

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest

Dwi Wahyu Triscowati, Bagus Sartono, Anang Kurnia, Dede Dirgahayu Domiri, Arie Wahyu Wijayanto

Dwi Wahyu Triscowati, Bagus Sartono, Anang Kurnia, Dede Dirgahayu Domiri, Arie Wahyu Wijayanto, "Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest," Proc. SPIE 11311, Sixth Geoinformation Science Symposium, 1131102 (21 November 2019); doi: 10.1117/12.2547216

SPIE.

Event: Sixth Geoinformation Science Symposium, 2019, Yogyakarta, Indonesia

Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest

Dwi Wahyu Triscowati^{*a,b}, Bagus Sartono^b, Anang Kurnia^b,
Dede Dirgahayu Domiri^c, Arie Wahyu Wijayanto^d

^aBPS-Statistics of Banyuwangi Regency, Banyuwangi Indonesia; ^bDept. of Statistics, IPB University, Bogor, Indonesia, ^cIndonesian National Institutes of Aeronautics and Space (LAPAN), Jakarta, Indonesia, ^dBPS-Statistics Indonesia, Jakarta, Indonesia

ABSTRACT

Food crops monitoring in developing countries such as Indonesia plays an essential role to support national goals in food security and self-sufficiency. One of the fundamental challenges is plant phase classification task which could help to estimate yield before harvest. In contrast to the conventional field survey method which required a large amount of human and capital resources, we explore a more scalable, inexpensive and real-time method using publicly available remote sensing data, i.e. Landsat-8 satellite. Landsat-8 provides rich spatiotemporal features which could support the detection of numerous vegetation and crop-related indices. However, to accurately classify the plant phase, the existing features require additional spectral pattern from different seasons. We found out the existence of temporal autocorrelation among features of food crops plant phase. The aim of this study is to propose a supervised random forest for features engineering to select the best multitemporal features for the classification of rice plant phase. In this study, we focus on the rice plant phase classification in Banyuwangi Regency, Indonesia as a case study. The ground truth data are the monthly kerangka sampel area (KSA) of average rice plant phase at the regency level which officially released by BPS-Statistics Indonesia. The experimental result shows the accuracy of 0.573 with one temporal feature. Furthermore, incorporating four consecutive temporal features gives higher accuracy gain to 0.727 which shows the temporal autocorrelation. Based on the extensive evaluations, our findings and contributions in this study include: (1) insight to capture the temporal autocorrelation to increase the model accuracy (2) a machine learning classification model which is not sensitive to multicollinearity. Our proposed method provides the potential benefit for the government and statistical agencies towards a more scalable agricultural survey.

Keywords: remote sensing, crops classification, temporal autocorrelation, landsat-8, random forest

1. INTRODUCTION

In many developing countries, food security and self-sufficiency have been declared as one of their national goals. Among many challenges to support such goals, food phase classification task is fundamentally potential to help the estimation of yield before harvest. Food crops, such as rice has been extensively cultivated to feed the world growing population. In Indonesia, with a population of more than 260 million people depending on rice as the staple food, rice self-sufficiency are also an important goal. Plant phase monitoring has been conducted by BPS-Statistics Indonesia based on conventional field surveys, namely Kerangka Sample Area (KSA)¹. In 2017, KSA firstly introduced the rice plants phase monitoring data. The field survey data of KSA is built on the powerful statistical method for sampling and estimation. Despite its high quality estimation, there are at least two major disadvantages of KSA field survey (1) cost-sensitive, as it requires the increase of sample size for larger area estimation which would depend on more human and capital resources, (2) lack of consideration to individual variations in non-sample areas, given that the estimation is merely based on the selected samples.

On the other hand, current development in remote sensing data and technology provides the potential of a more scalable, inexpensive and real-time benefit. Over the past decade, a huge amount of worldwide coverage remote sensing images has been publicly accessible. In particular, Landsat 8 satellite provides high-resolution multi-spectral datasets which include rich information of agricultural vegetation growth which freely available. This enable us to analyze the vegetation growth from the past to the present and then predict changes over time^{2,3,4}.

In this paper, we focus on the classification of rice plant phase in Banyuwangi Regency, Indonesia as our case study. In order to select the finest multitemporal features, we suggest a supervised random forest technique to create features

engineering. Based on the comprehensive evaluations, our contributions in this paper lays on two major points. First, we provide insight to capture the temporal autocorrelation to improve the accuracy of the model⁵. Second, we propose a machine learning-based classification model that is not susceptible to multicollinearity. Machine learning based method has gained powerful achievement in many supervised and unsupervised task^{6,7}.

