CLASSIFICATION OF RICE-PLANT GROWTH PHASE USING SUPERVISED RANDOM FOREST METHOD BASED ON LANDSAT-8 MULTITEMPORAL DATA

Dwi Wahyu Triscowati[1][*][2], Bagus Sartono[2], Anang Kurnia[2]
Dede Dirgahayu[3], Arie Wahyu Wijayanto[4]
[1]BPS of Banyuwangi Regency
[2]Departement of Statistics IPB University
[3]Indonesian National Institutes of Aeronautics and Space (LAPAN)
[4]BPS of Banyuwangi Regency
*e-mail: trisco@bps.go.id

Abstract. Data on rice production is crucial for planning and monitoring national food security in a developing country such as Indonesia, and the classification of the growth phases of rice plants is important for supporting this data. In contrast to conventional field surveys, remote sensing technology such as Landsat-8 satellite imagery offers more scalable, inexpensive and real-time solutions. However, utilising Landsat-8 for classification of rice-plant phase required spectral pattern information from one season, because these spectral patterns show the existence of temporal autocorrelation among features. The aim of this study is to propose a supervised random forest method for developing a classification model of rice-plant phase which can handle the temporal autocorrelation existing among features. A random forest is a machine learning method that is insensitive to multicollinearity, and so by using a random forest we can make features engineering to select the best multitemporal features for the classification model. The experimental results deliver accuracy of 0.236 if we use one temporal feature of vegetation index; if we use more temporal features, the accuracy increases to 0.7091. In this study, we show that the existence of temporal autocorrelation must be captured in the model to improve classification accuracy.

Keywords: *rice-plant classification, temporal autocorrelation, temporal features engineering, random forest, Landsat-8*

1 INTRODUCTION

Food-crop monitoring is important to answer the second goal of the SDGs, that is zero hunger. Rice is a vital commodity in Indonesia's food security programme, and achievement of this goal encourages rice-plant monitoring to support food security. Rice-plant monitoring has been conducted by BPS-Statistics Indonesia using conventional surveys based on framework sample areas (FSA)/ kerangka sampel area (KSA). Rice-plant phase data are collected every month for the selected KSA sample.

In addition to conventional data, free remote sensing data from Landsat-8 imagery can be obtained quickly and can be used to monitor rice-plant phase. This remote sensing data has been utilised in various fields, such as agriculture, where remote sensing data is used for rice-growth models (Parsa, Dirgahayu, Manalu, Carolita, & Harsanugraha , 2017), land cover classification (Kussul et al., 2017; Tong et al., 2018), crop type

classification (Azar et al., 2016), crop yield estimation (You, Li, Low, Lobell, & Ermon, 2017), and paddy rice mapping (Zhang, Zhang, & Zhang, 2018; Qiu, Lu, Tang, Chen, & Zou, 2017; Guan, Huang, Liu, Meng, & Liu, 2016).

Landsat-8 provides rich spatio-temporal features to support the detection of vegetation and plant-related indices. However, when analysis is carried out on all of the pixels in the images, the amount of data becomes large and unstructured. For this type of data, machine learning is recommended, because it will be difficult to devise models manually. It is also necessary to add temporal spectral patterns from one season to reduce the misclassification which may occur. This spectral pattern occurs because the vegetation index value of a period is influenced by the previous period, and this indicates the existence of temporal autocorrelation in features that must be treated to improve model accuracy.

In contrast to conventional field surveys that require large amounts of human and capital resources, we explore more scalable, inexpensive, and real-time methods using publicly available remote sensing data. In this study, we propose a supervised random forest method for features engineering to select the best multitemporal features for the classification of rice-plant phase. Random forest is a machine learning method that is not sensitive to multicollinearity. By using random forest, temporal features engineering can be derived as far as possible, from which the best features are selected using the variable importance plotting (varimplot) function.

In this study, we focus on the classification of rice-plant phase in Banyuwangi Regency, Indonesia, as a case study. The ground truth data are

the monthly KSA data for rice-plant phase at regency level officially released by BPS-Statistics Indonesia.

