Development of Mapping Design for Agricultural Features Extracted from LiDAR Datasets

Nerissa B. Gatdula [1], Mylene V. Jerez [1], Therese Anne M. Rollan [1], Ronalyn P. Jose [1], Coleen Dorothy U. Caranza [1], Joyce Anne Laurente [2] and Ariel C. Blanco [1,3]

[1]*Phil-LiDAR 2 Agricultural Resources Extraction from LiDAR Surveys, Training Center for Applied Geodesy and Photogrammetry, University of the Philippines, Diliman, Quezon City, Philippines*

[2]*Phil-LiDAR 1 Data Archiving and Distribution Component, Training Center for Applied Geodesy and Photogrammetry, University of the Philippines, Diliman, Quezon City, Philippines*

[3]*Department of Geodetic Engineering, University of the Philippines, Diliman, Quezon City, Philippines*

Keywords: Agriculture, LULC, Mapping Design, Post-classification, Geodatabase Schema, LiDAR, Minimum Mapping Unit, Resource Maps, Models, Automated Workflows.

Abstract: Methods for agricultural feature extraction were developed to produce detailed (crop-level) agricultural land use/land cover (LULC) maps from high resolution LiDAR datasets. As of February 2017, available LiDAR data in the Philippines covers 125,200.00 sq.km. or 42.43% of the land area of the Philippines. As part of product generation, definition of mapping design was considered. This includes algorithm for post-classification, development of geodatabase schema, and map layouts. Output maps in custom and 1:10,000 scale JPEG maps, shapefiles and KMZ files are distributed to local government units, national government agencies and other stakeholders for use in planning and other applications. Definition of LULC classes and types is in accordance with the standard codes of Bureau of Soils and Water Management while 1:10,000 is based on National Mapping and Resource Information Authority map indexes. Initial classified maps are maintained in high resolution layers. Detailed objects are refined by determining the Minimum Mapping Unit (MMU). The use of mapping design has standardized the output agricultural resource maps of implementing universities involved in the Phil-LiDAR 2 Program. Models and automated workflows were developed to improve the implementation of the map design.

1 INTRODUCTION

Agricultural land use/land cover (LULC) maps are produced by utilizing Light Detection and Ranging (LiDAR) Technology under the Phil-LiDAR 2 Program Project 1, funded by the Philippines' Department of Science and Technology (DOST) and monitored/co-managed by the Philippine Council for Industry, Energy and Emerging Technology Research and Development. This is to complement programs of the Department of Agriculture and to provide accurate, reliable, and detailed agricultural maps at the crop level. The Project utilizes LiDAR data acquired by the DREAM/Phil-LiDAR 1 Program, other remote sensing systems, and field measurements.

LiDAR, or 3D laser scanning, is an active remote sensing which measures point cloud data at a rate of 100,000 to 500,000 points per second. These high accuracy datasets and derivative layers enable accurate and detailed classification of features on the ground. As of February 22, 2017, available LiDAR data covers 42.43 percent (125,200.00 sq.km.) of the Philippines' land area. LiDAR datasets, which include Digital Surface Models (DSM), Digital Terrain Models (DTM), Orthophotos, and Classified LAZ, are available for distribution.

Agricultural resource maps produced by the Phil-LiDAR 2 Program are disseminated to local government units (LGUs), national government agencies (NGAs) and other partners in order to strengthen collaboration. These maps are used for planning, decision making and development needs. Combined with hazard maps and other thematic maps, vulnerabilities of agricultural resources are assessed.

The Program started in July 2014 with the University of the Philippines Diliman, through the

Gatdula, N., Jerez, M., Rollan, T., Jose, R., Caranza, C., Laurente, J. and Blanco, A.
Development of Mapping Design for Agricultural Features Extracted from LiDAR Datasets.
DOI: 10.5220/0006365902760283
In *Proceedings of the 3rd International Conference on Geographical Information Systems Theory, Applications and Management (GISTAM 2017)*, pages 276-283
ISBN: 978-989-758-252-3

Training Center for Applied Geodesy and Photogrammetry, leading fourteen universities in the nationwide inventory of natural resources. Processes implemented in relation to agricultural feature extraction and resource mapping generate various data layers and outputs in different forms. It was deemed necessary to develop a map design in the mapping of agricultural resources in the Philippines for harmonization and standardization of maps produced by agencies and universities.

