P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.28 No.2 April - June 2023
Page: [18-30]
Original research article

Selection of Multinomial Logit Models Based on Accuracy Reclassification of the Area Sampling Frame Labels

Yenni Kurniawati[1,2], Hari Wijayanto[2,*], Anang Kurnia[2], Dede Dirgahayu Domiri[3], Budi Susetyo[2]

[1]*Department of Statistics, Universitas Negeri Padang, Padang 25131, Indonesia*
[2]*Department of Statistics, IPB University, Bogor 16680, Indonesia*
[3]*Research Center for Remote Sensing, National Research and Innovation Agency, Jakarta 13220, Indonesia*

Received 16 September 2021; Received in revised form 23 December 2022
Accepted 16 January 2023; Available online 14 June 2023

ABSTRACT

A country needs accurate information about harvested area data to calculate national agricultural production precisely. In Indonesia, the size of the rice paddy harvested area is calculated based on observations from the Area Sampling Frame (ASF) survey. The observations were grouped into eight labels of land conditions: early vegetative, late vegetative, generative, harvest, land preparation, *puso*, non-rice in paddy fields, and non-rice fields. This paper develops several multinomial logit models based on the reclassification of ASF's labels. Furthermore, the models are built by utilizing the LANDSAT 8 spectral indices as a linear predictor. In addition, this study has also used imputation techniques to handle large missing data because the spectral indices data are not available due to cloud cover. The results showed the three best models for the classification of rice growth phases, i.e., the second, third, and fifth models. The second and third models are two classification models based on the reclassification of the ASF's labels. The third model is the recommended model in classifying the rice growth phase in the ASF survey because it has the highest balanced accuracy and can also increase the classification accuracy for the harvest phase. In general, the LANDSAT 8 spectral indices that give the most significant contribution to the harvest phase are EVI_t, $NDBI_{t-1}$, and $MNDWI_{t-1}$. In the future, this model can be used to classify rice growth phases using spectral indices from LANDSAT 8.

Keywords: Area sampling frames survey; Imputation techniques; Multinomial logit models; Reclassification; Spectral indices

***Corresponding author:** hari@apps.ipb.ac.id

1. Introduction

The Area Sampling Frame (ASF) is a method used by Statistics Indonesia-BPS to estimate the national rice harvested area since 2018. The sampling unit of the ASF survey is a segment in the form of a square area with a size of 300 m × 300 m. Each segment consists of 9 sub-segments measuring 100 m x 100 m, the midpoint of which is the observation point in the survey [1]. The ASF method is more scientific, objective, and accurate than the previous method, namely SP-Lahan [1, 2]. However, data collection on SP-Lahan is not based on measurements but uses the eye-estimation method, so the results tend to be biased (upward) and overestimated [1].

ASF surveys are carried out every month with observations labeled with numbers 1-8 for the following categories, namely: (1) early vegetative or vegetative 1, (2) late vegetative or vegetative 2, (3) generative, (4) harvest, (6) land preparation, (6) *puso*, (7) non-rice in paddy fields, and (8) non-rice fields [3, 4]. The first four categories are the rice growth phase used to calculate the harvest area for the current period up to the next three months. The harvest area for the next month can be estimated based on the calculation results of rice fields under the generative phase. Likewise, rice fields in the early and late vegetative phases, where each of these categories can describe the condition of the size of the harvested area in the next two to three months.

The surveyor sends ASF survey data in photos at each subsegment point within a radius of 10 m using the Android-based ASF application software [5]. The photos taken represent the condition of a subsegment. In practice, the coordinates of the subsegment are sometimes below locations that are difficult for the surveyor to visit, such as swamps, bushes, etc. Therefore, although ASF surveys can provide high-quality estimates, there are cost limitations because these surveys require high costs to observe a sample area of the segment in the field [4]. The Indonesian government has allocated a budget of about 64 billion Rupiahs (about US$ 4.35 million) per year to implement the ASF survey nationally in 2018 [5]. Another way to monitor rice fields cost-efficiently and quickly on a wide scale is through remote sensing.

Therefore, technological development such as monitoring data from satellite imagery can better observe ASF points because the range of vision is much broader than the camera photos taken by surveyors in the field. In addition, the remote sensing classification is much more cost-effective than the ASF survey because the data and software are freely available [5].

Remote sensing data such as LANDSAT 8 satellite imagery data are available for free on the Google Earth Engine (GEE) [4, 6]. GEE is a cloud-based platform to make it easy to access vast geospatial datasets using high-performance computing resources [6]. However, LANDSAT 8 has limitations in optical imagery, such as cloud disturbances that can reduce accuracy [7].

