

MONITORING AND ASSESSMENT OF AGRI-URBAN LAND CONVERSION USING MULTI-SENSOR REMOTE SENSING AND GIS TECHNIQUES

D. C. Fargas Jr. ^{1*}, G.A. M. Narciso ¹, A. C. Blanco ¹

¹ Department of Geodetic Engineering, College of Engineering, University of the Philippines Diliman, Quezon City 1101, Philippines – dcfargas@up.edu.ph, gmnarciso@up.edu.ph, acblanco@up.edu.ph

Commission III, WG III/7

KEY WORDS: Data fusion, LULC, Change Detection, Random Forest, Google Earth Engine

ABSTRACT:

Continuous agricultural land conversion poses threat to food security but this has not been monitored due to ineffectual policies. One of the Philippine provinces with a high rate of conversion is the rice-producing province of Cavite. To assess the spatiotemporal dynamics of agricultural land conversion in Cavite, this study aims to develop an operational methodology to produce Land Use and Land Cover (LULC) change maps using a multi-sensor remote sensing approach for decision making and planning. LULC maps were generated using Random Forest Classification of Landsat 8 and Sentinel-1 image collections. Spectral indices, combinations of radar polarizations (VV, VH), and their principal components were included to improve its accuracy. Conversion maps were generated by taking the bi-annual difference of LULC maps from 2016 to 2019. Accuracy was assessed using visual inspection with Google Earth Pro. Classification was carried out using single-sensor (optical or radar) and multi-sensor (optical and radar) approach in combination with three feature selection algorithms, namely, Sandri and Zuccolotto (2006), Liaw and Wiener (2015), Kursu and Rudnicki (2010). Multi-sensor and single sensor yielded similarly high overall accuracies (OA = 96%) with the exception of single-sensor radar approach (OA = 53%). Multi-sensor approaches exhibit high accuracies (Cumulative Accuracy = 91%) in detecting agricultural to built-up LULC change up to 5,000 square meters unlike single-sensor optical approach (Cumulative Accuracy = 76%). Among the multi-sensor approaches, the method of Liaw and Wiener (2015) remains to be superior as it only uses eight (8) variables.

1. INTRODUCTION

As an agricultural country, the Philippines had been reliant on rice, its most important agricultural commodity, for its economic growth and food security as it constitutes for about 15% of the gross value added in the agriculture industry while serving as the main source of food (David & Balisacan, 1995; Abdullah, Ito, & Adhana, 2006). Its economic contribution is exhibited by the overall land use of the country as nearly two-thirds of the country's farm lands produce rice (David & Balisacan, 1995). Moreover, the demand for rice has been increasing based from the average growth rate of the consumption per capita at 0.73% from 1996 to 2003 (David & Balisacan, 1995; Abdullah, Ito, & Adhana, 2006). To help support and sustain this valuable industry in the country, especially in attaining food-security, implies the need for effective monitoring systems and tools, and one way to implement that is through Remote Sensing and GIS technologies (Ravanera, 2018; David & Balisacan, 1995).

Various studies have provided substantial findings on the use of remote sensing in agricultural monitoring applications such as that of Torbick et al. (2017) which utilized optical imagery to monitor rice crop lands. A research of the Philippine Rice Research Institute in 2019, on the other hand, developed a model utilizing radar-based imagery to detect and monitor rice crops. However, using optical or radar imagery alone has been found to have limitations. Therefore, for agricultural RS applications, it was proposed to integrate both imageries as they could produce better results (Torbick et al, 2017; Boschetti et al, 2015; Joshi et al, 2016, Dusseux et al, 2014).

This study therefore aims to support the agriculture industry in the country by developing an operational methodology to produce Land Use and Land Cover (LULC) change maps, specifically focusing on agricultural land conversion, using optical and radar remote sensing datasets. This paper presents the following: (1) generation of LULC change maps of the study area using multi-sensor and single-sensor image analysis; (2) comparison of the approaches and results based on the accuracies of the generated LULC change maps; and (3) assessment of the spatio-temporal dynamics of agricultural land conversion in the study area.

Since free access to radar images was realized only recently, the timeframe of this study has been limited to 2016 and 2019 only. Moreover, the study focused on agricultural to built-up LULC changes only. Agricultural land, in this study, is defined to include cultivated and uncultivated croplands, grasslands and fallow lands. These sub-categories of agricultural lands are not distinguished in this study.

2. RESEARCH BACKGROUND

2.1 Integration of Optical and SAR Satellite Imagery

While both types of datasets have been utilized in LULC mapping applications, each has its own advantages and limitations. Optical remote sensing uses visible, near infrared (NIR), shortwave infrared (SWIR) sensors, and their derivative indices to assess and monitor vegetation growth and land cover dynamics (Lefsky and Cohen, 2003; Shaw and Burke, 2003;

* Corresponding author

Dusseux et al., 2014; Fieuzal, 2011). However, for temporal analysis, compositing/mosaicking is needed as optical imagery is more prone to cloud cover. This leads to sparse time series and ambiguous pixels due to the temporal differences and atmospheric variations of the optical images being mosaicked (Dusseux et al., 2014; Erasmí and Twele, 2009; Joshi et al., 2016; Torbick et al., 2017).