2 RELATED LITERATURE

The development of mapping design intends to standardize the output agricultural resource maps of implementing institutions involved in the Program. It is important that maps, through the map design, are able to effectively show the results of analysis. Developer must establish a 'good design' and consider the objective/s and end use of maps. Maps can represent and communicate the results of the analysis to wide range of users. Maps interact with users through the use of map products and how it is represented (Longley et al, 2005). Theoretically, map information is communicated to users through map designs. In practical terms, however, this is not easily achieved (ESRI, 1996).

In a map design process, considerations are enumerated by Robinson et al (1995) as follows: (1) *purpose*, determines what is to be mapped and how message is to be represented; (2) *reality*, defines the phenomena being mapped by limiting the design of the map; (3) *available data*, the specific data type or format affects the design; (4) *map scale*, controls how data should appear; (5) *audience*, wide range of user sees information differently; (6) *conditions of use*, environment on which map is to be used will define the design of the map; and (7) *technical limits*, digital and printed formats are usually processed and represented differently.

The standardization of output maps start with the standardization of procedure for output generation. In this light, development of algorithms for mapping design are considered.

2.1 Minimum Mapping Unit

Post-processing of initial classified maps include the determination of spatial grain or Minimum Mapping Unit (MMU) (Rutchey et al, 2009). MMU is the smallest entity size shown in a map. Several factors are considered in determining the smallest map unit: (1) *data resolution*, correponds to the ground dimension of a single pixel; (2) *map scale*, refers to the ratio between the map distance and ground distance; (3) *classification*, refers to the specific class type of an object; (4) *print size*, corresponds the physical dimension of the map paper; (5) *PPI*, the number of pixels within an inch of printed material; and (6) *viewing distance*, considers the distance of a person looking at a printed map, poster, signage, etc. on display (Spangrud, 2015).

Identified MMU should provide information without losing significant spatial information (Rutchey et al, 2009). In Phil-LiDAR 2 Program, MMU is applied to digital and printed formats, in custom-scale and 1:10,000 scale based on NAMRIA map index.

2.2 Agricultural LULC Schema and Classes

The Department of Agriculture - Bureau of Soils and Water Management (DA-BSWM) released in 2009 the standard codes for thematic mapping, including the classes for LULC maps. Mapping codes are grouped based on the most extensive dominated land use, percent dominant land use, most extensive associated land use, and percent associated land use. Percent distribution ranges from 50% to 100% for the dominant and below 5% to above 30% for the associated land use.

Geodatabase schema stores the spatial attribute data in table and polygon geometry which is maintained through structured query language (SQL) approach, a series of relational functions and operators. Schema is documented in a data dictionary wherein objects in a database, tables, fields in the table, and the relationship between fields and tables are well-defined. Attribute domains are applied to enforce the integrity of the dataset. (ESRI, 2016).

Implementation of proposed map design entails the use of Geographic Information System (GIS), models and automated workflows in order to standardize map production across universities.

3 METHODOLOGY

Mapping design for agricultural LULC maps, including the algorithm for post-classification, development of geodatabase schema, and map layouts, are considered. For coastal municipalities and cities, mangrove and aquaculture classes are integrated into the agricultural maps.

Models and automation workflows were developed to improve the implementation of the map design.

The general workflow for agricultural map design is shown in Figure 1, beginning from the initial classified image to the maps that will handed over to stakeholders. The classification image, showing the agricultural features extracted from LiDAR dataset using object based image analysis (OBIA), is exported as vector file from eCognition and post-classified in ArcGIS. Refinement and post-classification are done by determining the MMU and smoothening LULC polygons. Template geodatabase schema are applied to the refined shapefile to produce a series of LULC maps.