The accuracy of the estimated harvested area is essential and needed to obtain the accuracy of calculating the national rice production data. Therefore, this research developed an accurate classification model for the harvest class based on the ASF survey by utilizing LANDSAT 8 satellite imagery data. The satellite imagery data used are three spectral indices algorithms: Enhanced Vegetation Index (EVI) [8], Normalized Difference Bare Index (NDBI) [9], and Modified Normalized Difference Water Index (MNDWI) [10, 11]. These algorithms are often used to observe the greenness of rice plants, identify the harvest phase and indicate the water level present to identify the phase of rice when it is first planted. The form of the algorithms is:

$$EVI = \begin{cases} \dfrac{2.5(\rho_{NIR} - \rho_{Red})}{1 + \rho_{NIR} + 6\rho_{Red} - 7.5\rho_{Blue}}, & \text{if } \rho_{Red} < \rho_{NIR} \text{ or } \rho_{Blue} < \rho_{Red}, \\ \dfrac{1.5(\rho_{NIR} - \rho_{Red})}{0.5 + \rho_{NIR} + \rho_{Red}}, & \text{otherwise,} \end{cases} \quad (1.1)$$

$$NDBI = \frac{\rho_{SWIR_1} - \rho_{NIR}}{\rho_{SWIR_1} + \rho_{NIR}}, \quad (1.2)$$

and

$$MNDWI = \frac{\rho_{Green} - \rho_{SWIR_1}}{\rho_{Green} + \rho_{SWIR_1}}, \quad (1.3)$$

where $\rho_{Blue}, \rho_{Green}, \rho_{Red}, \rho_{NIR}$ and ρ_{SWIR_1} are the reflectance values of the blue, green, red, near-infrared, and $SWIR_1$ bands [3, 4, 9].

In estimating the growth phase of rice, a minimum of two multitemporal satellite image data is required, such as images in the t period and the previous temporal period (t-1 period) from several studies [3, 4, 12, 13]. Furthermore, the addition of three temporal satellite images (t, t-1, and t-2) on the features increases the accuracy of the rice growth phase model [3, 12].

Table 1. Data set with missing value.

Observation (i)	Y	X_1	X_2	⋯	X_p
1	Y_1	X_{11}	NA	⋯	NA
2	Y_2	NA	X_{22}	⋯	NA
⋮	⋮	⋮	⋮		⋮
n	Y_n	X_{1n}	NA	⋯	*Xpn*

EVI deviations, called $dEVI = EVI_t - EVI_{t-1}$ can also distinguish between the vegetative and generative phases. The vegetative is characterized by $dEVI > 0$, while the generative is present when $dEVI < 0$ [13]. Accordingly, this study used LANDSAT 8 spectral indices derived three times of 16-day temporal recording.

Satellite image data is not all clear due to cloud cover at that location, so the value of the image data is Not Available (NA); it is called missing data. One of the ways to handle missing data is by using imputation. Imputation in statistics is completing missing data with a value. The simplest imputation method replaces missing data with mean, median, or mode [14, 15]. However, mean imputation leads to problems in random perturbations of data distribution and underestimating variance, leading to biases in estimates [14, 16]. Imputation under the normality assumption is a viable tool to use when the missing data size is large [14].

The missing data in this study is called missing not at random (MNAR), because the probability of an observation being lost is generally caused by cloud cover or can be said to depend on the information that is not observed. By the time missing data is MNAR, valuable information is lost from the data, and there is no universal method to deal with lost data properly. The overall mean imputation method will yield biased results [17].

Therefore, in this study, the missing data value is a random value generated from a normal distribution with the mean and standard deviation of the spectral index recorded by the LANDSAT 8 satellite at times t, t-1, and t-2. Each of the spectral indices was imputed by generating random numbers that were assumed to follow a normal distribution,

$$\left(x_{ij(missing)} \mid Y = j\right) \sim N\left(\hat{\mu}_{ij(obs)}, \hat{\sigma}^2_{ij(obs)}\right),$$

with the *i*-th spectral index during a specific time in the *j*-th class (see Table 1). In the case of this study, for each missing spectral index in a certain period, the value is generated with a random value that depends on the conditions of each land condition based on the KSA label itself.

The classification model developed refers to the Generalized Linear Model (GLM) with logit as a link function. The logit link is a canonical function for π, which relates the expectation value of the random component (Y) to the systematic component $\sum_{j=1}^{P} x_j \beta_j$ [18]. In this study,

logit models were built based on the reclassification of the ASF label into a new category to explain the rice growth phase, especially the harvest phase.

The multinomial logit model assumes that a random variable (Y) has a multinomial distribution [19]:

$$g_j\left(\boldsymbol{\mu}_i\right)=\log\left(\frac{\mu_{ij}}{\left(1-\sum_{k=1}^{q-1}x_j\mu_{ik}\right)}\right)=\boldsymbol{X}_i\boldsymbol{\beta}, \quad (1.4)$$

where $\boldsymbol{\mu}_i=\left(\pi_{i1},...,\pi_{iq-1}\right)^T$, and π_{ij} is calculated using this form:

$$\pi_{ij}=\exp\left(x_i\beta_j\right)/\left(1+\sum_{k=1}^{q-1}\exp\left(x_i\beta_j\right)\right),$$

and the baseline-category probability is:

$$\pi_{iq}=1/\left(1+\sum_{k=1}^{q-1}\exp\left(x_i\beta_j\right)\right). \quad (1.5)$$

The confusion matrix in Table 2 could be used as a tool that functions to analyze whether the classifier can recognize different classes well [20]. The confusion matrix is generally defined as consisting of two rows and two columns that have True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). However, the confusion matrix is not limited to binary classification and can be used in multiclass classifiers [21].