2. METHODOLOGY

2.1 Rice phenological and spectral pattern

Rice has several growing phases in one planting season, namely land preparation (before planting), early vegetative, late vegetative, early generative (reproductive), final generative (ripening), and harvesting. During land preparation, paddy fields are flooded with the lots of water. Then in the initial vegetative phase, the puddle decreases and the greenness of the rice begins to be planted. The greenness of rice plant will increase in the final vegetative phase, then the yellow panicles begin to grow when entering the generative phase. During the harvest phase, the land becomes dry, brown and open^{8,9}.

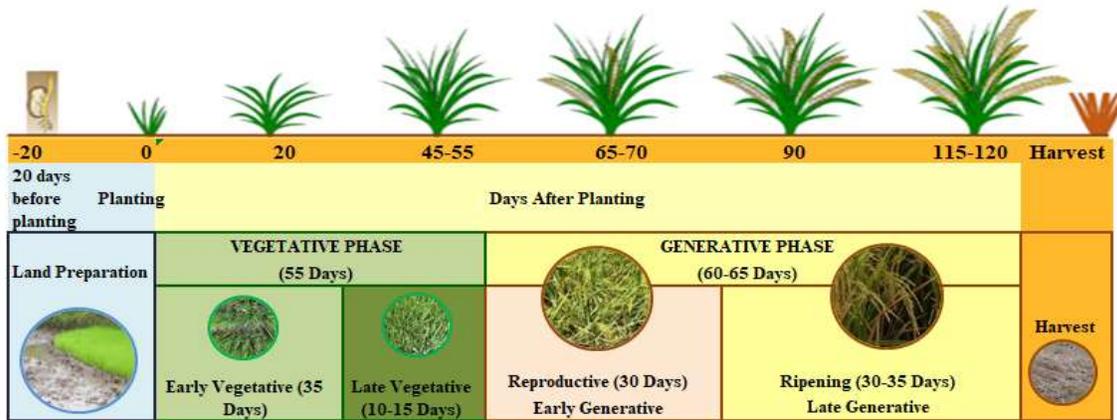


Figure 1. Phenology Rice Plant Phase

The venology of rice plant phase causes a spectral pattern of vegetation indexes from remote sensing in one planting period. The greenness of the plants is illustrated by the NDVI index and EVI, the level of standing water is indicated by NDWI, whereas open land is obtained from NDBI.

$$NDVI = \frac{NIR(b5) - (b4)}{NIR(b5) + (b4)} \quad (1)$$

$$EVI = \frac{NIR(b5) - (b4)}{1 + NIR(b5) + 6 * (b4) - 7,5 * (b2)} \times 2, \text{ if Red}(b4) < NIR(b5) \text{ or Blue}(b2) < Red(b4) \quad (2)$$

$$EVI = \frac{NIR(b5) - (b4)}{0,5 + NIR(b5) + (b4)} \times 1,5, \text{ other than Red} < NIR \text{ or Blue} < Red \quad (3)$$

$$NDWI = \frac{(b3) - SWIR1(b6)}{(b3) + SWIR1(b6)} \quad (4)$$

$$NDBI = \frac{SWIR1(b6) - NIR(b5)}{SWIR1(b6) + NIR(b5)} \quad (5)$$

During the land preparation phase, the EVI value is still low and NDWI is high, then the value of EVI starts to rise in the early vegetative phase and NDWI starts to fall. In the final vegetative phase, the EVI value rises steeply and reaches the maximum EVI peak, then the value drops again in the initial generative and final generative phases to harvest. At harvest time NDWI is very low and NDBI peaks occur as shown in Figure 2. Figure 2 is obtained from the extraction of certain samples.

In other types of food plants, this spectral pattern can be different. For example in corn or soybeans, variations in the value of NDWI will not be as high as rice plants. In non-paddy fields such as houses or forests, the values of EVI, NDVI,

NDWI, and NDBI tend to be stable. The different spectral patterns when puddles occur, the greenness begin to high, and the land become dry associated with each stage of rice plant phase, allows features of Landsat 8 imagery can be used to distinguish between rice and non-rice classes and rice plant phase classes.