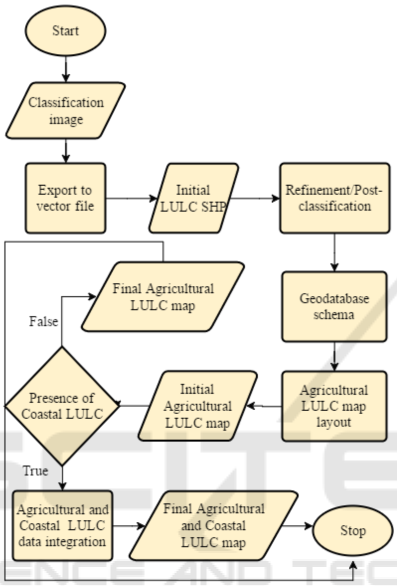

Figure 1: General workflow for the mapping design of agricultural LULC maps.

Definition of LULC classes and types followed the standard codes of Bureau of Soils and Water Management. The standard 1:10,000 scale maps are based on the National Mapping and Resource Information Authority map indexes. These were used so as not to deviate largely from the government's mapping standards.

Test data were used to show the performance of the developed mapping design.

3.1 Refinement/Post-classification Procedure

Initial LULC shapefiles were exported from eCognition and added as layers in ArcGIS. Determination of the MMU depends on the selection of the significant feature with the smallest area. A feature, or class, is considered significant when (1) a class is surrounded by other classes but are near clusters of the same class; (2) a class is abundant alongside a river or road; (3) a class serves as boundaries of a parcel; (4) a class is abundant even within built-up areas; and other related scenarios.

Figure 2 shows the general workflow for LULC post-classification and refinement. Implementation of MMU was carried out through the Elimination and Smoothing tools in ArcGIS.

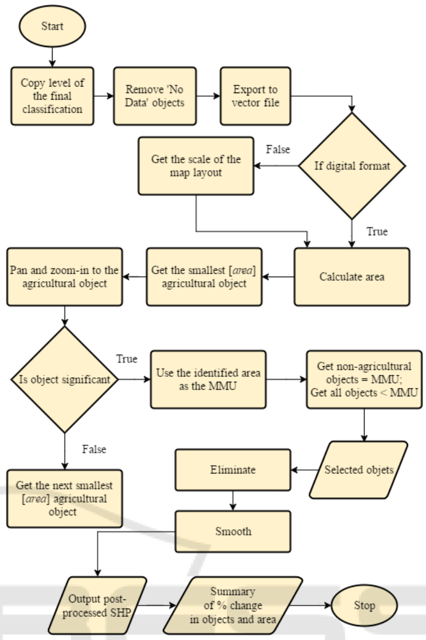

Figure 2: Workflow for the refinement/post-classification of LULC.

Objects are eliminated by determining the smallest significant unit and merging the insignificant objects with neighbouring polygons. Non-agricultural features equal to the identified MMU and features less than the identified MMU are removed. For printed maps, MMU can be computed by applying Equation 1.

$$\text{MMU Ground Area in m} = \left[\frac{(1\ \text{mm})\ \text{x}}{(1\ 000)}\right]^2 \quad (1)$$

As for the smoothing of LULC shapefiles, the maximum smoothing tolerance is set at 3 meters as higher tolerance can alter the shape and area of significant objects. Percent changes in number of objects and polygonal areas for the initial and post-processed/refined LULCs were computed.

3.2 Geodatabase Schema

A geodatabase schema was developed to standardize the output maps of agricultural feature extraction and LULC mapping of fourteen (14) implementing universities.

3.2.1 LULC Classes

Definition of LULC classes is based on the mapping standards developed by government agencies. Each class has assigned code, defined in the table domain and concatenated to come up with the unique identifier.