Radar remote sensing can capture surface information and produce dense time series regardless of weather and the Sun's presence (Idol, Haack and Mahabir, 2015; Erasmí and Twele, 2009). With varying dielectric constant to the surface, radar remote sensing can monitor soil conditions (Lefsky and Cohen, 2003; Fieuzal et al., 2011). However, it is limited to a single microwave. The inherent presence of speckle in the images results to uncertainties and poor accuracies during analysis (Idol, Haack and Mahabir, 2015; Joshi et al., 2016).

The combination of optical and radar-based datasets for land use and land cover mapping is really suggested as the synergism between both datasets can be utilized for improved analysis and results (Joshi et al., 2016; Erasmí and Twele, 2009).

2.2 Dimensionality Reduction and Variable Selection

Combining optical and radar-based datasets directly usually results to high data dimensionality and redundant information. This combination compromises interpretability. Principal Components (PC) analysis and Random Forest (RF) variable selection help in simplifying datasets with high dimensionality (King and Jackson, 1999; Genuer, Poggi and Malot, 2010; Argamosa et al., 2018). PC analysis summarizes variances of variables into uncorrelated dimensions producing different components with newly extracted information which are weighed using eigenvalues (King and Jackson, 1999). On the other hand, RF designates variables with importance scores based on the increase in the mean of the error of a tree in a forest when the variable is employed during a regression or classification process (Genuer, Poggi and Malot, 2010). A high variable importance signifies a greater contribution to a model's accuracy while a low variable importance suggests that a variable has no significant contribution to the model regardless of the permutation applied (Argamosa et al., 2018).

Various studies utilize RF for variable selection. In 2006, Sandri and Zuccolotto proposed RF variable importance which is measured by calculating the centroid of the variable importance values then selecting variables that are within a given threshold/average distance. This method has been effective as the noise variables tend to cluster together while significant variables appear as outliers (Sandri and Zuccolotto, 2006). On the other hand, Argamosa et al. (2018) uses a different implementation of variable selection following Liaw and Wiener (2015). In their study, the variables were arranged by increasing variable importance scores and grouped into different regions based on the similarity of scores as defined by a relatively flat slope (Liaw and Wiener, 2015). An abrupt change in slope denotes a new region and for each region, the variable having the highest variable importance score was selected (Liaw and Wiener, 2015). A variable was not selected in the first region due to its low importance (Liaw and Wiener, 2015). Kursá and Rudnicki (2010) used "shadow variables" which are generated with random pixel values. Z-scores of the variables are computed and all variables with an importance lower than the maximum Z-score value among the shadow variables were omitted (Kursá and Rudnicki, 2010).

2.3 Random Forest Classification

For multi-sensor remote sensing applications, Random Forest (RF) classifiers have been proven to produce superior and consistent accuracies even for large sets of data with high dimensionality (Torbick et al, 2016; Pal, 2003; Baghdadi and Zribi, 2016; Pelletier et al, 2016). Furthermore, RF requires shorter training time and easy stable parameterization because its parameters cause only small influence on the classification accuracy (Pelletier et al, 2016). However, RF shows possible overfitting of data but is still advantageous due to its ability to produce higher accuracies and efficiently process large and diverse datasets (Pal, 2003).

2.4 Dataset Derivatives

Various derivatives can be utilized for improved classification resulting in better LULC maps. For optical imagery, Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Normalized Difference Built-up Index (NDBI) and the Normalized Difference Bareness Index (NDBal) are among the indices which help discriminate land features (Ghosh, Chatterjee and Dinda, 2018; Li and Chen, 2014; Sharma, Ghosh and Joshi, 2012). NDVI utilizes the Red and Near infrared regions to emphasize greenness of land features making vegetation more visible compared to other features (Torbick et al, 2017; Szabo, Gacsi and Balazs, 2016). For water features, NDWI takes advantage of the near and short-wave infrared regions (Szabo, Gacsi and Balazs, 2016). NDBI is another index used to extract built-up features and lastly, the NDBal index discriminates bare soil features however, alternative to this index is the Bare Soil Index (BI), developed to model Forest Canopy Density (FCD) and exhibited better discrimination of the built-up and bare soil features (Macarof and Statescu, 2017; Li and Chen, 2014; Sharma, Ghosh, and Joshi, 2012; Rikimaru, Roy and Miyatake, 2002).

Similarly, radar-based imagery also has its derivatives and among the commonly used are the different combinations of polarizations such as the difference, mean and ratio of vertical and horizontal polarizations. These have been found to significantly increase classification accuracy (Abdikan et al, 2016).

3. MATERIALS AND METHODS

3.1 Study Area

The area of study is the rice-producing province of Cavite, located in the CALABARZON region of the Philippines (Figure 3). CALABARZON, one of the main producers of rice, fears for food security as agricultural lands are continuously converted due to urban development or expansion (Cabildo, 2017; Philippine Statistics Authority CAF, 2012).



Figure 1. Study Area (Cavite)

3.2 Materials

3.2.1 Data Processing

Cloud Computing - Google Earth Engine (GEE): Pre-processing of satellite imagery up to generation of conversion maps were all performed in Google Earth Engine. Google Earth Engine (GEE) was used as a tool for cloud computing and processing most especially for researchers who lack high-performance computational resources (Torbick et al, 2017).

Google Earth Pro: Generated KML files were overlaid in Google Earth Pro to validate conversion. Said Regions-of-interest (ROIs) were checked individually to determine conversion mapping accuracy for each category.

