

LAND-SOIL CHARACTERISTICS FOR MAPPING PADDY CROPPING INTENSITY USING DECISION TREE ANALYSIS FROM SINGLE DATE ALI IMAGERY IN MAGELANG, CENTRAL JAVA, INDONESIA

S. Arjasakusuma^{a,b} , P. Danoedoro^{a,b} , S. Herumurti^{a,b}, Y.A. Nugroho^a, P.A. Aryaguna^{a,b}

^a Remote Sensing Dept., Geographic Information Science Major, Faculty of Geography, Gadjah Mada University, Indonesia

^b Graduate School of Remote Sensing, Faculty of Geography, Gadjah Mada University, Indonesia

Article Info:

Received: 15 Dec 2016

in revised form: 1 May 2017

Accepted: 7 July 2017

Available Online: 30 Oct 2017

Keywords:

Decision Tree, Cropping Intensity, Paddy Field, Land Characteristics

Corresponding Author:

Sanjiwana Arjasakusuma

Gadjah Mada University

Email:

sanjiwana.arjasakusuma@uqconnect.edu.au

Abstract: Paddy field area and its cropping intensity are main information used to measure the crop production and the response of crop to changing climate conditions. Remote sensing technology has been widely used to map cropping pattern of paddy mostly using spectral analysis of multi-date multispectral data of remote sensing. However, the cropping intensity of paddy was also influenced by the characteristics of planted land to paddy field which defines the level of land suitability for planting paddy. This research aimed to map paddy rotation using single date ALI imagery by assessing the land and soil characteristics based on the land suitability parameters for planting paddy. Soil characteristics such as texture, acidity level, P205 (phosphor) and C-organic level collected from field work and terrain characteristics such as landform, surface water, and drainage density from visual delineation of SRTM 90 m were collected as inputs for the decision tree analysis to map the repetition of paddy planting throughout the year. The results showed the overall accuracy of $85\% \pm 8\%$ (95 % level of confidence) for the final paddy rotation map where 2-times paddy per year was mostly found in the study area.

Copyright © 2017 GJGP-UNDIP

This open access article is distributed under a Creative Commons Attribution (CC-BY-NC-SA) 4.0 International license.

Arjasakusuma, S., Danoedoro, P., Herumurti, S., Nugroho, Y.A., Aryaguna, P.A., (2017). Land-Soil Characteristics For Mapping Paddy Cropping Intensity Using Decision Tree Analysis From Single Date Ali Imagery In Magelang, Central Java, Indonesia. *Geoplanning: Journal of Geomatics and Planning*, 4(2), 187-200. doi: 10.14710/geoplanning.4.2.187-200

1. INTRODUCTION

Rice is the product of the paddy (*Oryza sativa*) which is the world dominant staple food for human. Around 480 million tons of rice were annually produced to supply the demand of global citizen especially the poor fulfilling up to 50 % of their calories (Muthayya, Sugimoto, Montgomery, & Maberly, 2014). Same situation occurred in Indonesia, rice is one of the main foods for most of the poor citizen of the country where 20-25 % of total expenditure was used for rice consumption (Timmer, 2004). Considering the large rice consumption, population increase and the declining of paddy field area due to land use conversion in Indonesia, accurate and annual information about the area of paddy field and the intensity of paddy planting is important to measure the potential of annual rice production. These two main parameters can be used to measure the actual and potential deficit of rice production so that further policy can be carried to prevent further problem such as inflation. In addition, information regarding crop intensity can be used as baseline information for measuring the impact of climate variability to the agricultural area as the dependence of agricultural area to climate conditions (Xiao et al., 2006) as well as the effect of intensified agriculture pressure to soil quality and biogeochemical cycles (Yan et al., 2014).

Remote sensing has been considered as valuable technology to provide important information regarding precise crop management with the ability to monitor the seasonality of crop and soil conditions and time dependence crop management (Moran, Inoue, & Barnes, 1997). Seasonal and time dependence cropping pattern in paddy field area is one of the main information that can be monitored using remote sensing. Various remote sensing methods have been employed to map the cropping pattern information of

paddy field using single date or multi temporal remote sensing data (Foerster, et al., 2012; Li, et al., 2012; Pan et al., 2015; Panigrahy & Sharma, 1997; Wang et al., 2015; Xiao et al., 2006). Analysis of paddy field distribution and crop rotation was mainly carried out using multi temporal data due to differences in spectral value in the paddy field during the phase of flooding, paddy growth and harvesting (Panigrahy & Sharma, 1997; Wang et al., 2015).

Current analysis of remote sensing data using multi temporal analysis relies on the high temporal remote sensing data in order to be able to catch the seasonality pattern in spectral values at paddy field area. High temporal resolution remote sensing usually have coarse spatial resolution which consequently mixed pixel. This becomes problematic when it is implemented on monitoring the smallholder agricultural area generally having smaller area compared to the remote sensing pixel size (Jain, et al., 2013). Meanwhile, medium spatial resolution remote sensing data tends to have less intensive revisit time to capture the seasonality in the paddy field (Wang et al., 2015).

Cropping intensity at detailed scale is influenced by the socio-economic factor such as farmer's decision and crop price, and climate such as precipitation (Biradar & Xiao, 2011). In addition to climate and socio-economic factor, land physical constraints such as soil and water availability to supplement rainfall plays role in the farmer's decision for cropping intensity (Osman, et al., 2015; Panigrahy & Sharma, 1997). Effort to map cropping pattern information using landscape-ecological approach from remote sensing data has been demonstrated to be able to map cropping intensity with decent accuracy (Danoedoro, n.d.)

In this study, we aimed to predict the potential cropping intensity information (1) by combining ancillary data such as soil characteristics and land morphology;(2) by employing data mining algorithm using decision tree analysis; and (3) by further measuring how accurate the potential cropping intensity map matches with the reality in the world. The decision tree algorithm was used because it gives comparable accuracy with the other remote sensing classification algorithm (Friedl & Brodley, 1997; XU, et al., 2005). Furthermore, it is also able to handle data with different scales, no statistical assumption and fast, with the output of tree model that is easy to be interpreted (Tso & Mather, 2009).

2. DATA AND METHODS

2.1. Study area

This study took place in Magelang districts, Central Java, Indonesia located between 110° 01' 51" - 110° 26' 58" E and 7° 19' 13" - 7° 42' 16" S with total area of 1112 km² (Figure 1). Furthermore, this area has various land morphologies which can be seen from the hill shade background in the Figure 1. Therefore, considering the variety of land morphology, this area is suitable for assessing the land physical characteristics relationships to paddy cropping intensity.

2.2. Data

2.2.1. Earth Observation - ALI Imagery

Earth Observation – Advanced Land Imagery (EO ALI) is an experimental satellite that was launched by NASA in November 2000. The main objective of this satellite was to test the new system of hyperspectral and multispectral sensor for the next Landsat mission. The data has similar orbit with previous Landsat system with 30 m spatial resolution. However ALI has more bands with extra Near Infrared and Middle Infrared bands sensors with the total of 9 bands of multispectral (Bicknell, et al., 1999). Details of EO ALI sensors and technical configuration are shown in the following Table 1.