2 MATERIALS AND METHODOLOGY
2.1 Location and data

This study was carried out in Banyuwangi Regency, which is one of the 'rice barns' of East Java Province. The distribution of sample locations can be seen in Figure 2-1.
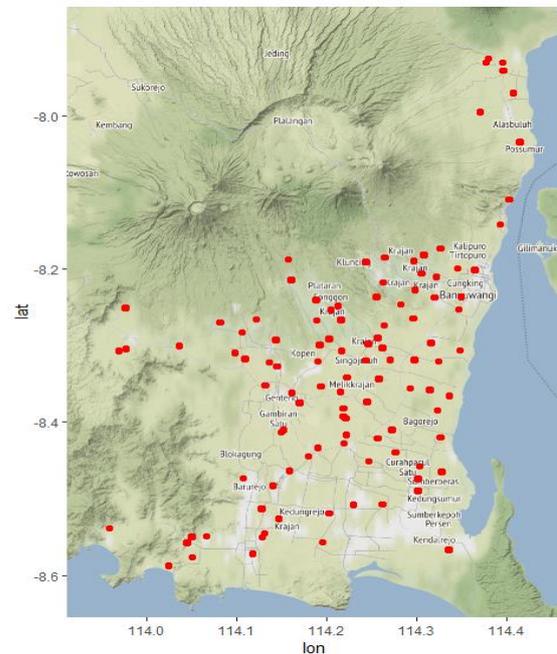

Figure 2-1: Location of KSA Sample

2.2 Label data and features

The label data used in this study follows the randomness principle because it is drawn from the KSA survey. The observations are grouped into nine classes: (1) early vegetative; (2) late vegetative; (3) early generative (reproduction); (4) harvest; (5) bare/land preparation; (6) puso; (7) non-rice in paddy fields; (8) non-rice fields; and (9) late generative (ripening).

The basic features used in this study are band 1 (coastal/aerosol), band 2 (blue), band 3 (green), band 4 (red), band 5 (near infrared [NIR]), band 6 (shortwave infrared [SWIR] 1), band 7 (SWIR 2) from selected Landsat-8, and

four vegetation indices, namely Enhanced Vegetation Index (EVI), Normalised Difference Water Index (NDWI), and Normalised Difference Built-Up Index (NDBI). This vegetation index was chosen based on Pusfatja LAPAN's research, in which EVI, NDWI and NDBI are used to detect rice-plant phase (Dirgahayu et al., 2015).

EVI is used to detect the green level. NDBI captures open/fallow land information which usually has a higher reflection in the SWIR area than NIR. For the rice-plant phases, NDBI is used to detect harvested land that still has crop residues and NDWI detects standing water. NDWI is used for early planting when the rice fields are still flooded.

if red < NIR or blue < red

$$EVI = \frac{band\ 5 - band\ 2}{1 + band\ 5 + 6 \times band\ 2 - 7.5 \times band\ 4} \times 2.5 \quad (2\text{-}1)$$

If other:

$$EVI = \frac{band\ 5 - band\ 2}{0.5 + band\ 5 + band\ 2} \times 1.5 \quad (2\text{-}2)$$

$$NDBI = \frac{band\ 6 - band\ 5}{band\ 6 + band\ 5} \quad (2\text{-}3)$$

$$NDWI = \frac{band\ 3 - band\ 6}{band\ 3 + band\ 6} \quad (2\text{-}4)$$

2.3 Temporal spectral patterns for features engineering

From several studies conducted by Pusfatja LAPAN of satellite imagery used to detect rice-plant phase, it was found that there was a spectral pattern for the EVI values in one rice-planting period (Dirgahayu, Noviar, & Anwar S, 2014). The spectral pattern found that during the early phase of planting it is estimated that NDWI value is high while NDBI and EVI values are low. In the late vegetative period, the NDWI value decreases and EVI rises to a maximum level then the decreases again in the early generative and late generative phases until harvest. At the time of