3.2.2 Satellite Data

This study utilized Landsat 8 and Sentinel-1 satellites. With a 16-day repeat cycle Earth coverage and having a total of 11 bands ranging from visible to infrared wavelengths with an average 30m spatial resolution, Landsat is able to provide timely and high-quality optical multispectral data which can be utilized for global and regional change detection and surface characterization applications (United States Geological Survey, 2020). Sentinel-1 offers SAR data with a 6 day repeat cycle at the equator which can be used for improved temporal change detection and analysis (ESA, 2019).

The integration of the optical and radar-based datasets requires the same temporal coverage for the images. The Google Earth Engine Data Catalog provides an overlap of satellite imagery for both sensors starting mid-2015 pushing the timeframe of the study from 2016 to 2019.

Optical-based Data: USGS Landsat 8 Surface Reflectance Tier 1 imagery derived from the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) were utilized in this study. Images were grouped by year to generate annual cloud-free composite images. Two image collections were formed, one for the year 2016 and another for 2019. The 'pixel_qa' band derived from the CFMASK algorithm detects pixels with cloud and cloud shadow attributes masking clouds in the surface reflectance image. Annual cloud-free composites were then formulated by taking the median of the multiple intra-year pixel values as its value.

Optical-based Spectral Indices were used to improve classification accuracy. NDVI (Rouse et al, 1974) distinguished vegetation while NDWI (Gao, 1996), BU (He et al, 2010) and BI (Rikimaru et al, 2002) distinguished water, built-up, and bare soil areas, respectively. NDVI, NDWI, BU, and BI are calculated using the following equations (Eq. 1-4).

$$NDVI = (NIR - RED)/(NIR + RED), \quad (1)$$

$$NDWI = (NIR - SWIR)/(NIR + SWIR), \quad (2)$$

$$BU = \left[\frac{SWIR - NIR}{SWIR + NIR} \right] - \left[\frac{NIR - RED}{NIR + RED} \right], \quad (3)$$

$$BI = \frac{RED + BLUE + GREEN}{RED + BLUE - GREEN} \quad (4)$$

where RED = pixel values of the red band
BLUE = pixel values of the green band
GREEN = pixel values of the green band
NIR = pixel values of the near infrared band
SWIR = pixel values of the shortwave infrared band

Radar-based Data: Level 1 C-band Synthetic Aperture Radar Ground Range Detected (SAR GRD) of Sentinel was utilized. Images were grouped by month before radiometric and geometric calibration. The polarizations 'VV' and 'VH' were selected before taking the mean value of each group. This was used to compute for other derivatives of the radar-based dataset.

$$Mean = \frac{VV+VH}{2}, \quad (5)$$

$$Ratio = \frac{VV^2}{VH}, \quad (6)$$

$$Difference = VV - VH, \quad (7)$$

3.3 Methods

3.3.1 Land Use and Land Cover Mapping

Land cover classes need to be defined before generating LULC maps (Baghdadi et al., 2016; Jin et al., 2018; Torbick et al., 2017). For this study, the land cover classes to be used are (1) Agriculture: cultivated and uncultivated croplands, grasslands and fallow lands, (2) Built-up: settlements, lands used for artificial or industrial use such as roads, (3) Bare soil: bare fields and mining sites, (4) Pure and mixed forests, and (5) Water: inland waters, lakes, reservoirs, and flooded lands.

According to Jensen and Lulla (1987), the minimum number of training pixels should be at least 10 times the number of variables used in the classification model. Since there are five land cover classes, the minimum training population was 50 pixels. However, to improve training for the LULC classification model, 250 training pixels were selected and a total of 150 pixels were used for testing. Training pixels were marked through high resolution satellite imagery as seen from Google Earth Pro. Marking of these features were also based on the familiarity and personal knowledge of the researchers on the study area eliminating confusion in classification of some pixels.

3.3.2 Principal Component Analysis (PCA)

Before Random Forest Classification, the number of variables for a single composite was seventy (70): ten (10) from the optical dataset and sixty (60) from the radar dataset. This was computationally expensive and time consuming for Google Earth Engine; thus, dimensionality reduction was implemented.

PCA allowed the creation of a linear combination of variables while retaining the influence of variables that were most important to the classification (King & Jackson, 1999). Google Earth Engine (GEE), however, has a limit of twelve (12) input variables when performing PCA. This was performed first on the optical dataset as it only had ten (10) variables, although this did not produce positive results as each variable was unique to its counterparts. The radar dataset was split into five (5) groups: one for each polarization derivative consisting of the month values intra-year. It appeared as that the variables in each group showed more relationship as a single Principal Component (PC) eigenvalue already accounted for more than 95% of the sum of the eigenvalues of all the PCs.

3.3.3 Random Forest Classification

Number of Decision Trees: The optimal number of trees was determined by slowly increasing it and watching for out-of-the bag error (OOB) convergence (Jin et al., 2018). Its effect on the overall accuracy of the model was also observed.

Models based on Variable Importance: The original seventy (70) variables were reduced to just twenty-six variables (26) after the implementation of PCA.

A model was constructed based on the method of dictating which variables were deemed to be important, some of which led to the reduction of many variables while one, none. Five (5) models were created: (a) the method of Sandri and Zuccolotto (2006), (b) the method of Liaw and Wiener (2015), (c) the method of Kursa and Rudnicki (2010), (d) a method inspecting whether a single-sensor (optical) would suffice to accurately perform land cover classification; (e) a combination of the three (3) former methods by taking the mode of the generated composites.