3.2.2 LULC Schema and Domains

Schema refers to the structure of the agricultural LULC dataset. Table structure consists of class, types and corresponding unique identifier. Data values are restricted by creating domains for 'Classification Types', 'LULC Classes', 'LULC Types', 'Crops', 'Farming System', 'Water body Types', and 'Road Types'.

3.2.3 Automation of Schema Implementation

Automation workflow for schema implementation was developed in order to generate series of maps more efficiently. Model is based on Python and utilizes the ArcGIS geoprocessing toolbox.

Required parameters are the Excel file, where LULC classes are specified, and template geodatabase schema, where LULC data is loaded (See Figure 3).

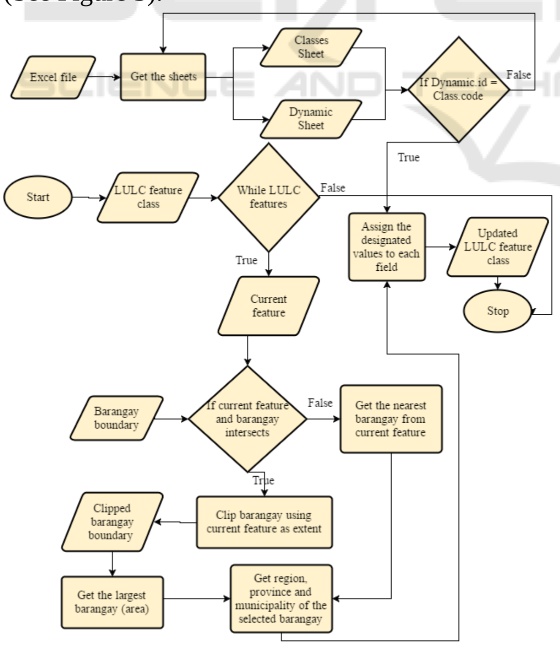

Figure 3: Automation workflow for schema implementation.

3.3 Data Integration

Agricultural and aquaculture resource datasets are combined for coastal areas. There are two cases in data integration: (1) coastal classification was not used as thematic layer in agricultural classification; and (2) coastal classification was used as thematic layer in agricultural LULC classification.

The integration aims to remove the overlapping agricultural and coastal classes. The developed workflow ensured that significant classes in coastal areas are retained. Insignificant classes, mostly 'Road', 'Bare/Fallow' and 'Building', are reclassified as auxiliary layer. Result of data integration is appended to the geodatabase schema (See Figure 4).

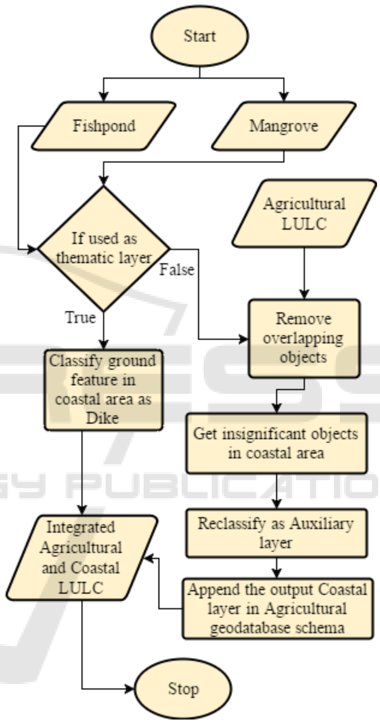

Figure 4: Workflow for agricultural and coastal LULC integration.

3.4 Map Layout

Final agricultural LULC files should be represented in a map series. Mapping templates in custom-scale and 1:10,000 were developed to standardize series of layouts for agricultural LULC maps. Maps in 1:10,000 scale are produced through Data Driven Pages (DDP) and ArcPy, while, maps in custom-scale are generated through ArcMap layout page.

3.4.1 DDP and ArcPy for 1:10,000 Scale Maps

DDP was used in the creation of series of layout pages from a single map document. The capabilities of DDP has been extended using ArcPy, a Python scripting module that automates the exporting and printing of maps. The iteration of map production is based on the selected 1:10,000 indexes with available LULC files (See Figure 5).