harvest, it is estimated that NDBI is high and NDWI is low. Besides EVI, vegetation indices that are commonly used to detect greenness are Normalized Difference Vegetation Index (NDVI). Seen that the spectral pattern of NDVI looks like EVI (Figure 2-2).
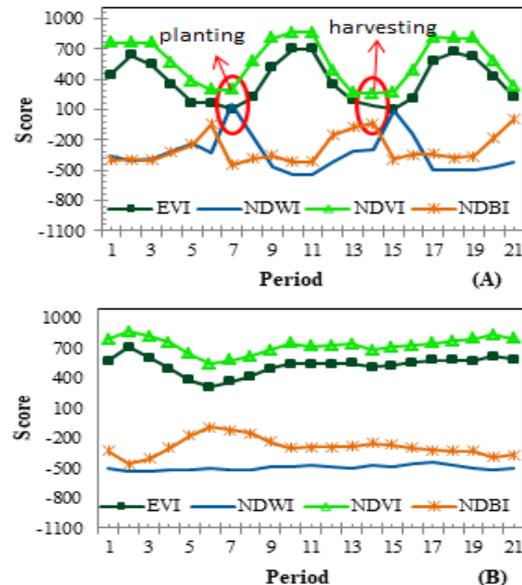

Figure 2-2: Examples of spectral patterns in EVI, NDVI, NDWI, NDBI From 21 Landsat periods in 2018: (A) lowland rice plants, (B) non-rice in paddy fields

2.4 Random forest

'Random forest' (Breiman, 2001) is the development of classification and regression trees (CART) by applying bagging and random features selection which randomly selects several features in each iteration. So many trees are produced by the iterations that the outcome resembles a 'forest'. The classification decision is taken from the most votes among the trees (Hastie, Tibshirani, & Friedman, 2009). The stages in producing a random forest are as follows:

1. Bootstrapping: taking a number of samples with replacement of the training set.

2. Subsetting: the selection of p features as a baffle to build a tree in which the value of p < q with q being all the features that exist. The selection of various p sizes allows better prediction

results if using different p values. The random forest parameter setting in this study uses the default random forest package and the default value for p is √q.

    3.   Repeat steps 1 and 2 until a tree is obtained.

    4.   The final prediction is the majority vote of all the trees.

Random forest method was chosen for this study because it can handle large amounts of data, is not sensitive to multicollinearity and so enables the modelling process to include as many variables as possible, is fast and not sensitive to overfitting, and can choose any explanatory variables that affect the predictions, using the varimplot function.

2.5   Classification approach

The classification framework in this paper is implemented as follows:

Step 1: Using Google Earth Engine, download Multitemporal Landsat 8 data corrected geometrically and radiometrically for cloud cover over 50 per cent and cloud shadow over 30 per cent. Landsat 8 data has NA value if there is cloud cover. There is no special treatment for this because random forest is a method that can handle NA value data. The Landsat 8 data downloaded consists of bands 1 to 7.

Step 2: Extract features for bands 1 to 7 with a grid size of 100m x 100m according to the coordinates of the KSA sample. Then add KSA label data so that the research sample data sets are obtained.

Step 3: Pre-process missing data, NDVI, EVI, NDBI and NDWI. From seven band features and four indices, we create temporal features such as band 1 period t to band 1 period t-4 (if period t is early June, then t-1 is the mid-May period etc.). We create this temporal feature for all of the bands and vegetation indexes. Then we create the feature differences

between periods and features from the regression line. For example, from EVI, we create EVI period t to EVI period t-3; then mean, minimum, maximum, varian, and regression EVI from 4 periods.

Step 4: Perform data exploration to identify whether the temporal spectral pattern of vegetation index matches our theory or not.

Step 5: Split the data into training data and testing data with a proportion of 70:30. Train the random forest model in the training data and then evaluate using the testing data. Perform model experiments using temporal features, feature differences between periods, and features from the regression line. This is intended to capture spectral patterns or overcome the presence of temporal autocorrelation in the vegetation index.

Step 6: Evaluate the classification model of the experimental results. Analyse important features in the rice-plant phase classification model using the varimplot function.