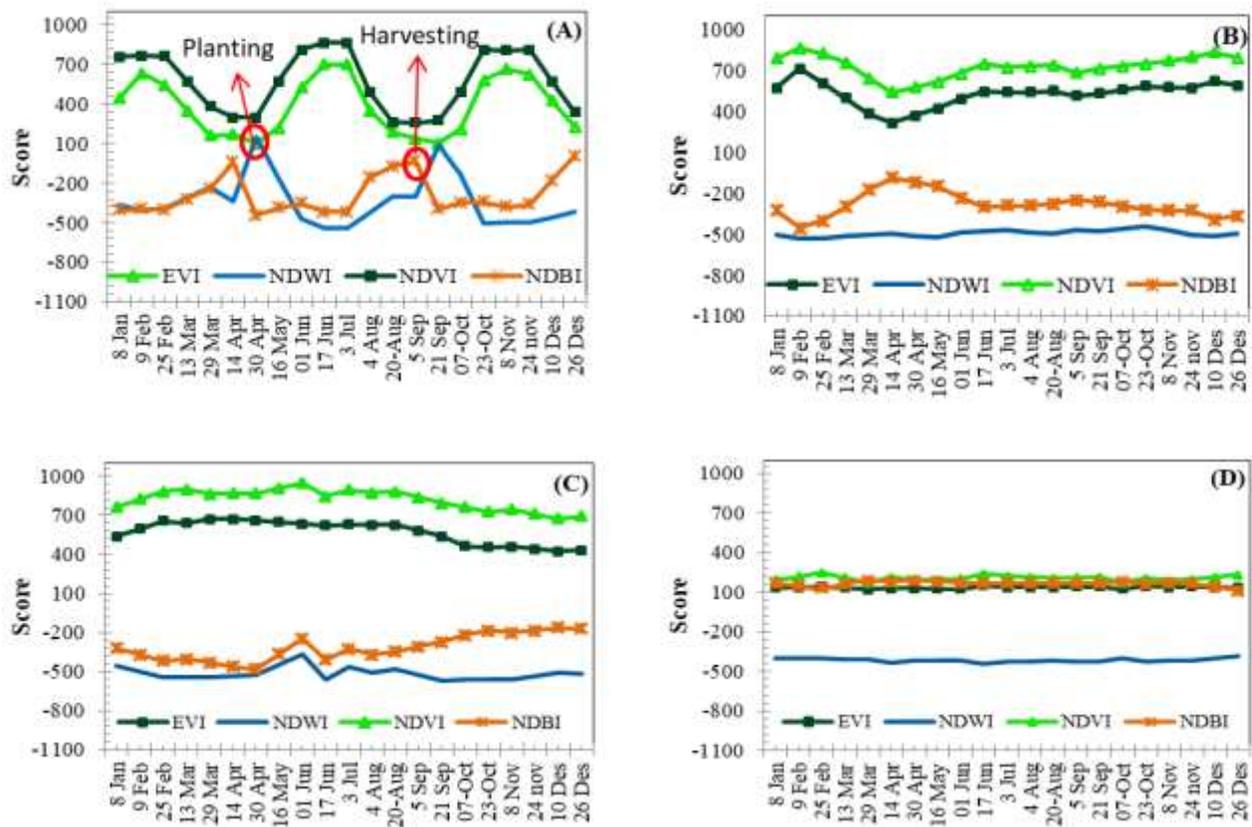


Figure 2. Examples of Spectral Patterns EVI, NDVI, NDWI, NDBI From 21 Landsat Periods in 2018. (a) Lowland Rice Plants, (b) Non-Rice In Paddy Fields, (c) Forests (Non Paddy Fields), and (d) Houses (Non Paddy Fields)

2.2 Random Forest

Random forest is the development of CART by applying bagging and random feature selection to CART, which randomly selects several features in each iteration¹⁰. Trees are produced as much as iterations so that the random forest resembles a forest. The classification decision is taken from the most votes among all trees. CART way to produce trees is to minimize errors in choosing which features are used as separators. This causes a lot of similarity in tree structure produced by CART, causing a high correlation between predictors. The combination of predictions from various models in the ensemble will work well if the predictions from the sub models are not related or at least the correlation is very weak. Stages in the Random Forest:

1. Bootstrapping, taking a sample of size n with replacement
2. Subsetting, choosing p features as a barrier to build a tree, $p < q$ with q are all the features available. Selection of size m based on the smallest value of $MSE = \left\{ \frac{1}{2} |\sqrt{q}|, |\sqrt{q}|, 2|q| \right\}$. But it is possible to predict better results if using different m values.
3. Repeat steps 1 and 2 until a tree is obtained
4. The final prediction is the estimated average of the k trees.
5. The selection of a stable tree based on the value of $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$, where \hat{y}_i is the average prediction of i and n is the number of observations.

2.3 Data

The data used in this study include label data and feature data. Label data comes from the kerangka sample area (KSA) survey, while feature data was extracted from Landsat 8 images based on rice phenology information and spectral pattern. The KSA data that we use from January until December 2018, while Landsat-8 from November 2017 until December 2018.

KSA is a monthly rice growing phase survey conducted by BPS. KSA is carried out to estimation and prediction of harvested area data every month in Indonesia. The estimation are calculated based on generalized observations of each sample. KSA is implemented based on the dot sampling method, with a sampling unit is an an area measuring 300 m x 300 m. Observation of the rice plant phase is conducted every month by the survey officer at the midpoint coordinates of the selected sample. In other words, each selected location will be observed every month.