3.3.4 Land Cover Change Detection

Conversion Map Generation: Conversion maps were produced by taking the bi-annual differences of LULC maps from 2016 to 2019. Pixels that did not change land cover or were not of agricultural land cover in the year 2016 were masked out or given a value of 0 representing no agricultural conversion present. Four (4) conversion maps were generated; this was done to compare conversion maps using the different models.

Spatial Filtering: Spatial filtering simplified raster maps before converting them into a vector format. A majority filter changes a raster value based on its contiguous neighbouring cells while boundary filter uses the expand and shrink method to clean ragged edges between zones (ESRI, 2014).

Categorizing Conversion: Final raster maps were converted into vector format. Shape areas of individual polygons were computed and categorized according to size: (a) 50,000 sqm and above; (b) 10,000 sqm to 50,000 sqm; (c) 5,000 sqm to 10,000 sqm; (d) 1,000 sqm to 5,000 sqm; (e) 500 sqm to 1,000 sqm. Each category was exported as a separate KML file and was imported into Google Earth Pro for accuracy assessment.

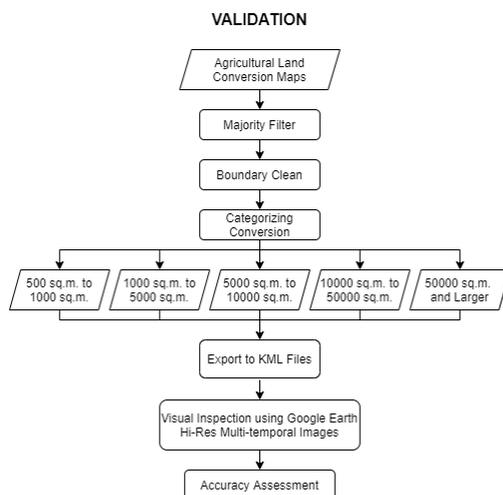


Figure 3. General Methodology – Validation

4. RESULTS AND DISCUSSION

4.1 Variable Reduction

4.1.1 Principal Component Analysis (PCA): The first principal component (PC1) of the VH, VV, Mean Backscatter, and Backscatter Ratio groups represented more than 90% of the total variation of all data. Thus, for these groups, the respective PC1s were used. This is not the case for the backscatter difference group, for which at least ten (10) PCs are needed to sufficiently describe variability. The twelve (12) backscatter difference variables were therefore retained instead of using the PCs to preserve all the information.

Table 1 shows the final pool of variables for the Random Forest Classification. The optical imagery dataset comprised of ten (10) variables: six (6) spectral bands and four (4) spectral indices. The radar-based imagery dataset had sixteen (16) variables: four (4) principal components and twelve (12) backscatter differences, one for each month of the year. The total number of variables in the pool was twenty-six (26).

Optical Imagery Variables	Radar-based Imagery Variables
Spectral Bands: B2, B3, B4, B5, B6, B7	PC1 - VV Backscatter PC1 - VH Backscatter PC1 - Mean Backscatter PC1 - Backscatter Ratio
Spectral Indices: NDVI, NDWI, BI and BU	12 Backscatter Difference

Table 1. Variables for reduced stacked image

Table 2 lists the variables used for each variable selection method. Liaw and Wiener’s method (2015) had the least variables at eight (8) while Kursa and Rudnicki’s method (2010) remained with twenty-six (26) variables. Variables labelled ‘diff_n’ refers to the backscatter difference for the nth month of the year.

Method	Count	Variables
Multi-sensor with Sandri & Zuccolotto (2006)	12	'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'BI', 'BU', 'NDVI', 'PC1_VH', 'PC1_VV', 'PC1_meanSAR'
Multi-sensor with Liaw & Wiener (2015)	8	'B2', 'B3', 'B5', 'BU', 'NDVI', 'PC1_meanSAR', 'diff_3', 'diff_11'
Multi-sensor with Kursa & Rudnicki (2010)	26	'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'BI', 'BU', 'NDVI', 'NDWI', 'PC1_VH', 'PC1_VV', 'PC1_meanSAR', 'PC1_ratioSAR', 'diff_0', 'diff_1', 'diff_2', 'diff_3', 'diff_4', 'diff_5', 'diff_6', 'diff_7', 'diff_8', 'diff_9', 'diff_10', 'diff_11'
Single-sensor Optical Dataset	10	'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'BI', 'BU', 'NDVI', 'NDWI'
Single-sensor Radar Dataset	16	'PC1_VH', 'PC1_VV', 'PC1_meanSAR', 'PC1_ratioSAR', 'diff_0', 'diff_1', 'diff_2', 'diff_3', 'diff_4', 'diff_5', 'diff_6', 'diff_7', 'diff_8', 'diff_9', 'diff_10', 'diff_11'

Table 2. List of Variables for each method

4.1.2 Summary of Image Classification Accuracies: Table 3 shows the accuracy assessment of the different methods under Random Forest Classification. All methods used the same number of decision trees (1000). The first four (4) methods showed very close overall accuracies (OA), suggesting that any one of these four (4) methods can be used to produce accurate LULC maps.

One thing to note, however, is the number of variables used in each method, the least of which is the method of Liaw & Wiener (2015) with just eight (8) variables. On the other hand, the method of Kursa and Rudnicki (2010) produced the best overall accuracy but required all twenty-six (26) variables to attain such accuracy.