2.6   Challenges in rice-plant classification

One of the main challenges in the classification of rice-plant phase in Indonesia is how to identify Landsat 8 data that can distinguish between rice and non-rice plants. The problem here is that the vegetation index for a period is only a greenness index, without any information about whether the plants are rice or non-rice. For example, in a rice field, there are three planting periods in one year, in that rice is planted in two periods and then interspersed with soya beans or corn once. If the EVI index obtained is 0.4, it cannot be ascertained whether the EVI shows green for rice or for non-rice (e.g. soya or corn). However, if we add information from the previous period it is expected that the information can show whether the plants are rice or

non-rice. The second challenge is that in large areas with small rice fields there are interspersed fields planted with various types of rice and non-rice plants with varying initial plantings. The third challenge is that the types of rice planted in Indonesia are very diverse and there are different spectral patterns among these types of rice.

3   RESULTS AND DISCUSSION

3.1   Data exploration

The exploration of boxplots shows information about band or vegetation index for the t period which is expected to differentiate between classes. It can be seen from the boxplots that NDVI is the most differentiated (class 2: late vegetative) compared to other classes. However, the value of class 2 is quite close to the value of class 9 (final generative). For the EVI index, the late vegetative class is also the most different from the other classes.

This inter-class boxplot shows that the spectral pattern for rice-plant phase is exactly like the rice-growth model developed by the Pusfatja LAPAN. It appears that the NDWI value at the beginning of planting in the early vegetative class is much higher than the value of the other class of NDWI. Meanwhile, the NDBI value that shows open land starts higher in phase 4 (harvest) and then increases further in phase 5 (fallow/land preparation). In class 7 (non-rice paddy fields), the NDWI is quite low and NDBI is quite high. This is consistent with the profile of non-paddy fields in Banyuwangi, which is dominated by corn and soya beans.

3.2   Accuracy and misclassification

Temporal feature experiments were carried out in three stages. The first stage is classification using one period features. The second stage uses two to four period features and the difference between periods to capture information on temporal spectral patterns for one season. The third stage uses the derivative features of the polynomial regression line performed at pre-processing.

The first stage classification was carried out using EVI for one period. When classification only used EVI for one period, the classification accuracy of the testing data was 0.2364. Misclassification often occurred between classes. For example, in the early vegetative class from 25 samples of testing data, only eight were correctly classified.

When the classification used EVI, NDWI and NDBI, classification accuracy increased to 0.4727, because there has been the decreasing misclassification in classearly vegetative class and non-rice fields. All fallow classes were still classified incorrectly.

Subsequent experiments were carried out with the addition of bands 1 and 7. These bands were not used in EVI, NDVI, NDBI and NDWI, so there was still information available that could be derived to capture vegetation signals. The classification results using band 1, band 7 and the four vegetation indices increased accuracy to 0.5636. This means that there is information from bands 1 and 7 that is useful for classifying rice-plant phase.