Sensitivity is the proportion of identified true defaulters, whereas specificity is the proportion of correctly identified non-defaulters. Sensitivity and specificity are used to analyze the model's classifier performance [22]. Accuracy is the proportion of true results, either true positive or true negative, in a population.

Table 2. Confusion matrix 2×2.

		Predicted Class	
		Yes	No
True Class	Yes	TP	FN
	No	FP	TN

Accuracy, sensitivity, and specificity are calculated in the following form:

$$\text{Accuracy}=\frac{TP+TN}{TP+FP+TN+FN}, \quad (1.6)$$

$$\text{Sensitivity}=\frac{TP}{TP+FN}, \quad (1.7)$$

$$\text{Specificity}=\frac{TN}{TN+FP}. \quad (1.8)$$

The balanced accuracy is the average of the sensitivity and the specificity. This index is better than regular accuracy [22]. The classification performance assessment of multiclass datasets (in a square matrix $q\times q$) was measured using average sensitivity and balanced accuracy [21]. Cohen's Kappa measures the agreement between the random approach and modeling approach, where each classifies N items into C mutually exclusive categories. If the approaches are in complete agreement, then $\kappa = 1$. Suppose there is no agreement among the approaches other than what would be expected by chance (as given by pe), κ. The main objective of this study was to analyze the best harvest phase based on five metrics. There are the overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy of the multinomial logistic regression model with two data, namely complete observed data and imputed data.

2. Materials and Method

2.1. Source of research data

This research data comes from 2 data sources: (1) The ASF survey data taken from BPS-West Java and (2) The LANDSAT 8 spectral indices data taken from Google Earth Engine (GEE). The ASF label was taken from observations in the April 2018 survey in West Java Provinces. It was used as a response variable, while the explanatory variable was taken from the LANDSAT 8 spectral index for three periods. The logit models analyzed the classification of ASF labels based on a

dataset of 2,025 ID segments, where a segment is divided into nine subsegments and nine observation points (see Fig. 1).

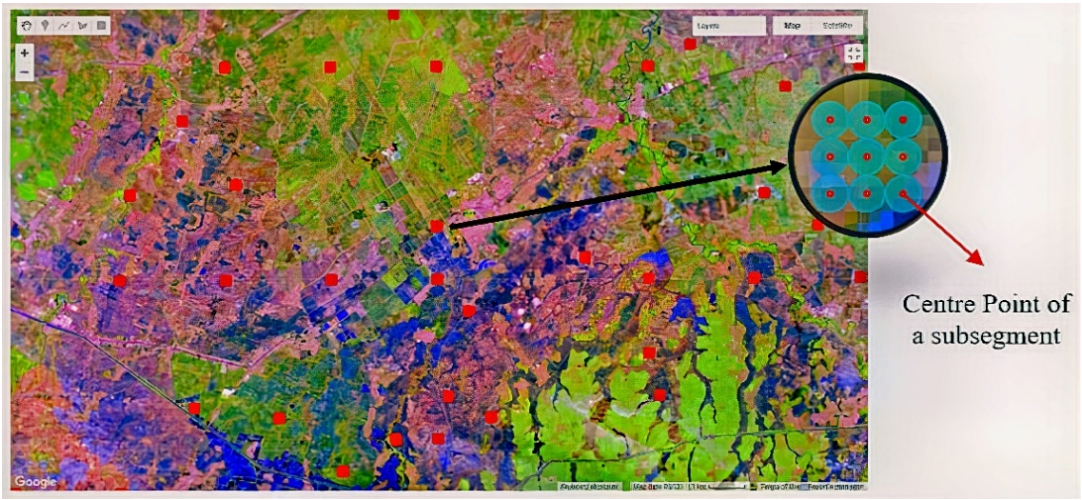

Fig. 1. The ASF sample point with size 100 m × 100 m at April 2018 in West Java Province based on LANDSAT 8 spectral indices using Google Earth Engine (GEE).

However, not all of these points have complete image recording data for all periods. In addition, several coordinates of the spectral index value are not available (NA) due to cloud cover. Therefore, an imputation technique is needed to handle missing data to improve the performance of the classification models.

2.2 Multinomial logit models based on ASF's data

In general, the multinomial logistic model for category j that utilizes the spectral indices is stated as follows:

$$\begin{aligned} g_j(\boldsymbol{\mu}_i) &= \log\left(\frac{\pi_{ij}}{\left(1-\sum_{k=1}^{q-1} x_j \mu_{ik}\right)}\right) \\ &= \beta_0 + \left(\mathbf{EVI}_t\right)_j \beta_{1j} + \left(\mathbf{NDBI}_t\right)_j \beta_{2j} \\ &+ \left(\mathbf{MNDWI}_t\right)_j \beta_{3j} + \left(\mathbf{NDBI}_{t-1}\right)_j \beta_{4j} \\ &+ \left(\mathbf{MNDWI}_{t-1}\right)_j \beta_{5j} + \left(\mathbf{NDBI}_{t-2}\right)_j \beta_{6j} \\ &+ \left(\mathbf{MNDWI}_{t-2}\right)_j \beta_{7j} + \left(\mathbf{dEVI}_1\right)_j \beta_{8j} \\ &+ \left(\mathbf{dEVI}_2\right)_j \beta_{9j}, \end{aligned} \tag{1.9}$$

where j is the rice growth phase category obtained from the ASF's label reclassification, the probability of the harvest phase being the baseline category (q) is calculated using Eq. (1.5).