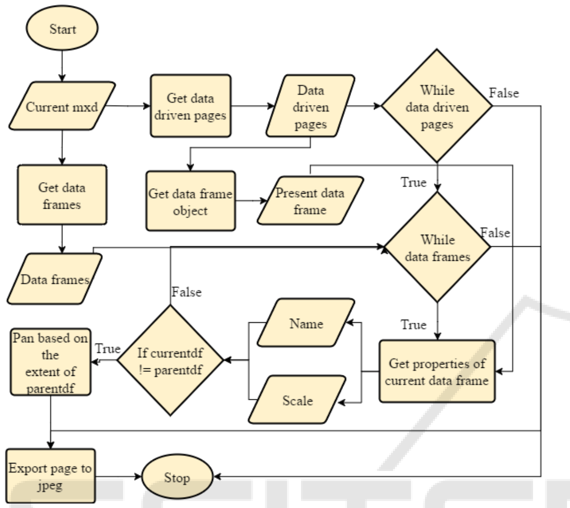

Figure 5: Automation workflow for the mapping of 1:10,000 scale maps.

3.4.2 Template Layout for Custom Scale Maps

Maps are also represented in custom-scale layout. Template was produced and distributed to universities to harmonize map production.

3.5 Production of GIS Files

The final and completed LULC files are handed over to LGUs, NGAs and other stakeholders. Vector files are provided in 1:10,000 scale shapefiles and custom-scale KML/KMZ file formats.

3.5.1 Clip per 1:10,000 Scale Model

Smaller regions of the final LULC files are produced for easier data handling. For better visualization, these shapefiles can be loaded in Google Earth Pro. Figure 6 shows the model for the clipping of 1:10,000 LULC files.

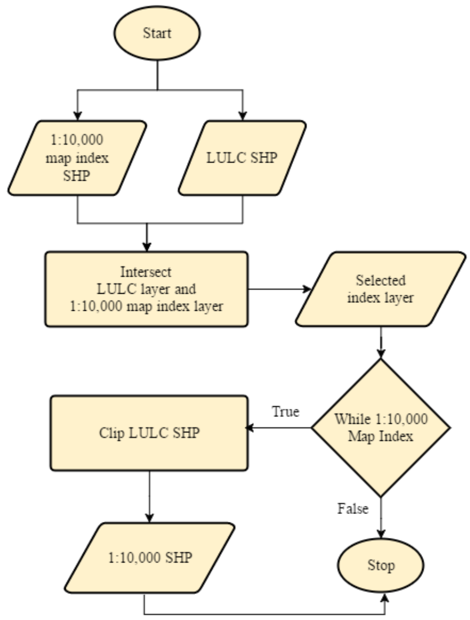

Figure 6: Automation workflow for the clipping of LULC in 1:10,000 scale grids.

3.5.2 LULC Shapefile to KML/KMZ Conversion Model

Keyhole Markup Language is an XML-based format used by Google Earth. Files can be in KML/KMZ file formats. Model for the conversion of shapefiles to KML/KMZ are developed (See Figure 7).

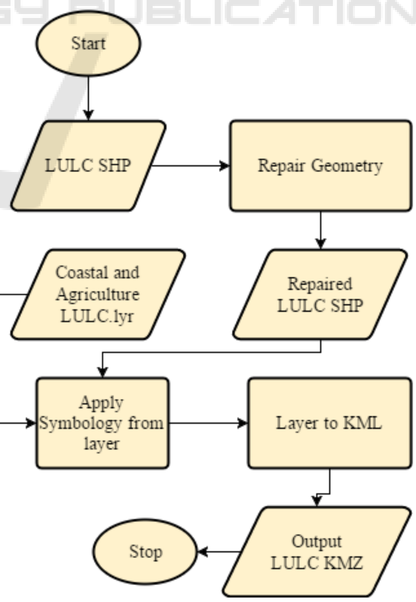

Figure 7: Automation workflow for the conversion of shapefile to KMZ/KML.