Based on the suspicion that bands 2 to 6 also contain information relating to plant phase, bands 2 to 6 were added to the model. Classification accuracy increased to 0.5727 with the best features being in band 1, NDWI, band 6, EVI, NDVI, band 5, band 7, NDBI, band 2, band 4, and finally band 3.
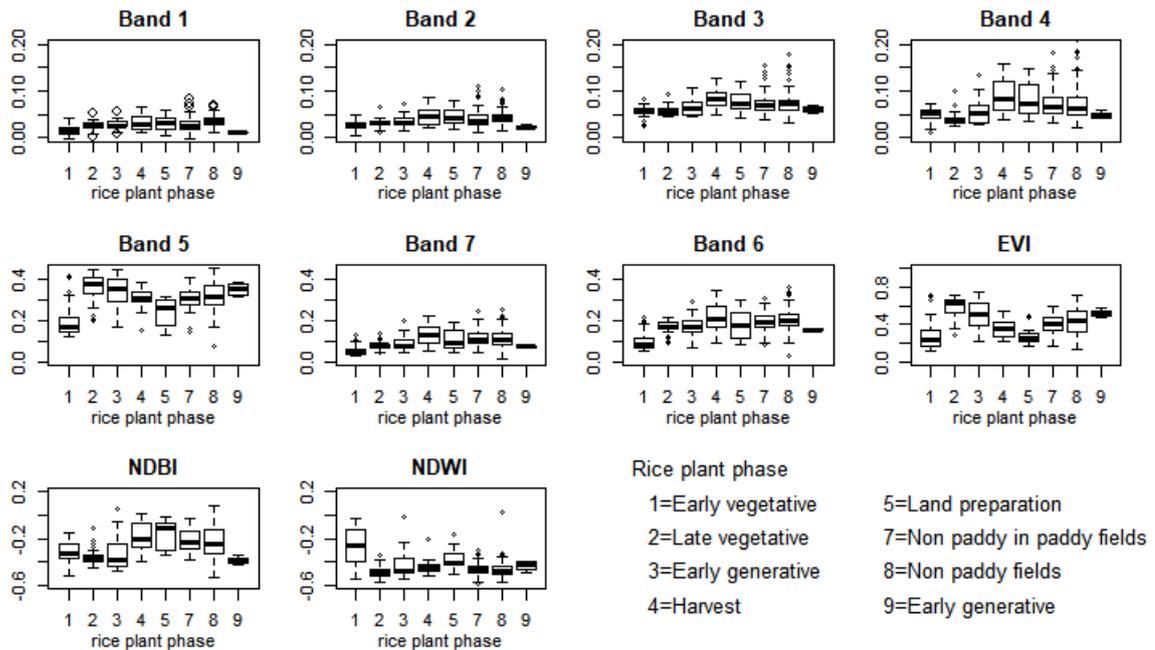
Figure 3-1: Distribution of band and vegetation index value per class period t

Based on the results of the experiment in stage 1 using one period features, the classification model with supervised random forest still cannot distinguish between classes of rice-plant phase. To help the model differentiate phases between classes, the second stage of the experiment using temporal features was carried out.

Table 3-1: Comparison of confusion matrix and classification model accuracy with some features

Features: EVI

Prediction	Reference							
	1	2	3	4	5	7	8	9
1	8	0	0	0	3	0	5	0
2	1	5	5	2	0	2	6	0
3	2	1	1	0	0	1	1	0
4	4	0	0	0	1	4	0	0
5	2	0	0	0	0	2	1	0
7	4	1	1	2	1	4	16	0
8	4	4	4	4	1	2	9	0
9	0	0	0	0	0	0	0	0
Overall accuracy: 0.2364								

Features: EVI, NDWI, NDBI

Prediction	Reference							
	1	2	3	4	5	7	8	9
1	14	0	1	1	2	0	1	0
2	2	5	3	0	0	0	1	0
3	3	0	1	1	0	0	1	0
4	0	0	0	2	1	2	0	0
5	0	0	0	0	0	0	0	0
7	3	1	1	2	0	7	11	0
8	3	5	1	2	3	6	23	0
9	0	0	0	0	0	0	1	0
Overall accuracy: 0.4727								

Features: 3 periods of band and vegetation index, the difference of vegetation index over 4 periods, coefficient regression EVI, maximum and minimum EVI over 4 periods

Prediction	Reference							
	1	2	3	4	5	7	8	9
1	17	0	1	1	1	0	1	0
2	3	10	3	0	0	0	0	0
3	2	0	0	0	0	0	0	0
4	0	0	1	5	1	2	0	0
5	0	0	0	0	0	0	0	0
7	2	1	2	1	2	11	2	0
8	1	0	0	1	2	2	35	0
9	0	0	0	0	0	0	0	0
Overall accuracy: 0.7091								

Classification accuracy increases when temporal features are added into the model. The classification accuracy using features of two periods, bands and vegetation index increased to 0.6364. The highest accuracy is 0.7091. Misclassification in the early vegetative class, late vegetative class, harvest and not-rice fields decreases. For the classification using features for two periods the increasing misclassification of non-rice in rice fields should be noted. This indicates that accuracy increases because the misclassification of some classes decreases, but the model is less able to distinguish between rice and non-rice in rice fields.