KSA consists of 6 classes of rice plant phase (early vegetative, late vegetative, generative, puso, harvest, fallow/land preparation) and 2 classes of non-rice (non-rice in rice fields, non rice fields). Non-rice in rice fields classes are included because there is a possibility of rice fields to be planted with rice or non-rice plants. Non-rice fields classes are included because there is a possibility of non-rice fields on KSA sample. This happens because the division of the entire area into an area measuring 300 m x 300 m is carried out using Landsat 8 imagery, so there is a possibility that the area assigned as rice field is apparently not rice field .

Generative classes in KSA have a long range, so in this study, generative classes are divided into early generative and late generative. If the observations at a sample point KSA in two consecutive months are generative phases, the observations in the second month are re-categorized as late generatives. However, the number of observations of this class was only a little. Other notes, because the puso rarely occurs, then from all of the selected samples, puso sample were not obtained. This causes puso to still be listed as a class for rice plant phase, but there is not puso samples in this study. Details of class and sample size can be seen in Table 1

Table 1. Class labels and number of samples

Code	Class	Explanation	Sample size
1	Early vegetative	Termination to seeding stage (1-35 days)	52
2	Late vegetative	Tillering to stem elongation (35-55 days)	39
3	Reproductive/early generative	Panicle initiation to booting and heading to flowering (55-80 days)	22
4	Harvest	When rice is being or has been harvested	25
5	Bare/ land preparation	After harvest or before transplanting	15
6	Puso	In the event of a pest or disaster, so that rice production is less than 11% of normal	-
7	Non rice in paddy fields	If paddy fields that are not cultivated for rice plants	70
8	Non paddy fields	If the observation on non-paddy fields such as forests, road settlements, water bodies, roads, etc.	136
9	Ripening/late generative	Milky and doughy and mature grain stage (80-105 days)	4

The set of features obtain from landsat 8 images downloaded using the cloud on the Google Earth Engine (GEE). the downloaded image criteria is geometric correction, atmosphere correction, and cloud cover less than 50%. The images taken from November 5, 2017 to December 26, 2018 period. The extracted features are bands 1 to 7, EVI, NDVI, NDWI, NDBI, multitemporal from t period until t-3 period. For example, if the KSA label is November, then period t for features is the image of November 24, 2018, t-1 is the image of November 8, 2018, t-2 is the image of October 23, 2018, and t-4 is the image of October 7, 2017. Based on the value of 4 period features, the variance from vegetation index, maximum EVI, and minimum EVI were calculated. The variance of the vegetation index, maximum EVI, and minimum EVI is used to capture the phenology pattern of rice plants or non rice plant and distinguish the rice plant phase. The

features used in modeling are period t bands 1 to 7, period t vegetation indexes, vegetation index variants, maximum EVI, and minimum EVI.

2.4 Classification Approach

Rice is usually planted in the paddy fields both irrigated, non-irrigated, or rainfed paddy fields. The paddy fields in one year are usually planted 1 or 2 times rice plants, sometimes 3 times, interspersed with other food crops if 1 or 2 times. Rice is also possible to be planted in the moor with three assumptions, which are planted with an upland system, land use conversion because the moorings are near the paddy fields, and there is a narrow paddy field mixed with moor so that it is not mapped as a small-scale paddy field (1: 100000) and recorded as a moor grid.

The rice planting methods that allow the same location to be planted alternately with other food crops make two important criteria that the rice plant phase classification model must have are (1) the ability to distinguish between types of rice plants and other plants and (2) the ability to distinguish rice plant phases in one planting period based on rice phenology. These criteria require the classes of rice plant phase model should consist of the rice plant phase, non-rice in paddy fields, and non-paddy fields. An extensive experiment will be conducted to investigate (1) features that distinguish rice and non-rice classes, (2) features that distinguish the rice plant phase class. The classification stages are detailed as follows, as shown in Figure 3.

Step 1: Download Landsat-8 from November 2017 until December 2018 from google earth engine (GEE). Landsat 8 is downloaded with certain cloud cover criteria and has been done radiometric and geometric corrections. Extract band 1 until band 7 from the multitemporal Landsat-8 at the coordinates of the KSA sample and add the KSA observations as label data. Calculating vegetation index features, vegetation index variants, maximum and minimum EVI. Extracting and manipulating data is using R software

Step 2: Exploring vegetation index patterns and vegetation index variants in 4 periods of each class, whether the vegetation index pattern according to theory of rice phenology in one planting period

Step 3: Splitting data into training and testing data. making a classification model using training data