The effect of the explanatory variable can be interpreted through the value of the Odds Ratio (OR), namely OR = $\exp(\beta_j)$. If OR = 1 there is no relationship between the two categories of variables. However, if OR > 1 the odds of success of first odd 1 are higher than the value of second odd as the baseline category, and vice versa.

2.3 Analysis procedure

The stages in this research are:

1. Create a shapefile (SHP) grid measuring 100 m × 100 m in R software based on the coordinates of the ASF sample of West Java Province. Then, add identity data and observation results in ASF labels for the West Java Province period in April 2018.

2. Extract the spectral indices data through Eqs. (1.1)-(1.3) as accurately as

LANDSAT 8 in 3 temporal periods of 16 days (t, t-1, t-2) where time t is the closest image data recording to the survey time in April 2018 used Platform Google Earth Engine (GEE).

3. Its preprocessing aims to produce data structures in building a classification model for rice growth phases. Preprocessing is done twice in the first stage to get the complete data structure in all periods of image recording. In this stage, cleaning the spectral indices data that is not available (NA) is carried out due to cloud cover at a certain period. Meanwhile, the second stage imputed the missing data using the distribution of the observed data with the assumption that $\left(x_{ij(missing)} \mid Y = j\right) \sim N\left(\hat{\mu}_{ij(obs)}, \hat{\sigma}^2_{ij(obs)}\right)$, for the *i*-th spectral index during a specific time in the *j*-th category.

4. Reclassify ASF labels as shown in Table 3 using several alternative new classes.

Table 3. Reclassification of ASF's labels for classification on logit models.

Types of Response Grouping	Definitions	Models
Label of ASF	$Y = \begin{cases} 1, & \text{Vegetative 1} \\ 2, & \text{Vegetative 2} \\ 3, & \text{Generative 1} \\ 4, & \text{Land Preparation} \\ 5, & \text{Puso} \\ 6, & \text{non-rice in puddy fields} \\ 7, & \text{non-rice fields} \\ 0, & \text{harvest} \end{cases}$	Multinomial Logistic Regression has 8 categories (The 1st Model)
Alternative 1	$Y = \begin{cases} 1, & \text{Pre-harvest}(\text{Veg 1, Veg 2, Gen}) \\ 2, & \text{Others}(\text{LP, Puso, SBP, BS}) \\ 0, & \text{harvest} \end{cases}$	Multinomial Logistic Regression has 3 categories (The 2nd Model)
Alternative 2	Clustering based on Biplot Analysis *	Multinomial Logistic Regression based on clustering of Biplot Analysis (The 3rd Model)
Alternative 3	$Y = \begin{cases} 1, & \text{harvest} \\ 0, & \text{not harvest} \end{cases}$	Binary Logistic Regression (The 4th Model)
Alternative 4 (Rice fields)	$Y = \begin{cases} 1, & \text{Vegetative 1} \\ 2, & \text{Vegetative 2} \\ 3, & \text{Generative} \\ 4, & \text{harvest} \end{cases}$	Multinomial Logistic Regression has 4 categories (The 5th Model)

* using library(ggbiplot) in R.

1. Analyze the logit models as the rice growth phase model based on the response categories in Table 3 for both the observed data and imputed data through the following stages:
   a. Estimate parameters using the maximum likelihood method at library caret in R.
   b. Test the parameters partially and simultaneously.
   c. Determine logit $\eta_j = g_j\left(\boldsymbol{\mu}_i\right)$ through Eq. (1.9).
   d. Interpret Odds Ratios.
2. Determine the best classification model for rice harvest classification based on the value of the overall accuracy, Kappa, sensitivity, specificity, and balanced accuracy for the dataset completely observed and dataset imputed using Eq. (1.6) up to (1.8).

Table 4. Total and proportion of complete data and missing data from recorded LANDSAT 8 satellite imagery for three temporal periods from ASF's sample in West Java Province.

	Labels of ASF	Time (t)		Time (t-1)		Time (t-2)	
		N	Prop.	N	Prop.	N	Prop.
Obs.	1	1227	(0.31)	3429	(0.87)	1132	(0.29)
	2	469	(0.25)	1554	(0.82)	825	(0.44)
	3	511	(0.21)	2203	(0.91)	639	(0.26)
	4	744	(0.22)	3113	(0.93)	802	(0.24)
	5	470	(0.26)	1637	(0.92)	409	(0.23)
	6	15	(0.26)	55	(0.96)	11	(0.19)
	7	268	(0.28)	847	(0.87)	411	(0.42)
	8	1110	(0.29)	3230	(0.85)	1582	(0.42)
Missing	1	2732	(0.69)	530	(0.13)	2827	(0.71)
	2	1425	(0.75)	340	(0.18)	1069	(0.56)
	3	1909	(0.79)	217	(0.09)	1781	(0.74)
	4	2612	(0.78)	243	(0.07)	2554	(0.76)
	5	1312	(0.74)	145	(0.08)	1373	(0.77)
	6	42	(0.74)	2	(0.04)	46	(0.81)
	7	706	(0.72)	127	(0.13)	563	(0.58)
	8	2673	(0.71)	553	(0.15)	2201	(0.58)