In class 5, all samples are not correctly predicted. This shows that the features used have not been able to capture the pattern of class 5. Class 5 consists of fallow after harvest (open and dry land) and land preparation conditions (flooded with water) so that class 5 is still classified as early vegetative, harvested and non-rice field classes. The early vegetative characteristic is the amount of standing water and the harvest characteristic is open, dry land, and both are similar to the conditions in the fallow class. Several experiments have been conducted to identify features to distinguish fallow/land preparation from early vegetative and harvest classes, but such features have yet to be found. If only using one feature, band 5 can classify fallow/land preparation class correctly; however, band 5 is not strong enough to catch fallow/land preparation classes.

Samples in class 3 are still misclassified as class 2. The problem is that because of the 16-day Landsat-8 period, there is the possibility of observing the t period at the beginning of class 3 and the t-1 period in class 2. When these conditions occur, the EVI difference becomes negative as a feature

of class 2. As a consequence, when the EVI difference is added to the model, there is an increasing misclassification of class 3 as class 2.

3.3 The best differentiating features

In general, the best features that distinguish rice and non-rice classes are NDWI variations, EVI variations, band 1, NDBI variations, NDVI variations and band 2. The value of the four vegetation indices varies considerably in rice classes due to growth in one planting period, while the variation in the vegetation index in the non-rice class is relatively stable. Band 1 is usually used in coastal studies. It is possible that in this study, band 1 was one of the best features for distinguishing between rice and non-rice because the observation KSA in the coastal area fell on non-rice classes such as ponds.
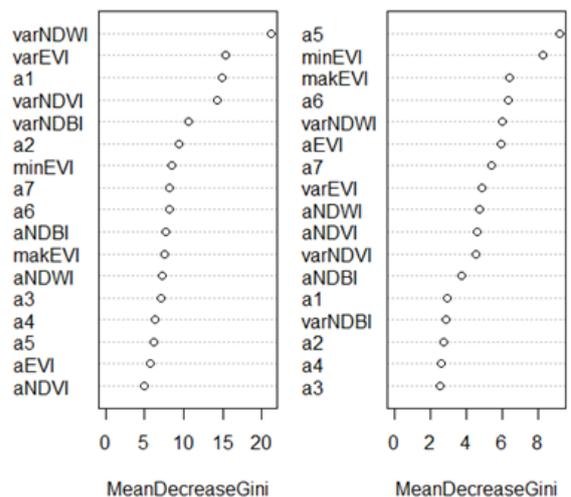

Figure 3-2: Findings of the best distinguishing features of a proposed model with random forest. (Left: rice classification and non-rice classification. Right: classification of rice-plant phase)

The best distinguishing features in the rice-plant phase classes are band 5, maximum and minimum EVI in four periods, band 6, and NDWI variations. Band 5 and band 6 have been

normalised to NDBI. However, apparently, the original value of the two bands is still a better differentiator than NDBI. It is suspected that the NDBI index still needs correction, or another form of normalisation from band 5 and band 6 is needed.

Figure 3-2 shows that classifications for different classes require different types of features. For the classification used to distinguish rice fields and non-rice fields, it is better to use feature variations of vegetation index rather than vegetation index. An example of the classification of rice fields and non-rice fields can be seen in Figure 3-3 and an example of maximum EVI class in rice fields can be seen in Figure 3-4.

being classified as rice fields, trees being classified as rice fields, and vice versa. Another error occurs due to the pixel scale, so that the border areas are classified according to the highest value of the pixel region.