Step 4: Evaluating prediction accuracy using testing data

Step 5: Exploring feature importance

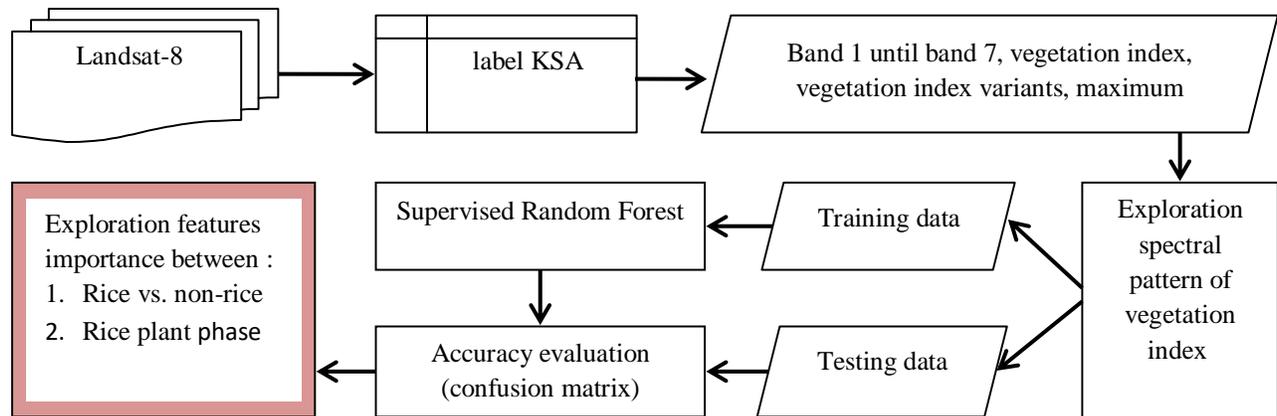


Figure 3. Stages of Rice Plant Phase Classification

3. RESULT

We plot EVI, NDVI, NDBI, and NDWI from 4 period in each class to check whether the vegetation index pattern is according to phenology rice plant phase in one planting period or is not. The Landsat-8 period is 16 days and the KSA period is 1 month, so there can be many combination of different classes between periods. For example when the period t is a early generative class, observations can occur in the early generative range of the end, middle, or beginning. If the observations are in the middle generative range, then the t-1 period is also the early generative class, the t-2 period is the late vegetative class, and the t-3 is also the late class. If the observations are in the beginning generative range, then the t-1 period is the late vegetative, the t-2 is the late vegetative, and the t-2 is the early vegetative. This can cause different

temporal spectral patterns in the same class, so the data becomes non-linear and machine learning is needed to automatically determine class boundaries for the classification process. The overall temporal spectral pattern of each class can be seen in Figure 2

As shown in Figure 4, in the initial vegetative class, the NDWI value in period t is quite high while EVI, NDVI, and NDBI are still low. In this early vegetative class, the period t-1 is likely to be land preparation, period t-2 and t-3 is harvest. It is seen that the NDBI value at t-3 is the highest compared to other periods, the EVI, NDVI, and NDWI values are also the lowest. This plot in the early vegetative class showed a data pattern in accordance with the phenology theory of the rice plant phase.

Variation in vegetation index from 4 periods is quite high in the rice plant phase class and relatively stable for non rice in rice fields rather class and non rice fields class. From this plot it is estimated that the variance value of the vegetation index will be a differentiator of the rice plant phase class with non-paddy in rice fields class and non rice fields.