2.2.2. Field Data

Field data was conducted to collect the input data for decision tree analysis and validation data for accuracy assessment. Input data that was collected from the field data were the soil characteristics such acidity level (pH), phosphor level (P2O5), C-organic level, soil permeability and soil texture while the collection of validation data included the land use data and cropping pattern mapping data to assess the accuracy of crop intensity map.

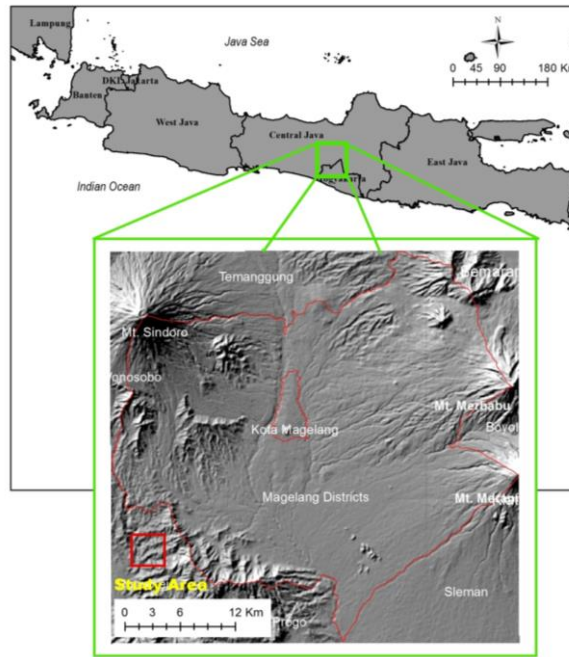


Figure 1. Map of study area, Magelang District, Central Java, Indonesia with SRTM 90 m hill shade as the background in the inset map (analysis, 2015)

Table 1. Technical configuration of EO-ALI (Lencioni, et al., 2005)

Characteristics	Specification	Details
Numbers of Bands (pixel size)	Panchromatic (0.48 – 0.69 μm)	1 band (10 m)
	Visible Near Infrared (0.433 – 0.89 μm)	6 bands (30 m)
	Shortwave Infrared (1.2 – 2.35 μm)	3 bands (30 m)
Orbit	Sun-synchronous with 705 km altitude aligned with Landsat 7 satellite	
Scanning system	Push-broom scanning	
Field of View (FOV)	15° cross-track by 1.26° in-track	
Swath width	37 km	
Temporal resolution	16 days	
Radiometric resolution	12 bits	

2.3. Methods

There are three main analyses in methods that cover 1.) identification of mapping unit boundary using visual inspection of land morphology from SRTM data, 2.) implementation of two-stage decision tree analysis on remotely sensed data where the first decision tree was used to extract paddy field area using based on at surface radiance-spectral values and the second decision tree was used to add cropping intensity information to the extracted paddy field area, and 3.) accuracy assessment of the paddy field area as well as the cropping intensity. Systematic general workflow of the aforementioned analyses is shown in the Figure 2.

2.3.1. Identification of Mapping Unit Boundary

In this study, land terrain unit acted as the mapping unit which is the smallest unit used in the analysis. The consideration of using terrain unit as the mapping unit was based on the assumption of homogenous terrain morphology having homogeneous lithological unit. Therefore, when the area was undertaken by the geomorphological processes such as erosion and sedimentation, it will create similar

output of land morphology depending on the resistance of rock material to the erosion process and the direction of the sedimentation. Terrain attributes were identified as important variables for soil properties mapping as used by (Elnaggar & Noller, 2009) which shows the importance of using of terrain unit as mapping unit.

The terrain was modeled by SRTM hill shade in which delineation was visually made by looking at the similar pattern of morphology such as slope, stream flow, hills and canyon. SRTM used in this research to generate hill shade has the spatial resolution of 90 m so that the mapping unit was identified at medium scale.

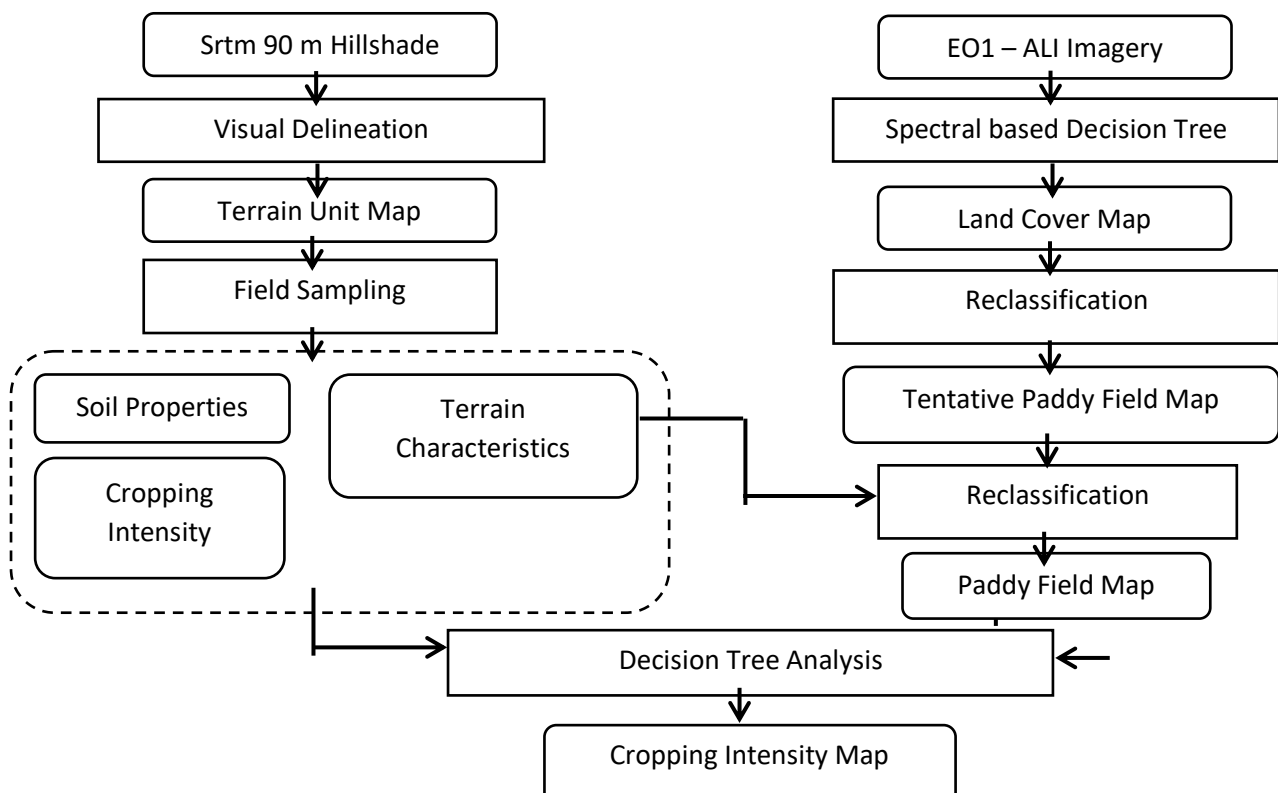


Figure 2. Workflow to extract cropping intensity from ALI – imagery using decision tree analysis (analysis, 2015)