4 RESULTS AND DISCUSSION

The initial classified shapefile has been refined to remove unnecessary objects and to set the minimum spatial grain. In MMU identification, the first agricultural class with the smallest area is observed. Table 1 shows the list of agricultural and non-agricultural classes commonly present in the classification.

Table 1: Common LULC classes.

Non-agricultural Class	Agricultural and other Significant Classes
Water	Crop Fields (i.e. Rice, Corn, Sugarcane)
Bare/Fallow	Crop Trees (i.e. Coconut, Banana, Mango)
Building	Mangroves
Road	
Shrubland	
Grassland	

The MMU should be the smallest significant object in LULC map. Figures 8, 9, 10, and 11 show the common scenarios where class is considered important.

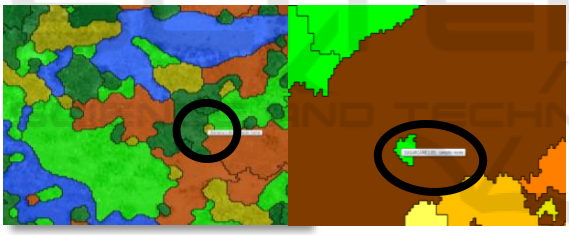

Figure 8: Banana object surrounded by other trees and shrubs but cluster of banana are present nearby (left); sugarcane within Barren parcel (right).

Figure 9: Abundant banana objects alongside rivers (left); abundant non-agricultural trees along roads (right).

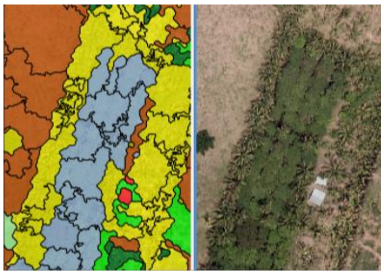

Figure 10: Coconut trees used as boundaries in a Mango plantation.

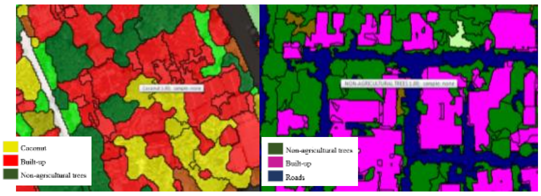

Figure 11: Abundant coconut objects in built-up area (left); abundant non-agricultural trees in built-up area (right).

The iterative application of the refinement process has been tested. Tables 2 and 3 show the percent change in number of objects and area. The percent change may vary depending on how segmentation in eCognition was done. Test Area A (Block 8H) has lower values compared to Test Area B (Block 1C) which may imply that segmentation is relatively good and removal of the salt and pepper effect was done prior to refinement.

Table 2: Percent change in number of objects after refinement.

LULC Class	Test Area A	Test Area B
Bare/fallow	1.61%	-34.45%
Building	0.84%	-29.33%
Developed	2.09%	-29.66%
Grassland	1.22%	0.44%
Mango	1.16%	0.00%
Non-agri	0.81%	-44.28%
Rice	5.00%	-14.77%
Road	1.85%	-45.49%
Water	3.99%	-60.44%

Table 3: Percent change in area after refinement.

LULC Class	Test Area A	Test Area B
Bare/fallow	0.04%	-0.11%
Building	-0.01%	2.02%
Developed	-0.01%	1.09%
Grassland	0.03%	0.17%
Mango	-0.08%	-0.23%
Non-agri	-0.09%	-0.74%
Rice	0.01%	0.04%
Road	0.10%	1.32%
Water	0.00%	0.06%

The developed geodatabase schema is applied to the post-classified/refined LULC shapefile. Table 4 shows the structure of the LULC datasets.

Table 4: Schema of the agricultural LULC map.