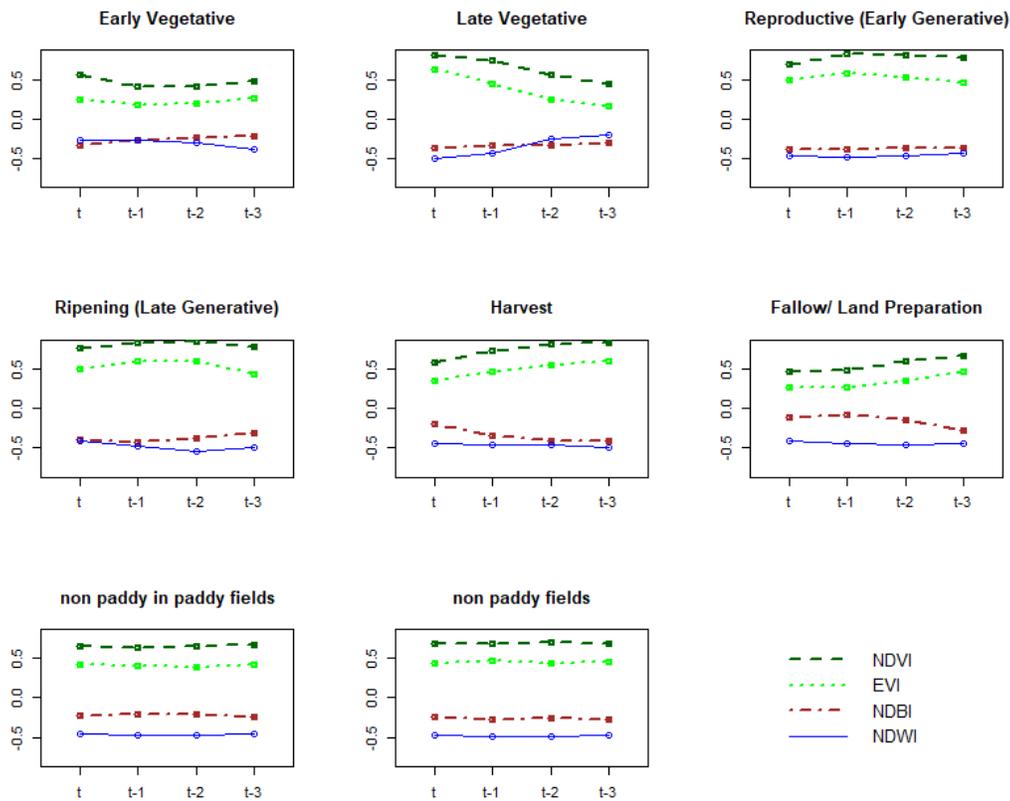


Figure 4. Spectral Pattern NDVI, EVI, NDBI, NDWI From 4 Period In Each Class

In the rice plant phase, it seems that the value of the vegetation index between classes looks different. For example, in the late vegetative class, the value of EVI and NDVI is higher than early vegetative, but the NDWI and NDBI values are lower. It is estimated that in the rice plant phase class, the value of the vegetation index will be more differentiating between classes than the vegetation index varian from 4 period. Furthermore, the classification model is carried out using 17 features, that is band 1 to 7 period t band, vegetation index (EVI, NDVI, NDBI, NDWI) period t, vegetation index variants, maximum EVI, and minimum EVI. Accuracy and misclassification of random forest models can be seen in Table 2.

Table 2. Accuracy of Classification Model

Prediction	Reference								Accuracy
	1	2	3	4	5	7	8	9	
1	19	0	1	0	2	0	0	0	86.36
2	1	9	2	0	0	0	0	0	75.00
3	1	0	2	0	0	0	0	0	66.67
4	0	0	0	5	1	2	0	0	62.50
5	0	0	0	1	0	0	0	0	0.00
7	2	1	1	1	0	9	2	0	56.25
8	2	1	1	1	3	4	36	0	75.00
9	0	0	0	0	0	0	0	0	0.00
Accuracy	76.00	81.82	28.57	62.50	0.00	60.00	94.74	0.00	
Overall accuracy : 0.7273					Kappa : 0.6437				
Overall accuracy if use one temporal features = 0,5727									

Random forest is an ensemble machine learning that is not sensitive to multicollinearity because each tree is produced using only part of the randomly selected features¹⁴. Therefore we can classify using several combinations of features to see the effect of a combination of features on the accuracy model and misclassification. For example when classification is only using EVI features, the model accuracy of 0.236. Misclassification occurs between rice and non-rice classes or between rice phase classes. When the classification was using EVI, NDVI, NDWI, and NDBI, the model accuracy increased to 0.491. Misclassification decreases especially in the non-rice fields class which is classified as non-rice in rice fields and the early vegetative class which is classified as another rice plant phase.

The accuracy of random forest model if use one temporal fitur (band 1 until band 7 and vegetation index), on the other hand we doesn't use fitur vegetation index variants, maximum EVI and minimum EVI from 4 period, the accuracy model only 0.573. The accuracy if we use temporal fitur is 0.727. So There is increasing of accuracy model if we use temporal features compare to the model if we doesn't use.

The highest accuracy is in class 8 (non rice fields) and the lowest accuracy is in class 5 (bare / land preparation). In class 5, all samples were classified as wrong. This indicates that the features used in the random forest have not been able to capture the class 5 pattern. This class 5 consists of fallow conditions after harvest (open and dry land) and land preparation conditions (flooded with water). As shown in Table 2, the class 5 is still classified incorrectly into the initial vegetative class, harvest, and not rice fields. The hallmark of the early vegetation class is the amount of standing water and the characteristics of the harvest class is open dry land. Both are similar to conditions in fallow class. Experiments using other features that are able to distinguish bare/land preparation from the early vegetative and harvesting need to be done to improve the accuracy of the model.

The accuracy of the early generative class is also still low, 28.57%. This early generative class is still misclassified into the early vegetative class, late vegetative, non-rice in paddy fields, and non rice fields. This misclassification occurs from all of the rice plant phase class become non rice in rice fields class and non rice field class. On the other hand, non rice in rice fields are still misclassified as harvest classes. This misclassification if it has not been able to be lowered will cause over or under estimate of agricultural yields. If many sample from the non rice in paddy field class is predicted as a harvest, then the estimated crop will over estimate. If many sample from the rice plant phase class is predicted as not rice or non-paddy fields, then the estimation of agricultural yields will under estimate.