Comparing the confusion matrices, Sandri and Zuccolotto's method (2006) presented a lower error of commission for agriculture and built-up land cover. Kursa and Rudnicki's method (2010) inched over the other methods in terms of commission error for bare soil land cover. The opposite happened when analyzing the omission errors. Kursa and Rudnicki's method (2010) performed best in classifying agriculture and built-up land cover.

Consequently, Sandri and Zuccolotto's method (2006) dominated the rest for bare soil land cover. The optical based dataset led in terms of omission error for the forest land cover. All methods yielded the same accuracy for water land cover as the first three (3) indicated fewer numbers of unmasked pixels due to the inclusion of the SAR dataset.

The single-sensor radar-based dataset exhibited very poor results. LULC change maps for this method were not generated due to its low classification accuracies.

Model	Class	Producer's Accuracy	User's Accuracy	Overall Accuracy (%)
Multi-sensor with Sandri and Zuccolotto (2006)	Agri	97.95	95.33	96.38
	Built Up	94.74	96.64	
	Bare Soil	96.53	92.67	
	Forest	95.54	100.00	
	Water	100.00	100.00	
Multi-sensor with Liaw and Wiener (2015)	Agri	97.95	95.33	96.06
	Built Up	95.24	93.96	
	Bare Soil	94.63	94.00	
	Forest	95.54	100.00	
	Water	100.00	100.00	
Multi-sensor with Kursa and Rudnicki (2010)	Agri	99.30	94.67	96.54
	Built Up	95.33	95.97	
	Bare Soil	95.30	94.67	
	Forest	95.54	100.00	
	Water	100.00	100.00	
Single-sensor Optical Dataset	Agri	97.35	98.00	96.13
	Built Up	90.32	93.33	
	Bare Soil	95.04	89.33	
	Forest	98.04	100.00	
	Water	100.00	100.00	
Single-sensor Radar-based Dataset	Agri	49.48	63.33	53.23
	Built Up	58.18	64.43	
	Bare Soil	62.26	22.00	

	Forest	42.25	52.67	
	Water	92.11	97.22	

Table 3. Random Forest Classification Statistics

4.1.3 Integration of Different Variable Selection Methods: An integrated LULC map was also generated in this study. It used the methods of Sandri and Zuccolotto (2006), Kursa and Rudnicki (2010) and Liaw and Wiener (2015) and were compared to one another at pixel level. The integrated map is based on the mode of the land cover class pixel values from the three LULC maps. As seen in Table 4, the different LULC maps agreed mostly with one another. Pixels which showed little or no agreement could be attributed to mixed land cover present in one pixel or uncertainties during random forest classification.

Agreement Level	# of Pixels	Percent Distribution
100% Agreement	11860054	91.88%
75% Agreement	1038932	8.05%
50% Agreement	4068	0.03%
No Agreement	5641	0.04%

Table 4. Agreement Map Summary

4.2 Mapping of Conversion of Agricultural Lands to Built-up

Figure 3 illustrates accurately classified agricultural land conversions within the study area. As seen in the figure, the region-of-interest (ROI) must have been agricultural land before or in 2016; by 2019, it must have been built-up land. The conversion in the ROI must have the same area, more or less, as the classified ROI. Errors in area overlaps and gaps were attributed to limitations in spatial resolution as long as they were kept to a minimum.

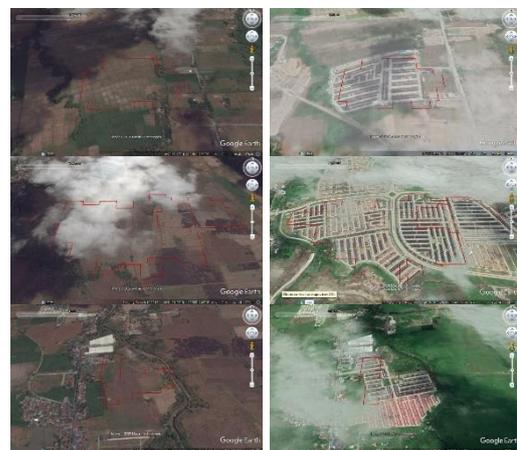


Figure 3. Samples of accurately classified regions-of-interest; From top to bottom: (a) Camella Bucandala, Imus, Cavite, (b) Lancaster Westwood, General Trias, Cavite, (c) Gentree Villas, General Trias, Cavite

The LULC conversion maps were not entirely accurate. Accuracy assessment of the detected conversion of agricultural lands to built-up was implemented using visual inspection with Google Earth Pro. Misclassifications of converted agricultural lands were also detected. Various cases can be attributed due to changes in image color despite showing the same land cover. Another factor was differences in phenological stages. These were attributed to the compositing of the optical datasets. The compositing of different Landsat 8 images with different temporal description resulted in the ambiguity of some regions in the output composite image.

Aside from misclassification, there were also cases of omitted agricultural land conversions. For some, while there were detected conversions, the size of the resulting ROI did not reflect the actual size of the conversion. This was attributed to the misclassification of the model between built-up and bare soil features. The limitation of the study to detect agricultural land conversions to built-up only could have resulted in the omission of other agricultural land conversion to built-up which were rather classified as conversion to bare soil due to the misclassification of the produced RF classification model.