Name	Description
CLASSIFICATION	Classification types
RESOURCE_TYPE	Resource map types
ID_CLASS	ID of LULC class
MAIN_CLASS	Main LULC class
OTHER_CLASS1	Other LULC class
OTHER_CLASS2	Other LULC class
CLASS_DESCRIPTION	LULC class description
ID_TYPE	ID of LULC type
MAIN_TYPE	Main LULC type
OTHER_TYPE1	Other LULC type
OTHER_TYPE2	Other LULC type
TYPE_DESCRIPTION	LULC type description
DATA_SOURCE	Source of dataset (e.g. LiDAR, Landsat)
DATASET_ ACQUIRED	Acquisition date of dataset
FARMING_SYSTEM	Farming system
CROP_PLANTING_PERIOD	Farming period of crops (e.g. Dec-Feb)
JAN	Crop planted in Jan
FEB	Crop planted in Feb
MAR	Crop planted in Mar
APR	Crop planted in Apr
MAY	Crop planted in May
JUN	Crop planted in Jun
JUL	Crop planted in Jul
AUG	Crop planted in Aug
SEP	Crop planted in Sep
OCT	Crop planted in Oct
NOV	Crop planted in Nov
DEC	Crop planted in Dec
AREA	Area of LULC
REGION	Region of the main City/Muni
PROVINCE	Province of the main City/Muni
CITYMUNI	Main City/Muni
BARANGAY	Main Barangay of the main City/Muni
REMARKS	Remarks

Identification of the main class should be based on the dominant crop in the area. Dominance is based on height for intercropping systems and on hectarage for mixed cropping systems.

The manual implementation of schema requires doing some of the processes repeatedly. Thus, automation workflow for the application of schema were developed. Based on benchmark testing, approximately 1,500 to 1,800 features were updated per hour. This translates to 25 to 30 features per minute. Processing time was observed to be highly dependent on the size and number of features.

In agricultural and coastal data integration, insignificant objects found in the coastal area are reclassified as auxiliary layer (See Figure 12).

Figure 12: Roads observed in fishpond area (left) and overlapping fishponds and bare objects (right).

Final LULC maps are produced in 1:10,000 scale and custom-scale JPEG files. Figure 13 shows sample agricultural and coastal LULC maps in custom-scale layout.

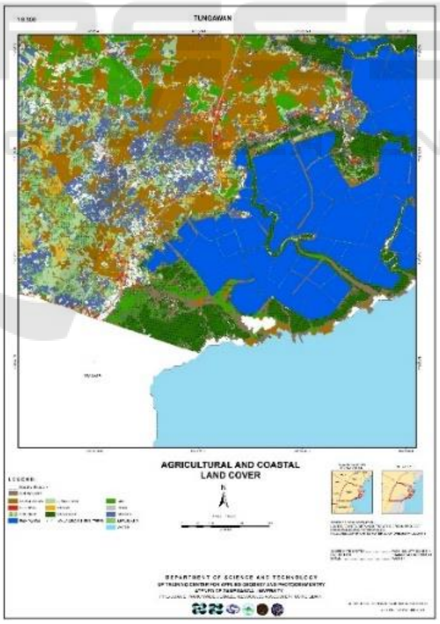

Figure 13: Custom scale layout of agricultural and coastal LULC map.

Vector files, in 1:10,000 scale shapefiles and custom-scale KML/KMZ file formats, are also generated (See Figures 14 and 15).

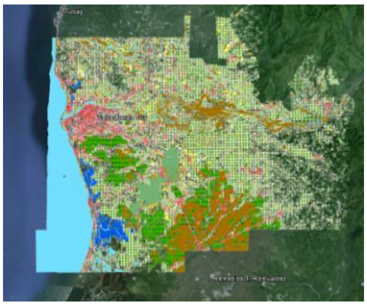

Figure 14: KML/KMZ maps in custom-scale.