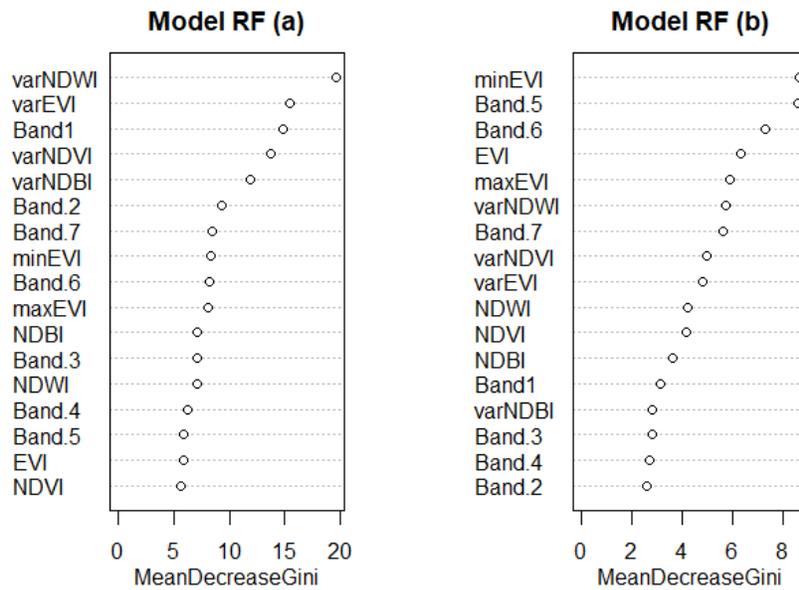


Figure 5. Important features from random forest model that (a) distinguish rice and non-rice classes, (b) distinguish the rice plant phase class

To provide an overview of important features, a random forest model is made into two models with different classes. In model A, there are 3 classes that are used, namely rice, non-rice in rice fields, and non-rice fields. We call model A the model that distinguishes rice vs. non-rice classes. Model B uses 7 all classes of rice phase, early vegetative, late vegetative, early generative, late generative, harvest, fallow/land preparation, and puso. We call model B as the model that distinguishes the rice plant phase class. Important features on both models can be seen in Figure 5.

Important features that can distinguish between rice and non-rice classes are vegetation index variant from 4 period and band 1. Vegetation index values in the paddy class are more varied than non paddy classes. NDWI values in the paddy class ranged from 0.0 to -0.6 and in the non paddy class ranged from -0.4 to -0.6. This NDWI shows the existence of standing water, so that the value of NDWI is high during the initial planting and decreases until harvesting. EVI values in the paddy class ranged from 0.1 to 0.7 and in the non paddy class ranged from 0.3 to 0.6. This EVI shows the level of greenness, so in rice classes, the value is low when the initial planting, rises to the maximum peak between the late vegetative phase with the early generative and decreases until harvesting. In non rice classes, there are a number of samples such as horticultural crops in paddy fields, forests, and shrubs that have high EVI values in the range of 0.6 and for 4 periods also remain in that range. There are also some samples such as houses and buildings that have low EVI values in the range of 0.3 and for 4 periods remain in that range.

Something surprising, band 1 was also a distinguishing feature of rice and non-rice classes. Band 1 of this Landsat 8 image is an aerosol or coastal band that is usually applied to marine monitoring. there are two assumptions why band 1 is one of the best differentiating features between rice and non-rice classes. First, maybe this is because some KSA samples are in non-rice fields such as ponds and lakes, so band 1 is useful for recognizing that sample. Second, this shows that band 1 can be explored as an index for plant monitoring.

An important feature of model A shows that if we are going to differentiate between paddy and non-paddy land, or to classify rice plants in a large area that allows observations of non-rice plants, it is better to use variations of vegetation indexes over several periods rather than EVI. If we classify rice and non-rice using only EVI, there is a possibility of misclassification, especially in samples with high EVI values such as non-rice plants in rice fields, forests, shrubs, fields covered with grass, etc.

Model B is used when the classification of rice plants is carried out on land that is always planted with rice. It is assumed that there are no non-rice plants planted on the land. From this model B we get information about the features that distinguish between phases in the rice class. The important features that differentiate between classes are minimum EVI from 4 period, band 5, band 6 EVI, and maximum EVI from 4 period. The EVI index, which is usually used to plant phase classification manually, turns out to be at number 4. The values of band 5 and band 6 have been used to calculate the NDBI value, but it turns out the original values of bands 5 and 6 are better at distinguishing the rice plant phase class. It is estimated that bands 5 and 6 can still be used to a new index that captures the value of open and dry land.

Based on the 5 best features of differentiation between rice phase classes, the characteristics of the early vegetative class are those 5 features are of low value. The characteristics of late vegetative class are the minimum EVI value is low, band 6 value is medium, and band 5, EVI, and the maximum EVI value is high. The characteristics of early generative class is the medium value of band 6 and the high value of other 4 features. The characteristics of harvest class is the scores are low on EVI, medium on band 5, and high on minimum EVI, band 6, and maximum EVI.