4.3 Spatio-temporal Assessment of Agricultural Land Conversion

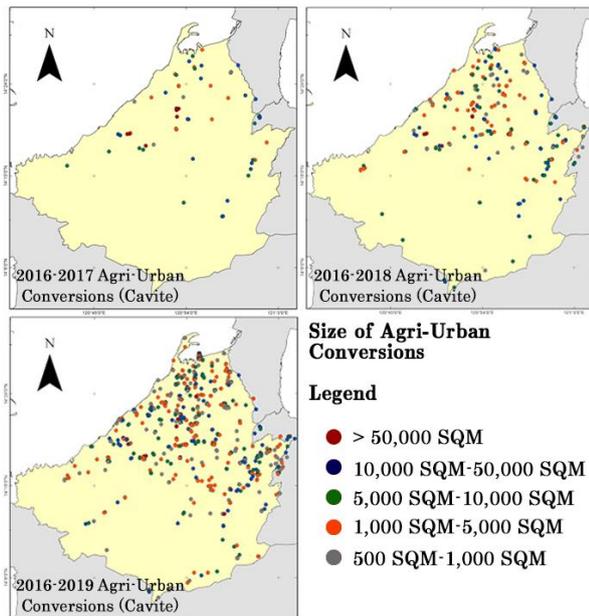


Figure 4. LULC change maps of Cavite from 2016 to 2017, 2016 to 2018, and 2016 to 2019.

From 2016 to 2019, the number of agriculturally converted lands continued to increase. Looking at the maps above, the locations of these lands are in the northern part of Cavite bounded in the same direction by National Capital Region (NCR), the seat of government and a metropolitan area in the country. Agri-urban conversions were classified by size to assess accuracy as detections become smaller. Table 7 shows that along with the number of detected conversions to increase, it can be implied that the total area of conversion also increases. From 59 conversions in the time frame 2016 to 2017, there were an additional 503 detected conversions in the next two years. Also, from just about 605,503 square meters being converted to built-up in the time frame 2016 to 2017, this rose to 1,950,588 in 2018 and 3,558,541 in 2019.

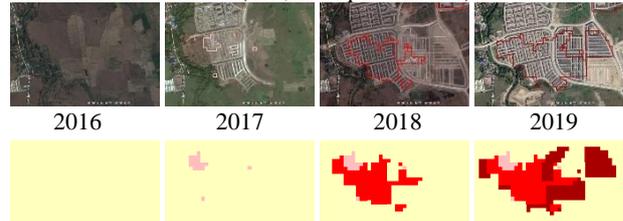
Time frame	Number of Detected Conversions	Total Area of Conversion (sq.m.)
2016-2017	59	605,503
2016-2018	209	1,950,588
2016 - 2019	562	3,558,541

Table 7. Summary of Agricultural Land Conversion from 2016 to 2019

Figure 5 shows examples of detected agricultural land conversion that increased in size over time. The regions in pink exhibited conversion that occurred in the year 2016 to 2017;

those in red showed conversion in the year 2017 to 2018. Finally, those in maroon - 2018 to 2019. The following changes in conversion are very common for subdivisions, usually starting as small projects and expanding as they progress.

Example A: Lancaster Westwood Subdivision, General Trias, Cavite (273,845 square meters)



Example B: Phist Park Homes and EEI ESG Yard, Tanza, Cavite (145,222 square meters)

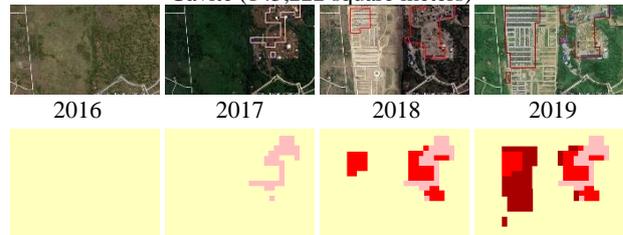


Figure 5. Examples of detected spatio-temporal conversions. Colors of polygons refer to timeframe of conversion – 2016 to 2017 (pink); 2017 to 2018 (red); 2018 to 2019 (maroon)

4.4 Accuracy Assessment of Detecting Agri-Urban Conversions

Comparison of Cumulative Accuracy

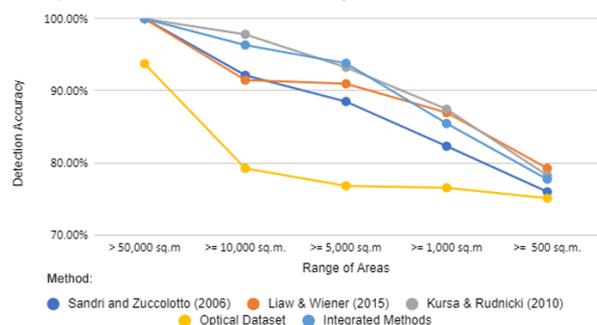


Figure 6. Cumulative Accuracies of different methods in detecting agri-urban conversions.

Figure 6 illustrates a graphical representation of the cumulative accuracies of the different methods as the sizes of the regions-of-interest (ROIs) decreased from greater than 5 hectares or 50,000 square meters down to a range from 500 square meters to 1,000 square meters. The four (4) multi-sensor methods: specifically, Sandri and Zuccolotto (2006), Liaw and Wiener (2015), Kursa and Rudnicki (2010) and the integration of methods, all began from an accuracy of unity and seemed to decrease linearly until ROIs of size 500 to 1,000 square meters where their accuracies lay within the range of 75.97 % to 79.23%. Comparing these methods, Liaw and Wiener's method (2015) and Sandri and Zuccolotto's method (2006) had the highest and lowest final cumulative accuracy, respectively. The use of an exclusively optical dataset remained at the bottom compared to all the methods used in the study; nevertheless, its use was still deemed acceptable for detecting conversions larger than 50,000 square meters.