2.3.2. Two Stages Decision Tree Analysis

Decision tree is a method to categorize data using hierarchical splitting procedures. This method offers the ability to comprehensively understand the relationship between the data attributes for data classification (Tso & Mather, 2009). The final classification is represented in a form of a model tree composed of root, interior, and terminal nodes depicting the decision stages that were taken to categorize the data. The creation of the tree model can be manually done based on the user’s knowledge or automatically using a set of training area or known as supervised decision tree or tree induction. Many algorithms have been developed for creating automatic tree model such as ID3, C.45, CART and See5.0.

See5.0/C5.0 algorithm applied in the software of see5® were used in this study in the formation of tree model. This algorithm was a development of the previous C4.5 algorithm which uses information or normalized gain ratios to perform the splitting process. The information gain was calculated based on the Shannon’s entropy measures. See 5.0 offers boosting techniques which are able to increase the model accuracy by repeatedly performing a tree-generating process. This process assigned weight to the training area with the ratio value of misclassified training area to the correctly labeled area.

In this study, two stages of decision tree induction were performed to produce land cover data and paddy rotation map. The first tree induction to produce land cover associated with paddy field and non-

paddy field employed the input value taken from each bands spectral values of the training area. The tentative map of the paddy rice distribution was further reclassified using the terrain unit derived from visual delineation of SRTM 90m using Boolean logic from field observation. The second stage of tree induction was carried out to map the frequency of paddy rotation by employing the set of training area from (1) land characteristics such as drainage density which qualitatively estimate stream flow density, relief, surface water and morphology and (2) soil characteristics such as texture, acidity level, P2O5 and C-organic level taken from the field work. The land characteristics that described the terrain configuration were considered useful to identify the distribution of soil properties (Sumfleth & Duttman, 2008) and water availability (Samson, Ali, Rashid, Mazid, & Wade, 2004). In the other hand, soil characteristics resembles the crop management such as cropping intensity and tillage which affect the dynamic of soil chemical, nutrient and quality (Liebig, Tanaka, & Wienhold, 2004; Mikha et al., 2006). Therefore, mapping cropping intensity using the combination of land-soil characteristics and remote sensing observation is possible to be conducted.

2.3.3. Mapping paddy field and reclassification process

To extract paddy field category from the single date ALI Imagery, cautious selection of training area were taken to map the land cover information from the imagery. Sixty training areas were taken to resemble various land cover visible in the 30 m of spatial resolution ALI Imagery. Average spectral value from each training area was used as the input for decision tree analysis with the standard deviation less than three to ensure the training area representing homogeneous land cover class. Land cover class information that was derived from the imagery was used as the base information for next classification of paddy and non-paddy field classes. To successfully deriving this information using single date imagery data, the possibility of visual appearance of paddy field at different states of condition should be taken into account. Paddy field which is recently planted will have water body that dominates spectral pattern meanwhile paddy field which is dominated with paddy cover had vegetation spectral pattern. Lastly paddy field that is recently harvested had bare soil typical spectral pattern (Figure 3). To accommodate those three characteristics of paddy field in single date imagery, different training classes were assigned to specifically record and training the algorithm to recognize these spectral pattern.

According to the crop calendar from the Indonesian Ministry of Agriculture, in the time when ALI imagery that was recorded in June 2004 during this study, the paddy fields in the area were mostly at the growth stage. In line with that, the paddy fields with the vegetation spectral pattern were mostly found in the imagery even though several unplanted and wet paddy fields were correspondingly found at this time. To capture the variability of paddy field in the study area, from 60 training areas that has been randomly collected, 23 of the training area were affiliated to paddy field in which 12 of the training areas are affiliated to the planted paddy fields dominated with vegetation cover, 4 training areas represent wet or moist soil covered paddy fields and 7 training areas affiliated to bare soil covered paddy fields. Those classes related paddy field were easily distinguished from other existing land cover classes in the 30 m spatial resolution of the imagery due to the homogenous shape and association with the pathway and the linearity of possible irrigation network. Although other non-paddy vegetation cover such as woody vegetation and smallholder plantation equally exists in the imagery, this vegetation can also be easily distinguished by looking at the tone and density of vegetation for woody vegetation and location for the smallholder non-paddy plantation. However, misclassification especially for smallholder plantation and vegetation covered paddy field was still possible due to the use of spectral information solely as an input during the classification process. Thus, additional post-processing was performed to rectify this misclassification using matching method based on the identified terrain unit.

Tree model was produced using See5.0(c) software which average values from each training area were employed as the base information to do the splitting process in decision tree in this study. There are 75 times of trial with 25 % pruning to avoid excessive leaf nodes in the output tree model so that over-fitting in the output model can be avoided. From 75 trials, there are 75 output tree models with different number of leaf node, splitting decision and error. Error of each model was counted by assessing the number of misclassified input training areas that did not match with the output tree model.

The Land cover classification can be linked into 2 main land use categories which are paddy and non-paddy fields. By performing this simple reclassification, tentative spatial distribution of paddy field was able

to be extracted. Visual assessment matched with the ground checking revealed misclassification in the tentative map especially between covered paddy field and dry smallholder plantation. Both have similar spectral signature due to vegetation coverage although both were associated with different landform unit. Based on ground checking, smallholder plantation is generally located at higher elevation with steeper slope and thin soil solum. With minimum need of irrigation for the majority of the crops in the smallholder plantation, planting these crops in this landform is feasible. On the other hand, developing paddy field in this area would be problematic as water availability is a limiting factor. With distinct differences in terms of landform for smallholder plantation and paddy field, binary matching using land terrain unit from the previous analysis was used to remove the falsely identified paddy field.