3. Results and Discussion

The ASF survey data used in this study mostly had missing data due to cloud cover, especially in periods t (mid to late April 2018) and t-2 (mid to late March 2018). As a result, more than 70% of the recorded LANDSAT 8 satellite images at (t-1) periods were Not Available (NA) see Table 3. At the same time, the LANDSAT 8 image recording condition in the initial period up to mid-April 2018 (t-1) was quite good because there was more than 80% completeness of the spectral indices data in each ASF category (see Table 3). As a result, the number of ASF points completely observed by LANDSAT 8 imagery in 3 periods (t, t-1, and t-2) is only 1,180 out of 18,225 subsegments. Deleting missing data risks losing important information that is useful in establishing a classification model.

The results of the ASF survey observation in April 2018 showed that there was a minority class (*puso*) with an average proportion of 0.4%. This condition will make more bias toward the majority class. Therefore, the ASF label reclassification technique can be used to avoid minority classes. Several models will be formed based on an alternative reclassification of ASF labels (in Table 3).

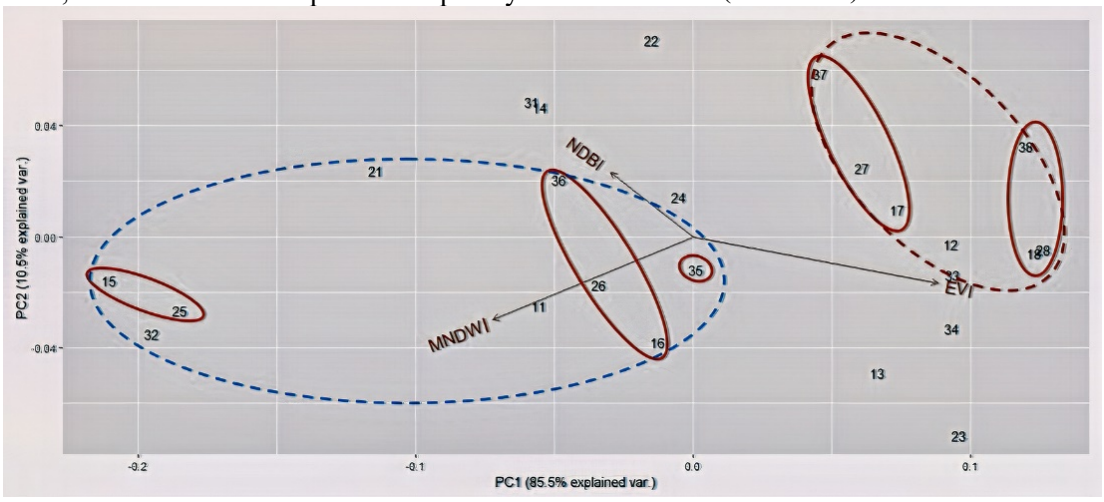

Fig. 2. Biplot of the ASF sample and the LANDSAT 8 spectral indices in West Java Province.

The Biplot in Fig. 2 shows the object's position, namely the ASF label for each period through a two-digit number. The first digit of the number is defined as the label of the image recording period, namely 1=t, 2=t-1, and 3=t-2. Furthermore, the second digit is ASF's labels of eight categories. For example, the number "27" means showing the position of the "7" category, namely non-rice in the paddy fields at the time of recording satellite imagery in period "2", namely t-1 (temporal one period before period t around the survey).

Table 5. Accuracy and Kappa of the classification model in Table 3.

Statistics		Multinomial Logistic Regression			Binary Logistic Regression	Models for Rice Fields
		Model 1	Model 2	Model 3	Model 4	Model 5
Complete Case	Accuracy	0.49	0.65	0.55	0.87	0.65
	Kappa	0.35	0.42	0.39	0.47	0.52
Imputation*	Accuracy*	0.58	0.66	0.61	0.83	0.77
	Kappa*	0.49	0.45	0.51	0.31	0.69

Based on the Biplot, it can be seen that the 7th and 8th grades for each period are clustered on the right, marked by high EVI values. The 5th and 6th class positions are on the left with high MNDWI scores (see Fig. 2). Based on the Biplot, two new categories are formed. There are the fifth and the sixth new class. The fifth new class is a merger of land preparation (the 5th class) and puso phases (the 6th class). Meanwhile, the sixth new class is a class based on the merger of the non-rice in the paddy fields class and non-rice fields class.

Meanwhile, the growth phase of rice in classes 1 to 4 is generally scattered or not clustered based on the characteristics of spectral indices in the three time periods of LANDSAT 8 satellite recording. For example, in the late vegetative phase (class 2) at the time of observation period t, the number "12" is in the right position close to EVI; this characterizes a high greenness index (EVI). In contrast, the observation 16 days before (number "22") is close to NDBI, which characterizes a higher NDBI value, and the observation 32 days before time t (3=t-2) or number "32" is characterized by a high water index (MNDWI). The late vegetative phase and harvest have a reasonably short time. Therefore, the positions of these two categories are seen to be far apart in the three image recording periods.