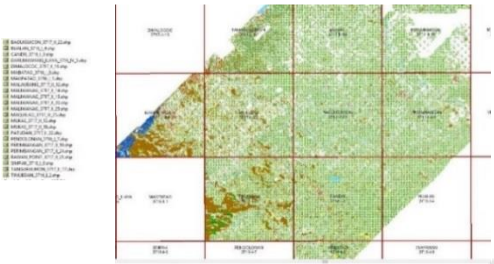

Figure 15: LULC shapefiles in 1:10,000 subset files.

5 SUMMARY AND CONCLUSIONS

The development of mapping design was considered in the production of agricultural LULC to ensure the standardization of maps disseminated to various stakeholders. The design has been useful in the management of spatial information and maintenance of LULC database. Model and scripts using ArcGIS, ArcPy and Python are utilized in the production of LULC maps, resulting in faster turn-around from data to map products.

ACKNOWLEDGEMENTS

The research is made possible through the funding of Philippines' Department of Science and Technology and monitoring of the Philippine Council for Industry, Energy and Emerging Technology Research and Development. Technical facilities and support were provided by the UP College of Engineering, the UP Department of Geodetic Engineering and the UP Training Center for Applied Geodesy and Photogrammetry. The research is also supported by the Philippines' Department of Agriculture through the Information Technology Center for Agriculture and Fisheries.

The authors acknowledge everyone who has contributed to the research.

REFERENCES

Buckley, A. (2008). *Guidelines for minimum size for text and symbols on maps.* Retrieved from https://blogs.esri.com/esri/arcgis/2008/01/16/sizefortextandsymbolsonmaps/

Bureau of Soils and Water Management (2009). *Manual on Map Standards & Symbols for Soil & Water GIS.*

Carating, R., Manguerra, J., Samalca, I., Pascual, N., Albano, F. (2009). *Manual on Map Standard & Symbols for Soil & Water GIS. Bureau of Soils and Water Management.*

Conley, E. P. (2013). *2013 ESRI International User Conference. Generalization for Multi-scale Mapping*. San Diego, California.

Environmental Systems Research Institute, Inc. (1996). *Introduction to Map Design.* Retrieved from http://healthcybermap.org/HGeo/res/intrcart.pdf

Environmental Systems Research Institute, Inc. (2007). *ArcGIS Desktop Help*. Retrieved from http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Simplifying_and_smoothing_features.

Environmental Systems Research Institute, Inc. (2016). *The architecture of a geodatabase.* Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/manage-data/geodatabases/the-architecture-of-a-geodatabase.htm

Environmental Systems Research Institute, Inc. (2016). *An overview of attribute domains.* Retrieved from https://pro.arcgis.com/en/pro-app/help/data/geodatabases/an-overview-of-attribute-domains.htm

Longley, P., Goodchild, M., Maguire, D., & Rhind, D. (2005). Geographic Information Systems and Science. *ISBNs: 0-470-87000-1*.

Philippe Rigaux, M. S. (2002). *Spatial Database with Application to GIS*

Rutchey, K., & Godin, J. (2009). Determining an appropriate minimum mapping unit in vegetation mapping for ecosystem restoration: a case study from the Everglades, USA. *Landscape Ecol 24:1351–1362.* Springer Science+Business Media B.V

Spangrud, D. (2015). *A Question of Scale, Resolution, and MMU.* Retrieved from https://blogs.esri.com/esri/esri-insider/2015/05/21/a-question-of-scale-resolution-and-mmu/ on 21 Jan 2016

Stohlgren, T.J. (2006). *Measuring Plant Diversity: Lessons from the Field*. Oxford University Press, USA.

Stocksigns (2012). *What size sign should I use? A viewing distance guide.* Retrieved from http://blog.stocksigns.co.uk/size-sign-viewing-distance-guide/

T. W. Lillesand and R. W. Kiefer (2004). *Remote Sensing and Image Interpretation.* New York: Wiley, p. 157.

USGS Land Cover Institute. *NLCD 92 Land Cover Class Definitions*. Retrieved from http://landcover.usgs.gov/classes.php.