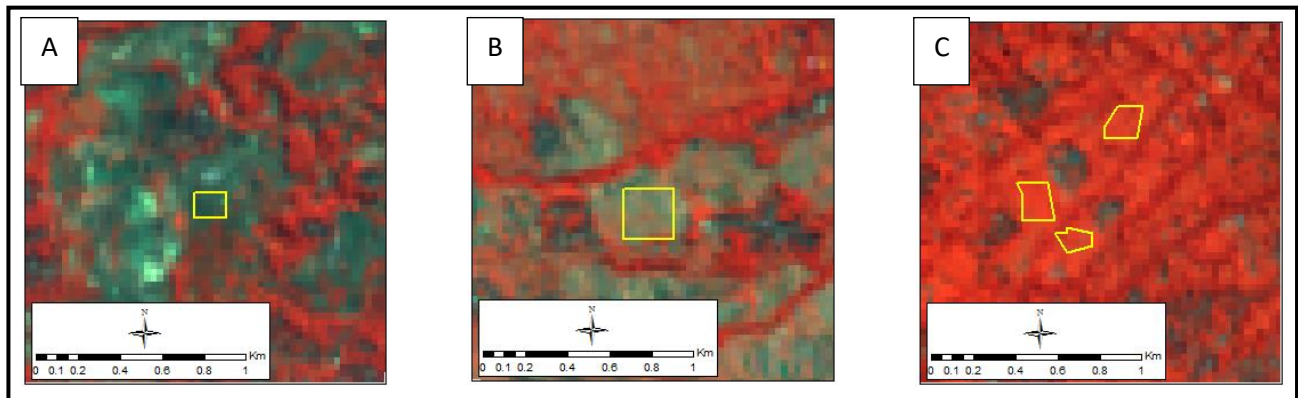


Figure 3. Different visual appearance of paddy field at different states of condition as appeared in ALI multispectral imagery false color composite such as (a). paddy field at flooding stage, (b). paddy field with bare soil reflectance, and (c). paddy field with vegetation (analysis, 2015)

2.3.4. Accuracy Assessment

Accuracy assessment was carried using error matrix based calculation developed by (Olofsson, Foody, Stehman, & Woodcock, 2013) in order to derive the producer's, user's and overall accuracy in percentage. The confidence interval for the producer's, user's and overall accuracy were calculated using this method which gives deviation of error range for the accuracy. As accuracy assessment of the map was calculated from statistical analysis using sample data representing the entire population in the map, the confidence interval calculation was necessary. Therefore, the accuracy value should not be merely represented by a single number but the margin of error should be included in the accuracy assessment. This was to give the value of uncertainty of the final accuracy number (Olofsson et al., 2013). The research implemented conventional error matrix as the baseline information of the accuracy calculation. However instead of using normal pixel or point count calculation in the conventional error matrix (Table 2), another error matrix was developed using the estimated area proportion for class i, j (P_{ij}) in the cell entries (Table 3).

Table 2. Conventional error matrix using sample counts (n) with additional information of Area and Area Proportion (W_i) to transform sample counts into estimated area proportion

Class	1	2	..	Q	Total	Area	W_i
1	n_{11}	n_{21}	..	n_{q1}	$n_{1.}$	A_1	$A_1/A_{.Tot}$
2	n_{12}	n_{22}	..	n_{q2}	$n_{2.}$	A_2	$A_2/A_{.Tot}$
..
q	n_{1q}	n_{2q}	..	N_{qq}	$n_{q.}$	A_q	$A_q/A_{.Tot}$
Total	$n_{.1}$	$n_{.2}$..	$n_{.q}$	N	$A_{.tot}$	1

Table 3. Modified error matrix using estimated area proportion (p) in the cell entries (Olofsson et al. 2013)

Class	1	2	..	Q	Total
1	p11	p21	..	pq1	p1.
2	p12	p22	..	pq2	p2.
..
Q	p1q	n2q	..	Pqq	pq.
Total	p.1	p.2	..	p.q	P

Here, rice field data was collected using ground checking combined with Google Earth© observation while cropping intensity validation was collected by performing interview to the local farmers. There are 6171 pixels for validation area from the ground checking and extrapolation process for validating rice field area and 42 points of validation data for estimating the accuracy of cropping intensity map.

3. RESULTS AND DISCUSSION

3.1. Identification of Mapping Unit Boundary

There are approximately 26 main land units that were identifiable by visual interpretation from SRTM 90 m hillshade data (Figure 4). These land units represented various morphology, surface roughness and slope steepness indicating different process and intensity of geomorphological processes occurring in the background. The study area was dominated by main geomorphological processes such as volcanic (V), fluvial (F) and denudation (D). Paddy field seems to be centered on fluvial landform with least steep slopes and good water availability in the surface. While in the upper volcanic neck and denudation hill with rough surface roughness and intensive stream flow indicating the undergoing process of gully erosion and less capability of soil to hold water, dry smallholder plantation are often discovered. The association of land physical unit characteristics with paddy field and dry plantation is useful for performing reclassification process to further eliminate possible misclassification in the land use mapping.

3.2. Mapping and Reclassification of Paddy Field

First stage of decision tree analysis resulted in 75 tree models with different nodes and error. From these trees the best model having least nodes and lower error was chosen (Figure 5). This particular model had 12 nodes and 8.25 % of error indicating that from the total of 60 training areas, 5 of them were misclassified or did not match with the output tree model. The output of tree model depends highly on the input training area so that changes in training area will correspondingly change the output of the tree model. Therefore, the binary splitting decision in the tree model is area-specific and cannot be applied to other imagery and other areas.

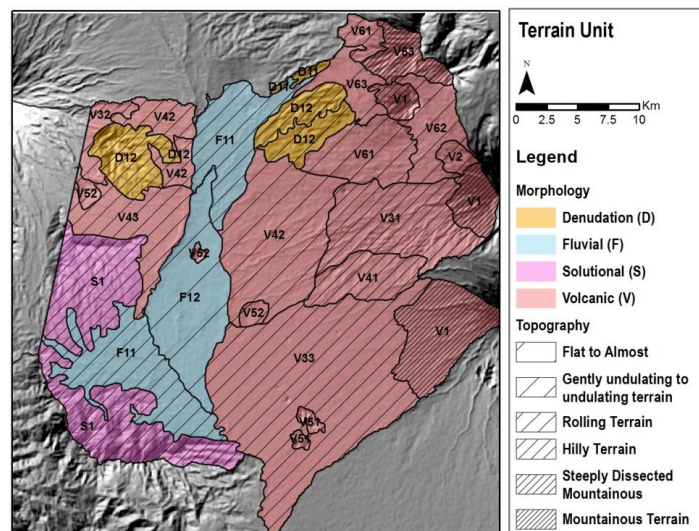


Figure 4. Map of terrain unit identified from SRTM 90 m hillshade data and their corresponding land morphology and topography attributes (analysis, 2015)