In contrast, the early vegetative and generative phases have a relatively wide period. There are about 1-35 days after planting (DAT), while the generative phase is 55-105 days after planting (DAT). This condition caused the second position of the category not to spread too far in three periods.

The imputation results increase the accuracy overall and Kappa value of the multinomial logistic regression model, except for the binary logistic regression model (the 4th Model). The third model with imputed data can be reasonably good based on overall accuracy, Kappa, and balanced accuracy, with 0.61, 0.51, and 0.74, respectively (see Tables 5 and 6). The best models based on four statistical accuracy values are the second, third, and fifth.

The second and third models were obtained by reclassifying the eight ASF survey labels. They gave better accuracy results than the first model built from eight categories of ASF surveys. The second model is a model that explains the conditions before and during harvest but has not been able to explain the phase of rice growth. In comparison, the third model is a model that can explain the four phases of

rice growth based on the logit equation obtained (see Table 7).

The third model is proposed as the best rice growth phase model for reclassifying ASF labels. It has good overall accuracy and balanced accuracy in classifying all classes, especially in the harvest phase. This third model consists of 6 response categories, namely (1) early vegetative, (2) late vegetative, (3) generative, (4) harvest class, (5) land preparation and *puso*, (6) non-rice in the paddy fields, and non-rice field classes. New classes are formed based on the results of the biplot analysis. It has good accuracy and balanced accuracy in classifying all classes, especially in the harvest phase.

Conversely, the first and fourth models do not perform well because the first model has a minority class, and binary categorization for the fourth model is also ineffective in improving the model's performance. At the same time, the fifth model is a model that has good accuracy in classifying rice growth phases specifically for rice fields.

The EVI_t , NDBI_{t-1}, and MNDWI_{t-1} are the most significant spectral indices in distinguishing the harvest phase from other rice growth phases in Table 6. For every 0.1 increase in the spectral index to the harvest class classification, the probability that the subsegment is in the harvest phase compared to other phases can be seen from $OR^* = 1/\exp\left(\beta_{pj} \times 0.1\right)$, where $p = 1, \ldots, P$ is the number of explanatory variables, and $j=1,\ldots, q\text{-}1$ is the number of categories for each logit equation.

Table 6. Performance of classification model for each response category.

Model	Data	The performance of classification	Categories of Response Variable								Mean
Model 1st	Complete case	Class	Class: 4*	Class: 1	Class: 2	Class: 3	Class: 5	Class: 6	Class: 7	Class: 8	
		Sensitivity	0.58	0.40	0.57	0.09	0.11	NA	0.00	0.74	0.36
		Specificity	0.90	0.88	0.90	0.96	0.98	1.00	1.00	0.74	0.92
		Balanced Accuracy	0.74	0.64	0.73	0.53	0.55	1.00	0.50	0.74	0.64
	Imputation	Sensitivity	0.64	0.78	0.61	0.62	0.18	0.00	0.01	0.64	0.44
		Specificity	0.89	0.88	0.96	0.94	0.98	1.00	1.00	0.85	0.94
		Balanced Accuracy	0.77	0.83	0.79	0.78	0.58	0.50	0.50	0.75	0.69
Model 2nd	Complete case	Class	Class: 0*	Class: 1	Class: 2						
		Sensitivity	0.49	0.63	0.72						**0.61**
		Specificity	0.95	0.82	0.65						**0.80**
		Balanced Accuracy	0.72	0.72	0.69						**0.71**
	Imputation	Sensitivity	0.44	0.76	0.63						**0.61**
		Specificity	0.93	0.77	0.75						**0.81**
		Balanced Accuracy	0.69	0.76	0.69						**0.71**
Model 3rd	Complete case	Class	Class: 4*	Class: 1	Class: 2	Class: 3	Class: 5	Class: 6			
		Sensitivity	**0.56**	0.39	0.54	0.06	0.22	0.78			0.43
		Specificity	**0.92**	0.90	0.91	0.98	0.98	0.70			0.90

		Balanced Accuracy	**0.74**	0.65	0.72	0.52	0.60	0.74		0.66
	Imputation	Sensitivity	**0.59**	0.77	0.58	0.58	0.18	0.70		**0.57**
		Specificity	**0.90**	0.89	0.97	0.95	0.98	0.83		**0.92**
		Balanced Accuracy	**0.74**	0.83	0.78	0.76	0.58	0.76		**0.74**
Model 4th	Complete case	Class	Class: 1*							
		Sensitivity	0.98							0.98
		Specificity	0.40							0.40
		Balanced Accuracy	0.69							0.69
	Imputation	Sensitivity	0.96							0.96
		Specificity	0.29							0.29
		Balanced Accuracy	0.62							0.62
Model fifth (Rice Field)	Complete case	Class	Class: 4*	Class: 1	Class: 2	Class: 3				
		Sensitivity	0.75	0.53	0.70	0.34				**0.58**
		Specificity	0.87	0.85	0.83	0.91				**0.87**
		Balanced Accuracy	0.81	0.69	0.77	0.63				**0.72**
	Imputation	Sensitivity	0.77	0.82	0.64	0.81				**0.76**
		Specificity	0.91	0.91	0.95	0.93				**0.92**
		Balanced Accuracy	0.84	0.86	0.80	0.87				**0.84**

Note: * for the class of interest, that is harvest class (4).