The overall accuracies of the multi-sensor methods remained close to one another, and the generated LULC Maps and Conversion maps appeared similar. The range of their neighbouring accuracies signified that using any of these methods can be used interchangeably without sacrificing too much on its accuracy.

One thing that distinguished the multi-sensor methods from one another is the number of variables used in forming the model. Sandri and Zuccolotto's method (2006) used twelve (12) variables while Liaw and Wiener's method (2015) utilized eight (8) variables. These two (2) methods proved to be the most efficient as fewer number of variables translate to a lower time and computational complexity. Kursa and Rudnicki's method (2010) proved to be exhausting as it uses all twenty-six (26) variables for it to produce relatively the same results. Finally, the integration of the mentioned three (3) methods was the most burdensome as it required performing the first three methods before taking the mode of the generated LULC maps. Nevertheless, it remained superior in detecting land conversions from agricultural to built-up land cover up in areas 5,000 to 10,000 square meters.

5. CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The research was able to develop a multi-sensor remote sensing approach using Landsat 8 and Sentinel-1 imageries to generate LULC maps of the province of Cavite, Philippines from 2016 to 2019 with overall accuracies at 96%. By employing GIS techniques, LULC change maps were then produced allowing the assessment of the spatio-temporal dynamics of agri-urban land conversion through the dynamic detection and quantification of conversion of agricultural lands to built-up. Using the produced LULC change maps, it was also found out that while single-sensor approach, specifically using optical satellite imagery, can also produce similar image classification accuracies to that of the multi-sensor approach, the developed multi-sensor approach proved to be superior when producing the LULC change maps having the capability to detect agri-urban conversions across ranges of areas.

5.2 Recommendations

As the study is limited to land cover change for agricultural to built-up only, it is recommended to extend the analysis of land cover change to other land cover classes as well. The methodology can also be further improved by developing a component that would allow identification of cultivated and uncultivated agricultural lands. Time-series analysis can also be integrated which may be used for estimation of crop yield. By doing so, the potential of the study could be maximized especially in generating comprehensive spatio-temporal analysis of agricultural lands which could be used by the government, most especially those involved in the agricultural sector in addressing the problem of food security and self-sufficiency in the country.

REFERENCES

Abdullah, A.B., Ito, S. and Adhana, K., 2006, March. Estimate of rice consumption in Asian countries and the world towards 2050. In *Proceedings for Workshop and Conference on Rice in the World at Stake* (Vol. 2, pp. 28-43).

Argamosa, R.J.L., Blanco, A.C., Baloloy, A.B., Candido, C.G., Dumalag, J.B.L.C., DImapilis, L.L.C. and Paringit, E.C., 2018. MODELLING ABOVE GROUND BIOMASS OF MANGROVE FOREST USING SENTINEL-1 IMAGERY. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4(3). doi:10.5194/isprs-annals-IV-3-13-2018

Baghdadi, N. and Zribi, M., 2016. *Land Surface Remote Sensing in Agriculture and Forest*. Elsevier.

Balisacan, A.M. and Ravago, M.L.V., 2003. The rice problem in the Philippines: trends, constraints, and policy imperatives.

Boschetti, M., Nelson, A., Nutini, F., Manfron, G., Busetto, L., Barbieri, M., Laborte, A., Raviz, J., Holecz, F., Mabalay, M.R.O. and Bacong, A.P., 2015. Rapid assessment of crop status: An application of MODIS and SAR data to rice areas in Leyte, Philippines affected by Typhoon Haiyan. *Remote sensing*, 7(6), pp.6535-6557. doi: 10.3390/rs7060653

Cabildo, J., Subingsubing, K. & Reysio-Cruz, M., 2017. Many farms lost to land conversion. *Inquirer.NET*. Available at: <https://newsinfo.inquirer.net/876377/many-farms-lost-to-land-conversion> [Accessed June 20, 2020].

Chuvieco, E., 2016. *Fundamentals of satellite remote sensing: An environmental approach*. CRC press.

David, C.C. and Balisacan, A.M., 1995. *Philippine rice supply demand: prospects and policy implications* (No. 1995-28). PIDS Discussion Paper Series.

Dusseux, P., Corpetti, T., Hubert-Moy, L. and Corgne, S., 2014. Combined use of multi-temporal optical and radar satellite images for grassland monitoring. *Remote Sensing*, 6(7), pp.6163-6182. doi:10.3390/rs6076163

Erasmí, S. and Twele, A., 2009. Regional land cover mapping in the humid tropics using combined optical and SAR satellite data—A case study from Central Sulawesi, Indonesia. *International Journal of Remote Sensing*, 30(10), pp.2465-2478. doi:10.1080/01431160802552728

ESA, 2019. Sentinel-1 - Data Products - Sentinel Handbook. Retrieved January 24, 2020, from <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/data-products>

ESRI, ArcGIS for Desktop. (2014.). Smoothing zone edges with Boundary Clean and Majority Filter. Retrieved April 20, 2020, from <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/smoothing-zone-edges-with-boundary-clean-and-majority-filter.htm>