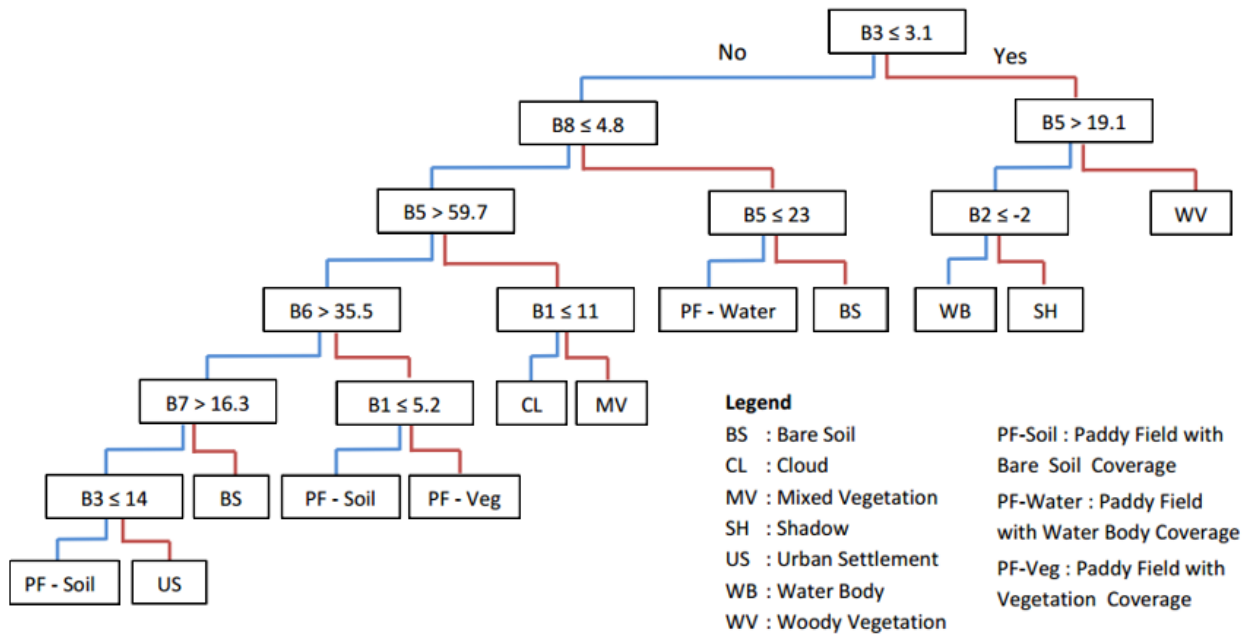


Figure 5. Output tree model produced by using average spectral value of different land cover in the training area (analysis, 2015)

The association of paddy field to certain land terrain characteristics (Figure 3) was developed based on the ground checking in order to map the paddy field. In the study area, we observed that paddy fields were mostly distributed in flat to least steep with less surface stream flow. This indicated low gully erosion in this area indicating thick soil solum and the ability to sustain water. These characteristics are mostly associated with fluvial process even though some paddy fields were also found in volcanic and denudation landforms such as low to middle volcanic plain, and denudation landform. Using these observations, the association was transformed into binary rules (Table 4) to rule out the false identified paddy field from the tentative map.

Table 4. Binary matching to rule out false identified paddy field from the tentative map (analysis, 2015)

Class	Mapping Unit									
	D11	D12	F11	F12	S1	V1	V2	V31	V32	
Non-Paddy Field	+	+	+	+	+	+	+	+	+	
Paddy Field	-	+	+	+	-	-	-	-	+	

Class	Mapping Unit									
	V33	V41	V42	V51	V52	V53	V61	V62	V63	
Non Paddy Field	+	+	+	+	+	+	+	+	+	
Paddy Field	+	-	+	-	+	-	-	+	+	

Legends : + = exist, - = not exist

The resulted map showed the removal of pre-identified paddy field especially in the area with higher latitude and coarse surface roughness which agricultural related land cover were associated to smallholder non-paddy plantation (Figure 6). This process reduced the detected paddy field although it still has to be justified by conducting the accuracy assessment. It is worth noted that the use of the binary matching resulted in a discrete distribution of where paddy fields are located. While in the real world, the boundary might be fuzzy even though it is still controlled by the landform characteristics.

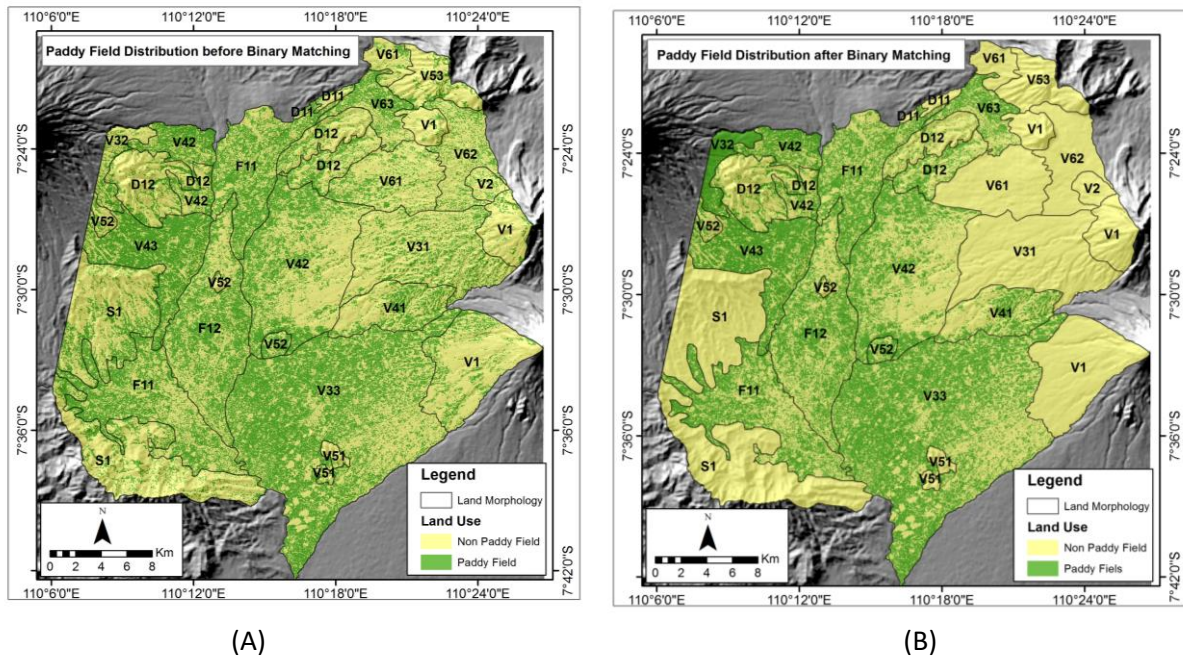


Figure 6. Paddy field distribution derived by using (A) decision tree model employing the spectral values and further corrected by using (B) binary matching by considering the land morphology – land use relationships

3.3. Mapping Paddy Cropping Intensity

The remaining paddy fields after reclassification process were labeled with the paddy cropping intensity. This analysis was performed using similar decision tree analysis with land and soil characteristics as the training data to select the significant variables influencing the paddy cropping intensity. Eighteen soils samples were taken from different landform where paddy field was found. These samples carrying the information of land physical and soil characteristics were used as an input for the decision tree analysis. Seventy five trials of decision tree analysis were executed to produce 75 decision tree models. To assess one best output model, error and the number of leaf nodes were used as the criteria. Here, tree model with 5.6 % error and 6 leaf nodes were selected where the phosphate level (P2O5), soil texture and pH were significant variables that influence paddy field cropping intensity (Figure 7).