The classification of rice fields in Banyuwangi based on the maximum EVI class can be seen in Figure 3-4. It appears that most of the rice fields in Banyuwangi are in the high EVI class. The highest value of maximum EVI is usually related to rice productivity. This is in accordance to Banyuwangi's position as the fourth-largest producer of rice in East Java. The greenest areas are in the Pesanggaran and Tegaldlimo districts, which are sub-districts with high rice productivity.
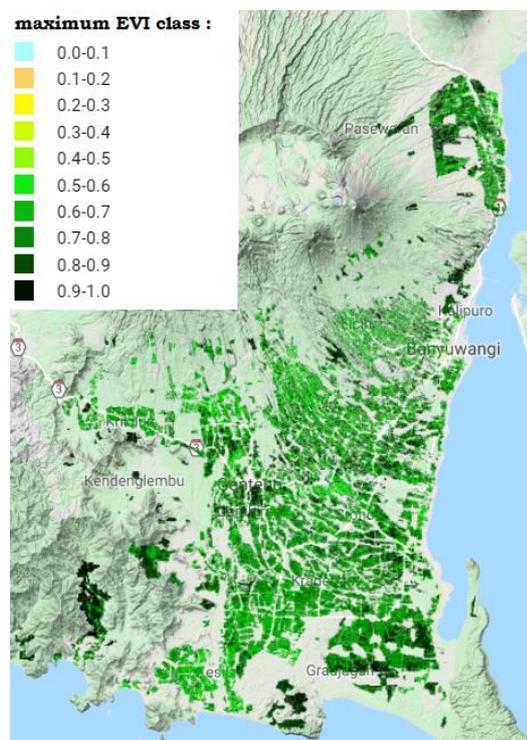

Figure 3-4: EVI maximum class for rice fields

It can be seen from Figure 3-3 that misclassification still occurs. The example of misclassification is houses

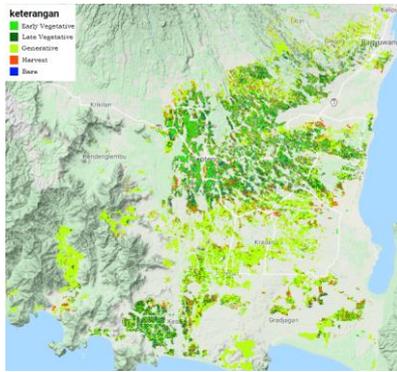Figure 3-3: Rice field and non-rice field classifications

Figure 3-5: Example of rice-plant phase classification from random forest model

4 CONCLUSION

In this study, intensive experiments have been carried out using the random forest method to investigate the best multitemporal features for the classification of rice-plant phase. Some important points can be summarised as information to improve rice-plant phase classification: (1) classification accuracy increases when temporal features are included in the model; (2) the highest accuracy is in classes 1, 2 and 8 because all of these classes have special characteristics distinguishing them from other classes; (3) the lowest accuracy is in class 3 and class 5, with class 3 being still classified into class 2, and class 5 being still misclassified as class 1 and 4, because it consists of dry and wet fallow conditions.

As input for further research, studies could focus on: (1) class 5 being divided into wet fallow and dry fallow classes, or increasing the number of samples and looking for other features that can capture both of these classes; (2) increasing the number of class 3 samples and identifying more sensitive features which could be used distinguish class 3 from classes 2 and 4; (3) developing methods that not only read the maximum and minimum values of vegetation indices in four periods, but also accommodate positions of lag, as it is expected that such a model would be better able to capture temporal spectral patterns and accuracy would be improved; (4) other satellite images with a short period could be added to obtain information about maximum and minimum turning points, so the EVI difference between t-periods and t-1 periods becomes more reflective of the condition of each class.

AUTHOR CONTRIBUTIONS

Classification Of Rice-Plant Growth Phase Using Supervised Random Forest Method Based On Landsat-8 Multitemporal Data. Lead Author: Dwi Wahyu Triscowati; Co-Author: Bagus Sartono, Anang Kurnia, Dede Dirgahayu, and Arie Wahyu Wijayanto.