4. CONCLUSIONS

In this study, we performed extensive experiments using supervised random forest method to investigate best multi-temporal features for the classification of rice plant phase. From the results, several important points can be summarized as insight for improving the classification model of rice plant phase, including (1) classification accuracy increases when temporal features are included in the model, (2) the highest accuracy is in class 8 (non-paddy fields) because the spectral pattern of this class is not quite varied, (3) the lowest accuracy is in class 5 (bare / land preparation), (4) we need to do further experiments with other features that can distinguish bare/land preparation from the early vegetative and harvesting to improve the accuracy of the model in class bare/land preparation and other features that can distinguish generative from late vegetative to improve the accuracy of the model in early generative class, (5) features engineering have an important role for accuracy improvement because different classification classes require different set of features.

REFERENCES

- [1] BPS-Statistics Indonesia, [pedoman pelaksanaan uji coba sistem kerangka sampel area (KSA)], BPS-Statistics Indonesia, Jakarta (2015).
- [2] JRC Scientific And Policy Reports, [Combining crop models and remote sensing for yield prediction : concepts, applications and challenges for heterogeneous , smallholder environment], European Commission, Roma (2012).
- [3] Asgarian, A., Soffianian, A., and Pourmanafi, S, "Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: a case of central iran using multi-temporal landsat 8 imagery," *Computers and Electronics in Agriculture* (127), 531–540 (2016).
- [4] Clark, M. L, "Comparison of simulated hyperspectral hyperspectral and multispectral landsat 8 and sentinel-2 imagery for multi-seasonal, regional land-cover mapping," *Remote Sensing of Environment* (200), 311–325 (2017).
- [5] Triscowati, D.W., Sartono, B., Kurnia, A., Domiri, D.D., Wijayanto, A.W., "Classification of Rice Plant Phase Using Supervised Random Forest Based On Multitemporal Data Landsat-8 Satellite," *Proc Sinas Inderaja LAPAN* (2019)
- [6] Wijayanto, A. W. & Takdir. Fighting cyber crime in email spamming: An evaluation of fuzzy clustering approach to classify spam messages. in *2014 International Conference on Information Technology Systems and Innovation (ICITSI)*, 19–24 (2014).
- [7] Wijayanto, A. W. & Murata, T. Flow-Aware vertex protection strategy on large social networks. in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2017* (2017).
- [8] Dirgahayu, D., Parsa, I. M., Silvia, Hartini, S., Budoyo, S., Indriawan, K., Ernawati, [Litbang pemanfaatan data penginderaan jauh untuk pemantauan pertumbuhan tanaman padi di lahan sawah (studi kasus pulau kalimantan)], Pusfatja LAPAN, Jakarta (2015)
- [9] Kontgis, C., Schneider, A., Ozdogan, M., "Mapping rice paddy extent and intensification in the vietnamese mekong river delta with dense time stacks of landsat data," *Remote Sensing of Environment* (169), 255–269 (2015).

- [10] Breiman, L., "Random forest," *Machine Learning* (45), 5–32 (2001).
- [11] Ghimire, B., Rogan, J., Galiano, V. R., Panday, P., Neeti, N., "An evaluation of bagging , boosting , and random forests for land- cover classification in cape cod, massachusetts, USA," *GIScience & Remote Sensing* (49), 623–643 (2012)
- [12] Jiang, Z., Huete, A.R., Didan, K., Miura, T., "Development of a two-band enhanced vegetation index without a blue band," *Remote Sensing of Environment*, 13–24 (2008).
- [13] Millard, K., Richardson, M.," On the importance of training data sample selection in random forest image classification: a case study in peatland ecosystem mapping," *Remote Sensing* (7), 8489–8515 (2015).
- [14] Gregorutti, B., Michel, B., Pierre, P.S.," Correlation and variable importance in random forests," *Stat Comput* (27), 659–678 (2017)
- [15] Du, P., Samat, A., Waske, B., Liu, S., Li, Z., "Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features," *ISPRS Journal of Photogrammetry and Remote Sensing* (105), 38–53 (2015).
- [16] Belgiu, M., Dragut, L., "Random forest in remote sensing: a review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing* (114), 24–31(2016).
- [17] Peerbhay, K.Y., Mutanga, O., Ismail, R., "Random forests unsupervised classification: the detection and mapping of solanum mauritanum infestations in plantation forestry using hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (vol. 8, no. 6), 3107-3122 (2015).
- [18] Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K, Congalton, R.G., Yadav, K., Huete, A., "A 30-m landsat-derived cropland extent product of australia and china using random forest machine learning algorithm on google earth engine cloud computing platform," *ISPRS Journal of Photogrammetry and Remote Sensing* (144), 325–340 (2018).