The selected soil parameters of phosphate level, acidity level and soil textures in the tree model indicated the significance of those parameters for supporting cropping intensity although the relationships between cropping intensity and the parameters are most likely indirect. The parameters detected in the tree model are often used for indicating the chemical and physical properties needed to study the soil nutrients in the paddy fields among other parameters (Cho & Han, 2002). In particular, P2O5 is often used in the chemical fertilizer in which its amount of substance in the soil is needed to sustain and/or to improve paddy productivity (Inthavong, Fukai, & Tsubo, 2011; Mishima, Taniguchi, & Komada, 2006). In addition, the productivity of paddy fields is also influenced by the water holding capacity factor. One of the parameters that control the water availability is soil texture besides the topographic configuration and rainfall intensity (Inthavong et al., 2011). Furthermore, soil texture with less sand will have higher productivity due to its capability to preserve waters, nutrients and high electric conductivity (Mzuku et al., 2005).

In the above tree model, sandy loam was the threshold category of soil texture which is still able to be planted maximum 2 times per year. The other condition of soil texture with P2O5 below 97 could only yield rice for 1 time per year according to the model in the left side branch. In the right side of the branch, in order to obtain maximum 3 times per year planting intensity, pH level close to neutral (equal or above 5) and high soil nutrient of phosphate (equal or above 97) is needed. Otherwise, land is only able to be planted once per year due to the acidic condition of the soil. Output tree model showed that soil characteristics have more influence compare to land physical characteristics that were also included as the training area. The conclusion of whether soil characteristics has more influence to paddy cropping intensity compared to land physical characteristics cannot be uniquely justified based on this analysis. This is due to the nature of

decision tree analysis that is sensitive to the input training area. Thus, the need for more samples to become more stable is indispensable. However, these selected variables in the tree results seems to align with the statement of soil characteristics that influenced more on the paddy field productivity compared to the location or the spatial distribution of paddy field (Keersebilck & Soeprapto, 1985). However, similar to the previous decision tree model, this output model is highly sensitive to the input of training area. It cannot be implemented for another study using different input even though the same parameters were used.

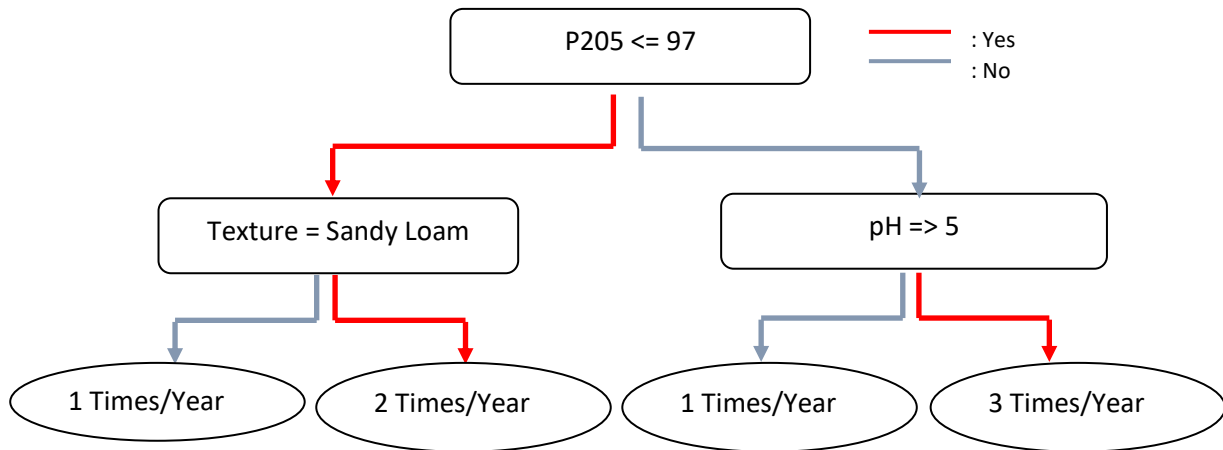


Figure 7. Output tree model to map the cropping intensity using land physical and soil characteristics as the input training data (analysis, 2015)

Cropping intensity using these rule sets from decision tree analysis showed that 2 times a year paddy cropping is common in this area which occupies 47 % (13,520 ha) of the total paddy field (Figure 8). The boundary of different crop intensity category in the map is distinct due to the use of land terrain unit as the smallest mapping unit for mapping cropping intensity. Two to three times a year, paddy seems to be clustered around the alluvial plain and at the lower level of Merapi volcanic slopes.

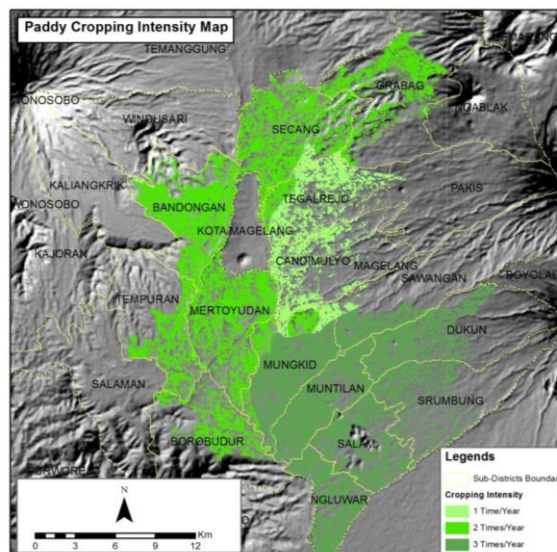


Figure 8. Final cropping intensity map where dominant of 2 times a year paddy field was found in the area, the distribution of different types of cropping intensity appear distinct due to the use of terrain unit as the mapping unit (analysis, 2015)

3.4. Accuracy Assessment

Accuracy assessment was performed to both tentative and reclassified maps of paddy field distribution using 6171 pixel count based on the field observation and extrapolated using Google Earth imagery. Based on the calculation performed, the tentative map produced the overall accuracy of 80 % ± 1% with the 79 % ± 1.5 % and 76 % ± 1.4 % for user accuracy and producer accuracy, respectively (Table 5). After using landform association with the paddy field existence to remove the possible false identified paddy field from the tentative map, 9 % increase in the overall accuracy has been achieved (Table 6). The overall accuracy for the reclassified map is now 89 % ± 1 % with 86 % ± 1.2 % and 82 % ± 1.6 % for the user accuracy and producer accuracy. This accuracy is at par with the accuracy of rice field mapping of different studies that were able to map rice field at the accuracy around 78 - 90 % (Gumma, Thenkabail, Maunahan, Islam, & Nelson, 2014; Wang et al., 2015). The increasing of accuracy after the post processing is the proof that using the relation between landform and land use was able to remove possible paddy field mistake in the map.