Table 7. Estimation of multinomial logit model for the 3rd model and 5th model.

Model	logit	Coef/ OR*	Int.	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Model 3rd	η_1	Coef	7.23*	-18.85*	-6.80*	2.39	-6.05*	-0.55*	8.59*	2.96*	18.68*	9.21*
		OR*	0.49	**6.59**	1.97	0.79	1.83	1.06	0.42	0.74	0.15	0.40
	η_2	Coef	-0.53	-7.66*	-7.81*	-3.96*	-8.64*	-1.20*	6.44*	5.22*	17.49*	12.55*
		OR*	1.05	2.15	2.18	1.49	**2.37**	1.13	0.53	0.59	0.17	0.29
	η_3	Coef	-9.67*	8.63*	-8.99*	-0.93*	-11.65*	-2.40*	2.94*	1.74*	0.25	3.70*
		OR*	2.63	0.42	2.46	1.10	**3.21**	1.27	0.75	0.84	0.98	0.69
	η_5	Coef	6.64*	-14.38*	-2.50*	2.95*	-1.62*	-0.37*	6.65*	1.45*	12.37*	5.49*
		OR*	0.51	**4.21**	1.28	0.74	1.18	1.04	0.51	0.87	0.29	0.58
	η_6	Coef	-3.45*	4.49*	-3.25*	-3.34*	-1.55*	-5.53*	9.15*	-2.29*	3.26*	2.86*
		OR*	1.41	0.64	1.38	1.40	1.17	**1.74**	0.40	1.26	0.72	0.75
Model 5th	η_1	Coef	1.01	-20.64*	-11.64*	-1.46	-17.58*	-7.23	0.16*	-0.63	0.23*	0.37
		OR*	0.90	**7.88**	3.20	1.16	5.80	2.06	0.98	1.07	0.98	0.96
	η_2	Coef	-1.04	-15.74*	-5.88*	-4.97	-18.98*	-7.55	0.12*	4.03	0.25*	0.13*
		OR*	1.11	4.83	1.80	1.64	**6.67**	2.13	0.99	0.67	0.98	0.99
	η_3	Coef	-6.87*	-3.54	0.47	5.07	-23.79*	-17.59*	0.10*	4.11	0.12*	1.79
		OR*	1.99	1.42	0.95	0.60	**10.79**	5.81	0.99	0.66	0.99	0.84

Note: OR* is an Odds Ratio for Harvest Class, X_1: EVI_t, X_2: $NDBI_t$, X_3: $MNDWI_t$, X_4: $NDBI_{t-1}$, X_5: $MNDWI_{t-1}$, X_6: $NDBI_{t-2}$, X_7: $MNDWI_{t-2}$, X_8: $dEVI_1$, X_9: $dEVI_2$.

As an illustration in the fifth model, the $\mathbf{EVI}_t$ increase of about 0.1 can be impacted to classify a sub-segment area that tends to be categorized into the harvest phase compared to the early vegetative phase of $1/\exp(-20.64\times 0.1)=7.88$ times. The first logit in the third model shows the tendency of segments to be in the harvest phase compared to the early vegetative when an increase in the $\mathbf{EVI}_t$ of 0.1 is 6.59 times.

4. Conclusion

This study indicates that the multinomial logistic models, namely the third and fifth models, are the best models for classifying rice growth phases based on the ASF survey. The fifth model is a model that can adequately classify the specific harvest phase for rice fields with an accuracy of classification based on the sensitivity and specificity values of 0.77 and 0.91, respectively. The fifth model has been able to classify the harvest class and other phases reasonably well, although the fifth model for all categories is still around 65%. The effect of the spectral indices on the rice growth phase class can be seen through the odd ratio value for each logit equation. The fifth model also has three logit equations (η_1,η_2,η_3). These are describing the logit value for early vegetative, late vegetative and generative phases. The comparison category is the harvest phase. In logit 1 (η_1), it can be seen that the most significant influence of LANDSAT 8 spectral indices on the harvest phase compared to the vegetative phase 1 is the index $\mathbf{EVI}_t$ with a value of $1/\exp(\beta_{11}\times 0.1)=7.88$ times for each addition of $0.1\times\mathbf{EVI}_t$.

The third model is the recommended model for classifying rice growth phases based on the reclassification of ASF labels with six new categories based on the proximity of objects or phases during three periods in the biplot analysis. Two new categories were formed based on the Biplot, namely the fifth class and sixth class. The fifth class is a merger of land preparation and *puso* classes, while the sixth class is a merger of non-rice in paddy fields and non-rice fields classes. This model performs well, especially in classifying the harvest class, with sensitivity, specificity, and balance accuracy values of 0.59, 0.90, and 0.74. In addition, the use of imputation techniques in the third model can increase the overall accuracy by 6%. However, this result is not satisfactory because the overall accuracy is still below 80%. So, it is necessary to do further research to increase the accuracy, such as adding random effects to the model or applying other imputation methods.