Fieuzal, R., Duchemin, B., Jarlan, L., Zribi, M., Baup, F., Merlin, O., Hagolle, O. and Garatuza-Payan, J., 2011. Combined use of optical and radar satellite data for the monitoring of irrigation and soil moisture of wheat crops. *Hydrology and Earth System Sciences*, 15(4), pp.1117-1129. doi:10.5194/hess-15-1117-2011

Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3), pp.257-266. doi:10.1016/s0034-4257(96)00067-3

- Genuer, R., Poggi, J.M. and Tuleau-Malot, C., 2010. Variable selection using random forests. *Pattern recognition letters*, 31(14), pp.2225-2236. doi:10.1016/j.patrec.2010.03.014
- Ghosh, S., Chatterjee, N. D., & Dinda, S., 2018. Relation between urban biophysical composition and dynamics of land surface temperature in the Kolkata metropolitan area: a GIS and statistical based analysis for sustainable planning. *Modeling Earth Systems and Environment*. doi:10.1007/s40808-018-0535-9
- He, C., Shi, P., Xie, D. and Zhao, Y., 2010. Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sensing Letters*, 1(4), pp.213-221.
- Idol, T., Haack, B., & Mahabir, R., 2015. Comparison and integration of spaceborne optical and radar data for mapping in Sudan. *International Journal of Remote Sensing*, 36(6), 1551–1569. doi:10.1080/01431161.2015.1015659
- Jensen, J. R., & Lulla, K., 1987. Introductory digital image processing: A remote sensing perspective. *Geocarto International*, 2(1), 65–65. doi:10.1080/10106048709354084
- Jin, Y., Liu, X., Chen, Y., & Liang, X., 2018. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. *International Journal of Remote Sensing*, 1–21. doi:10.1080/01431161.2018.1490976
- Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T. and Reiche, J., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing*, 8(1), p.70.
- King, J. R., & Jackson, D. A., 1999. Variable selection in large environmental data sets using principal components analysis. *Environmetrics*, 10(1), 67–77. doi:10.1002/(sici)1099-095x(199901/02)10:1<67::aid-env336>3.0.co;2-0
- Kursa, M. B., & Rudnicki, W. R., 2010. Feature selection with the Boruta package. *J Stat Softw*, 36(11), 1-13.
- Lefsky, M. A., & Cohen, W. B., 2003. Selection of Remotely Sensed Data. *Remote Sensing of Forest Environments*, 13–46. doi:10.1007/978-1-4615-0306-4_2
- Li, S., & Chen, X., 014. A new bare-soil index for rapid mapping developing areas using landsat 8 data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(4), 139.
- Liaw, A., & Wiener, M., 2015. Package 'randomForest'. Breiman and Cutlers random forests for classification and regression. Version 4.6-12. CRAN. Software.
- Macarof, P., & Statescu, F., 2017. Comparasion of NDBI and NDVI as indicators of surface urban heat island effect in landsat 8 imagery: a case study of Iasi. *Present Environment and Sustainable Development*, (2), 141-150.
- NASA Harvest, 2019. The GEO Global Agricultural Monitoring (GEOGLAM) Initiative, Retrieved September 03, 2019, from <https://nasaharvest.org/project/geo-global-agricultural-monitoring-geoglam-initiative>
- Pal, M., 2003. Random forests for land cover classification. *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*. doi: 10.1109/igarss.2003.1294837
- Pelletier, C., Valero, S., Inglada, J., Champion, N., & Dedieu, G., 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187, 156–168. doi: 10.1016/j.rse.2016.10.010
- Philippine Rice Research Institute, 2019. PRISM - Philippine Rice Information System, Retrieved October 28, 2019, from <https://prism.philrice.gov.ph/>
- Ravanera, R., 2018. Governance of Agricultural Lands, Ancestral Domains, and Aquatic Resources in the Philippines: CSO Land Monitoring Report in the Philippines 2018. *Coalition for Agrarian Reform and Rural Development (ANGOC) and Land Watch Asia (LWA)*. 2019, 10.
- Rikimaru, A., Roy, P. S., & Miyatake, S., 2002. Tropical forest cover density mapping. *Tropical ecology*, 43(1), 39-47.
- Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS, 3rd Earth Resource Technology Satellite (ERTS) Symposium, 1, 48-62
- Sandri, M., & Zuccolotto, P., 2006. Variable Selection Using Random Forests. *Data Analysis, Classification and the Forward Search*, 263–270. doi:10.1007/3-540-35978-8_30
- Sharma, R., Ghosh, A., & Joshi, P. K., 2013. Spatio-temporal footprints of urbanisation in Surat, the Diamond City of India (1990–2009). *Environmental monitoring and assessment*, 185(4), 3313-3325.
- Shaw, G. A., & Burke, H. K., 2003. Spectral imaging for remote sensing. *Lincoln laboratory journal*, 14(1), 3-28.
- Szabó, S., Gácsi, Z., & Balázs, B., 2016. Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories. *Acta Geographica Debrecina Landscape & Environment*, 10(3-4), 194-202.
- Torbick, N., Chowdhury, D., Salas, W., & Qi, J., 2017. Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. *Remote Sensing*, 9(2), 119. doi: 10.3390/rs9020119
- United States Geological Survey, 2020. EarthExplorer - Home, Retrieved April 23, 2020, from <https://earthexplorer.usgs.gov/>
- Vibhute, A. D., & Gawali, B. W., 2013. Analysis and modeling of agricultural land use using remote sensing and geographic information system: a review. *Int. J. Eng. Res. Appl.(IJERA)*, 3(3), 081-091.