Table 5. Accuracy before post processing using landform approach (analysis, 2015)

		Reference			Area (Ha)	Wi	Class Accuracy		CI	
		Paddy Field	Non Paddy Field	Total Sample			Ui	Pi	Ui	Pi
Map	Paddy Field	2353	620	2972	43287.97	0.431	0.79	0.76	0.015	0.014
	Non Rice Field	599	2600	3199	57033.25	0.569	0.81	0.84	0.014	0.010
	Total Sample	2951	3219	6171	100321.21	Overall Accuracy : 0.80 ± 0.01				

Table 6. Accuracy after reclassification using landform approach (analysis, 2015)

		Reference			Area (Ha)	Wi	Class Accuracy		CI	
		Paddy Field	Non Paddy Field	Total Sample			Ui	Pi	Ui	Pi
Map	Paddy Field	2646	425	3071	33790.78	0.337	0.86	0.82	0.012	0.016
	Non Rice Field	306	2794	3100	66530.76	0.663	0.90	0.93	0.010	0.006
	Total Sample	2951	3219	6171	100321.55	Overall Accuracy : 0.89 ± 0.01				

Table 7. Accuracy of Cropping Intensity of Paddy Field (analysis, 2015)

		Reference				Area (Ha)	Wi	Class Accuracy		CI	
		1 Times/Year	2 Times/Year	3 Times/Year	Total Sample			Ui	Pi	Ui	Pi
Map	1 Times/Year	4	0	0	4	3699.05	0.129	1.00	0.66	0.00	0.21
	2 Times/Year	4	20	5	29	13520.70	0.470	0.69	1.00	0.17	0.00
	3 Times/Year	0	0	9	9	11555.41	0.402	1.00	0.83	0.00	0.11
	Total Sample	8	20	14	42	28775.16	Overall Accuracy : 0.85 ± 0.08				

Second accuracy assessment was performed to estimate the accuracy of cropping intensity map using decision tree analysis from 42 samples collected by interviewing local farmers on the area. Here, overall accuracy of the final cropping map showed 85 % ± 8 % which is good for the classification using medium resolution imagery (Table 7). However, in the individual class accuracies, producer accuracy for cropping intensity of 1 times of paddy per year and user accuracy for class 2 times a year is rather moderate to low with big deviation for the confidence interval. Producer accuracy of 1 times paddy per year was 66 % ± 21% and user accuracy for class 2 times a year was 69 % ± 17 %. This moderate to low accuracy was compared to related research using multirate imagery that able to map 90-97 % of producer accuracy of crop rotation (Nguyen, 2013; Panigrahy & Sharma, 1997). The low accuracy for individual class might be caused by the coarse unit of mapping that was used for generalizing the cropping intensity for paddy field within the specific landscape.

The drawback of using 90 m SRTM as the base data for identifying the boundary of land terrain is that only major terrain can be identified resulting in general landform mapping unit. Cropping intensity in paddy field was mostly influenced by the availability of water so that if the paddy field has a good irrigation system, then 2 or 3 times paddy per year can be achieved. However, variation of cropping intensity within clustered paddy field in the same area also possible. Topography controls the distribution of the surface water flows so that the distribution of water in the same paddy field may be different. In slope or terraced paddy fields will consistently receive water compared to paddy fields in the plain and higher elevation which will sustain less water especially in the dry season. In term of irrigation management, terraced paddy field will arranged to be keep inundated especially in the dry season to prevent the landslide which gives the farmer the chance to plant another crop (Adachi, 2007). In addition, further assessment to include to climate data such as rainfall and temperature in the analysis to the dynamic of cropping intensity can be performed with the dependence of paddy field to water availability.

4. CONCLUSION

The accuracy of predicted cropping intensity map revealed the possibility of using the landform approach and characteristics. It was used to improve the identification of paddy field and its cropping intensity attributes with respectable accuracy. Decision tree was used as the algorithm to develop the rule set from the collected training areas. Binary matching using landform approach was able to increase the overall accuracy of 9 % indicating the benefit of using logical relationships between landform and land use/land cover to omit the misclassification from the result. Rule set developed from decision tree analysis using soil parameters was also able to map cropping intensity with respectable level. However, future assessment using additional parameters such as monthly rainfall and temperature data is required to improve the result as well as socio-economic parameters such as the availability of labor to work on the land. In addition, the infrastructures of the paddy field such as irrigation systems require more detailed data in the future to show the irrigation systems. Detailed land morphology delineation could also be made by using higher spatial resolution of terrain data such as SRTM 30 m or Aster GDEM to enhance the detail of land unit map and identify the micro-topography that affect the water distribution especially in rain-fed paddy field. The main drawback of this study lies particularly in the algorithm of decision tree, although decision tree can give respectable accuracy, the algorithm was strongly sensitive to the input of training area where small changes in the input can entirely change the output model. This also suggests that collection of training area should be taken carefully and the more number of samples should be taken to give more valid output of the tree model. Further assessment using ensemble method, random forest or neural network analysis also could be explored in this study application.

5. ACKNOWLEDGMENTS

The authors would like to acknowledge Graduate School of Remote Sensing, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia for the support in this study. The authors also would like to extend the gratitude to Statistical Agency (BPS) of Magelang Districts, Central Java, Indonesia and BAPPEDA Magelang Districts for the permission to conduct field research and the support of the statistical data. The author is grateful for being granted with the *Beasiswa Unggulan DIKTI* for providing support during the study. We also thank anonymous reviewers for the comments and suggestions which significantly improve the quality of the manuscript.

6. REFERENCES

- Adachi, S. (2007). Agricultural technologies of terraced rice cultivation in the Ailao Mountains, Yunnan, China.
- Bicknell, W. E., Digenis, C. J., Forman, S. E., & Lencioni, D. E. (1999). EO-1 Advanced Land Imager. In W. L. Barnes (Ed.), *Earth Observing Systems {IV}*. SPIE. [[Crossref](#)]
- Biradar, C. M., & Xiao, X. (2011). Quantifying The Area and Spatial Distribution of Double- And Triple-

