah Ilmiah Pertanian*, vol. 44, no. 1, pp. 72–76, Feb. 2019, doi: 10.31602/zmip.v44i1.1573.

[16] K. Kadir and O. R. Prasetyo, “Peramalan Luas Panen Padi Indonesia Dengan Model ETS (Error, Trend, Seasonal),” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 9, no. 1, p. 7, 2021, doi: 10.24252/msa.v9i1.19666.

p-ISSN 2798-253X
e-ISSN 2798-3153