REFERENCES

Azar, R., Villa, P., Stroppiana, D., Crema, A., Boschetti, M., & Brivio, P. A. (2016). Assessing In-Season Crop Classification Performance Using Satellite Data: A Test Case In Northern Italy. *European Journal of Remote Sensing, 49*, 361–380. doi: 10.5721/EuJRS20164920

Belgiu, M. & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24–31. doi:10.1016/j.isprsjprs.2016.01.011

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32. Retrieved from https://link.springer.com/content/pdf/10.1023/A:1010933404324.pdf

Dirgahayu, D., Noviar, H., & Anwar, S. (2014). Model pertumbuhan tanaman padi di pulau Sumatera menggunakan data EVI MODIS multitemporal. *Deteksi Param. Geobiofisik Dan Diseminasi Penginderaan Jauh* (7):333–343.Dirgahayu, D., Parsa, I. M., Silvia, Hartini, S., Budoyo, S.,

Indriawan, K., & Ernawati (2015). *Litbang Pemanfaatan Data Penginderaan Jauh Untuk Pemantauan Pertumbuhan Tanaman Padi Di Lahan Sawah: Studi Kasus Pulau Kalimantan [R & D utilization of remote sensing data for monitoring rice plant growth in rice fields: Case study of Kalimantan Island].* Jakarta: Pusfatja LAPAN.

Ghimire, B., Rogan, J., Galiano, V. R., Panday, P., & Neeti, N. (2012). An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing, 49*(5), 623–643. doi: 10.2747/1548-1603.49.5.623

Gregorutti, B., Michel, B., & Saint-Pierre, P., (2017). Correlation and variable importance in random forests. S*tatistics and Computing, 27*(3), 659–678. doi: 10.1007/s11222-016-9646-1

Guan, X., Huang, C., Liu, G., Meng, X., & Liu, G. (2016). Mapping rice cropping systems in Vietnam using an NDVI-based time-series similarity measurement based on DTW distance. *Remote Sensing, 8*(19), 1–25. doi: 10.3390/rs8010019

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning (data mining, inference, and prediction).* New York: Springer.

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* 14(5):778–782. doi:10.1109/LGRS.2017.2681128.

Millard, K. & Richardson, M. (2015). 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. https://doi.org/10.3390/rs70708489

Mishra, N. B., Crews, K. A., Miller, J. A., & Meyer, T. (2015). Mapping vegetation morphology types in southern Africa savanna using using MODIS time-series

metrics: A case study of Central Kalahari, Botswana. *Land, 4*, 197–215. doi: 10.3390/land4010197

Parsa, I. M., Dirgahayu, D., Manalu, J.,Carolita, I., & Harsanugraha, K. W. (2017). Uji Model Fase Pertumbuhan Padi Berbasis Citra Modis Multiwaktu Di Pulau Lombok [Test model of multi-time modular image-based rice growthphase in Lombok Island]. J*urnal Penginderaan Jauh 14*(1), 51–64.

Qiu, B., Lu, D., Tang, Z., Chen, C., & Zou, F. (2017). Automatic and adaptive paddy rice mapping using Landsat images: Case study in Songnen Plain in Northeast China. *Science of The Total Environment.* 598:581–592. doi:10.1016/j.scitotenv.2017.03.221.

Teluguntla, P., Thenkabail, P., Oliphant, A., Xiong, J., Krishna Gumma, M., Congalton, R. G., … Huete, A. (2018). 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*(2), 325–340. doi: 10.1016/j.isprsjprs.2018.07.017

Tong, X., Xia, G., Lu, Q., Shen, H., Li, S., You, S., & Zhang, L. (2018). Learning transferable deep models for land-use classification with high-resolution remote sensing images. *NASA Astrophysics Data System (Subject Computer Vision and Pattern Recognition) 7*, 1–25.

You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian process for crop yield prediction based on remote sensing data. *31st AAAI Conference on Artificial Intelligence*, 4559–4565. Zhang, C., Zhang, H., & Zhang, L. (2018). An automated paddy rice extent extraction with time stacks of Sentinel data : A case study in Jianghan Plain, Hubei, China. *7th International Conference on Agro-Geoinformatics (Agro-Geoinformatics) Agustus 2018*, 1–6. doi: 10.1109/Agro-

Geoinformatics.2018.8476119