- Cropping Croplands In India With Multi-Temporal MODIS Imagery In 2005. *International Journal of Remote Sensing*, 32(2), 367–386. [[Crossref](#)]
- Cho, J.-Y., & Han, K.-W. (2002). Nutrient Losses From a Paddy Field Plot in Central Korea. *Water, Air, and Soil Pollution*, 134(1/4), 215–228. [[Crossref](#)]
- Danoedoro, P. (n.d.). Spatial, Temporal and Ecological Consideration in Satellite Image-Based Mapping of Crop Rotational Aspect: Cases From Several Parts Of Central Java.
- Elnaggar, A., & Noller, J. (2009). Application of Remote-Sensing Data and Decision-Tree Analysis to Mapping Salt-Affected Soils over Large Areas. *Remote Sensing*, 2(1), 151–165. [[Crossref](#)]
- Foerster, S., Kaden, K., Foerster, M., & Itzerott, S. (2012). Crop Type Mapping Using Spectral-temporal Profiles and Phenological Information. *Computers and Electronics in Agriculture*, 89, 30–40. [[Crossref](#)]
- Friedl, M. A., & Brodley, C. E. (1997). Decision Tree Classification of Land Cover from Remotely Sensed Data. *Remote Sensing of Environment*, 61(3), 399–409. [[Crossref](#)]
- Gumma, M. K., Thenkabail, P. S., Maunahan, A., Islam, S., & Nelson, A. (2014). Mapping Seasonal Rice Cropland Extent and Area in the High Cropping Intensity Environment of Bangladesh Using MODIS 500m Data for the Year 2010. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 91, 98–113. [[Crossref](#)]
- Inthavong, T., Fukai, S., & Tsubo, M. (2011). Spatial Variations in Water Availability, Soil Fertility and Grain Yield in Rainfed Lowland Rice: A Case Study from Savannakhet Province, Lao PDR. *Plant Production Science*, 14(2), 184–195. [[Crossref](#)]
- Jain, M., Mondal, P., DeFries, R. S., Small, C., & Galford, G. L. (2013). Mapping Cropping Intensity of Smallholder Farms: A Comparison of Methods using Multiple Sensors. *Remote Sensing of Environment*, 134, 210–223. [[Crossref](#)]
- Keersebilck, N. C., & Soeprapto, S. (1985). Physical Measurement in Lowland Soils Techniques and Standardization. *Soil Physics and Rice*. *International Rice Research Institute, Los Baños, Laguna, Philippines*, 99–112.
- Lencioni, D. E., Hearn, D. R., Digenis, C. J., Mendenhall, J. a., & Bicknell, W. E. (2005). The EO-1 Advanced Land Imager : An Overview. *Lincoln Laboratory Journal*, 15(2), 165-180.
- Li, P., Feng, Z., Jiang, L., Liu, Y., & Xiao, X. (2012). Changes in Rice Cropping Systems in the Poyang Lake Region, China during 2004-2010. *Journal of Geographical Sciences*, 22(4), 653–668. [[Crossref](#)]
- Liebig, M. A., Tanaka, D. L., & Wienhold, B. J. (2004). Tillage and Cropping Effects on Soil Quality Indicators in the Northern Great Plains. *Soil and Tillage Research*, 78(2), 131–141. [[Crossref](#)]
- Mikha, M. M., Vigil, M. F., Liebig, M. A., Bowman, R. A., McConkey, B., Deibert, E. J., & Pikul, J. L. (2006). Cropping System Influences on Soil Chemical Properties and Soil Quality in the Great Plains. *Renewable Agriculture and Food Systems*, 21(01), 26–35. [[Crossref](#)]
- Mishima, S., Taniguchi, S., & Komada, M. (2006). Recent Trends in Nitrogen and Phosphate Use and Balance on Japanese Farmland. *Soil Science and Plant Nutrition*, 52(4), 556–563. [[Crossref](#)]
- Moran, M. S., Inoue, Y., & Barnes, E. M. (1997). Opportunities and Limitations for Image-based Remote Sensing in Precision Crop Management. *Remote Sensing of Environment*, 61(3), 319–346. [[Crossref](#)]
- Muthayya, S., Sugimoto, J. D., Montgomery, S., & Maberly, G. F. (2014). An Overview of Global Rice Production, Supply, Trade, and Consumption. *Annals of the New York Academy of Sciences*, 1324(1), 7–14. [[Crossref](#)]
- Mzuku, M., Khosla, R., Reich, R., Inman, D., Smith, F., & MacDonald, L. (2005). Spatial Variability of Measured Soil Properties Across Site-Specific Management Zones. *Soil Science Society of America Journal*, 69(5), 1572. [[Crossref](#)]
- Nguyen, T. T. H. (2013). Earth Observation for Rice Crop Monitoring and Yield Estimation: Application of Satellite Data and Physically Based Models to the Mekong Delta.
- Olofsson, P., Foody, G. M., Stehman, S. V, & Woodcock, C. E. (2013). Making Better Use of Accuracy Data in Land Change Studies: Estimating Accuracy and Area and Quantifying Uncertainty Using Stratified Estimation. *Remote Sensing of Environment*, 129, 122–131. [[Crossref](#)]
- Osman, J., Inglada, J., & Dejoux, J.-F. (2015). Assessment of a Markov Logic Model of Crop Rotations for Early Crop Mapping. *Computers and Electronics in Agriculture*, 113, 234–243. [[Crossref](#)]
- Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., ... Liu, J. (2015). Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and*

- Geoinformation*, 34, 188–197. [[Crossref](#)]
- Panigrahy, S., & Sharma, S. A. (1997). Mapping of Crop Rotation Using Multidate Indian Remote Sensing Satellite Digital Data. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 52(2), 85–91. [[Crossref](#)]
- Samson, B., Ali, A., Rashid, M., Mazid, M., & Wade, L. (2004). Topographic Position Influences Water Availability in Rainfed Lowland Rice at Rajshahi, Northwest Bangladesh. *Plant Production Science*, 7(1), 101–103. [[Crossref](#)]
- Sumfleth, K., & Duttman, R. (2008). Prediction of Soil Property Distribution in Paddy Soil Landscapes using Terrain Data and Satellite Information as Indicators. *Ecological Indicators*, 8(5), 485–501. [[Crossref](#)]
- Tso, B., & Mather, P. M. (2009). *Decision Trees (2nd Editio ed.)*. Boca Raton: Taylor and Francis Group.
- Timmer, P. (2004). Food Security in Indonesia: Current Challenges and the Long-Run Outlook. *{SSRN} Electronic Journal*. [[Crossref](#)]
- Wang, J., Xiao, X., Qin, Y., Dong, J., Zhang, G., Kou, W., ... Zhang, Y. (2015). Mapping Paddy Rice Planting Area in Wheat-rice Double-cropped Areas through Integration of Landsat-8 OLI, MODIS and PALSAR images. *Scientific Reports*, 5(1). [[Crossref](#)]
- Xiao, X., Boles, S., Froking, S., Li, C., Babu, J. Y., Salas, W., & Moore, B. (2006). Mapping Paddy Rice Agriculture in South and Southeast Asia using Multi-temporal MODIS Images. *Remote Sensing of Environment*, 100(1), 95–113. [[Crossref](#)]
- XU, M., WATANACHATURAPORN, P., VARSHNEY, P., & ARORA, M. (2005). Decision Tree Regression for Soft Classification of Remote Sensing Data. *Remote Sensing of Environment*, 97(3), 322–336. [[Crossref](#)]
- Yan, H., Xiao, X., Huang, H., Liu, J., Chen, J., & Bai, X. (2014). Multiple Cropping Intensity in China Derived from Agro-meteorological Observations and MODIS Data. *Chinese Geographical Science*, 24(2), 205–219.