The magnitude of the tendency of the subsegment to be in the harvest phase compared to the early vegetative phase when the EVI_t value increased by 0.1 was 6.59 times. The odds ratio EVI_t value for logit 1 in the 3rd model is smaller than in the fifth model. In general, from the model, it can be concluded that the indices that give the most excellent chance of a subsegment being in the harvest phase compared to other phases are $\mathbf{EVI}_t$, $\mathbf{NDBI}_{t-1}$, and $\mathbf{MNDWI}_{t-1}$, because when < 0, then odds (harvest phase) > 1.

Acknowledgments

The authors thank Statistics Indonesia-BPS West Java for the cooperation and for providing information regarding the ASF survey data in West Java Province, Indonesia.

Declaration

Authors have worked together on a publication and contributed equally.

References

[1] Ruslan K. Improving Indonesia's food statistics through the area sampling frame method. Jakarta: Center for

Indonesian Policy Studies; 2019.

[2] Prasetyo OR, Kadir, Amalia RR. A pilot project of area sampling frame for maize statistics: Indonesia s experience. Stat J IAOS 2020;36: 997-1006.

[3] Triscowati DW, Sartono B, Kurnia A, Dirgahayu D, Wijayanto AW. Classification of rice-plant growth phase using supervised random forest method based on Landsat-8 multitemporal data. Int J Remote Sens Earth Sci 2020;16:187–-97.

[4] Triscowati DW, Sartono B, Kurnia A, Domiri DD, Wijayanto AW. Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest. Proc. SPIE 11311, Sixth Geoinf. Sci. Symp., vol. 1131102, 2019, p. 1-10.

[5] Gandharum L, Mulyani ME, Hartono DM, Karsidi A, Ahmad M. Remote sensing versus the area sampling frame method in paddy rice acreage estimation in Indramayu regency, West Java province, Indonesia. Int J Remote Sens 2021; 42:1738-67.

[6] Mab P, Kositsakulchai E. Water balance analysis of tonle sap lake using weap model and satellite-derived data from google earth engine. Sci Technol Asia 2020;25: 45-58.

[7] Gandharum L, Mulyani ME, Hartono DM, Karsidi A, Ahmad M. Remote sensing versus the area sampling frame method in paddy rice acreage estimation in Indramayu regency, West Java province, Indonesia. Int. J. Remote Sens., vol. 42, Taylor & Francis; 2021, p. 1738-67.

[8] Zhao R, Li Y, Ma M. Mapping paddy rice with satellite remote sensing: A review. Sustainability 2021;13:1-20.

[9] Parsa M, Dirgahayu D, Harini S. Development of Paddy Field Classification Method Based on Multi-Temporal Indeces of Landsat Images. J Penginderaan Jauh Dan Pengolah Data Citra Digit 2019;1: 35-44.

[10] Arjasakusuma S, Kusuma SS, Rafif R, Saringatin S, Wicaksono P. Combination of Landsat 8 OLI and Sentinel-1 SAR time-series data for mapping paddy fields in parts of west and Central Java Provinces, Indonesia. Int J Geo-Information 2020; 9:1-17.

[11] Mansaray LR, Huang W, Zhang D, Huang J, Li J. Mapping rice fields in urban Shanghai, southeast China, using Sentinel-1A and Landsat 8 datasets. Remote Sens 2017;9:1-23.

[12] Marsuhandi AH, Soleh AM, Wijayanto H, Domiri DD. Pemanfaatan ensemble learning dan penginderaan jauh untuk pengklasi-fikasian jenis lahan padi. Semin Nas Off Stat 2020; 2019: 188-95.

[13] Dirgahayu D. Aplikasi inderaja untuk mendeteksi awal tanam padi menggunakan data EVI MODIS multitemporal. Crespent Press 2014:39-48.

[14] Almasi I, Salehi M, Moradi M. A new algorithm to impute the missing values in the multivariate case. J Iran Stat Soc 2020;19:133-43.

[15] Buuren S van. Flexible imputation of missing data. Second edi. Boca Raton, Florida: CRC Press; 2018.

[16] Longford NT. Statistics for Social Science and Public Policy. New York, NY 10013, USA: Springer Science+Business Media, Inc.; 2005.

[17] Donders ART, Heijden GJMG Van Der, Stijnen T, Moons KGM. Review : A gentle introduction to imputation of missing values 2006;59:1087-91.

[18] Srisuradetchai P, Junnumtuam S. Wald confidence intervals for the parameter in a bernoulli component of zero-inflated poisson and zero-altered poisson models with different link functions. Sci Technol Asia 2020;25:1-14.

[19] Agresti A. An introduction to categorical analysis. 3rd Editio. Hoboken, NJ: John Wiley & Sons; 2019.

[20] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with application in R. New York, NY 10004, USA: Springer Science+Business Media, Inc.; 2021.

[21] Ballabio D, Grisoni F, Todeschini R. Multivariate comparison of classification performance measures. Chemom Intell Lab Syst 2018;174:33-44.

[22] Di Biase L, Di Santo A, Caminiti ML, De Liso A, Shah SA, Ricci L, et al. Gait analysis in parkinson's disease: An overview of the most accurate markers for diagnosis and symptoms monitoring. Sensors 2020;20:1-22.