TIME-SERIES LAND COVER AND LAND USE MONITORING AND CLASSIFICATION USING GIS AND REMOTE SENSING TECHNOLOGY: A CASE STUDY OF BINH DUONG PROVINCE, VIETNAM

PhD DISSERTATION

BUI DANG HUNG

Supervisor:

DR. HABIL. MUCSI LÁSZLÓ

associate professor

Doctoral School of Geosciences

Faculty of Science and Informatics

University of Szeged

SZEGED, 2023

Table of contents

List of tables

List of figures

Abbreviations and acronyms

AA	Agriculture with annual plants
AD	Allocation disagreement
AER	Annual expansion rate
AP	Agriculture with perennial plants *(for land use)*
AP	Annual plants *(for land cover)*
AREA_MN	Mean Patch Size
AUC	Area under the curve
BL	Barren land
BL_H	Bare land with high albedo
BL_L	Bare land with low albedo
BOA	Bottom of atmosphere
BPA	Basic Probability Assignment
BU_H	Built-up with high albedo
BU_L	Built-up with low albedo
CAD	computer-aided design
CONTAG	Contagion Index
CORINE	European Union's Coordination of Information on the Environment
DEM	Digital elevation model
DF	Decision Forest
D-S	Dempster-Shafer
ECR	Expansion contribution rate
FLS	Full Lambda Schedule
FoM	Figure of merit
GADM	Database of Global Administrative Areas
GIS	Geographic information system
GLCM	Gray-level co-occurrence matrix
GR	Grass
GRD	Ground Range Detected

IC	Industry and commerce
IJI	Interspersion and Juxtaposition Index
IS	Impervious surface
IW	Interferometric Wide Swath
JAXA	Japan Aerospace Exploration Agency
L-8	Landsat-8
LCCS	Food and Agriculture Organisation's Land Cover Classification System
LCM	Land Change Modeler
LiDAR	Light Detection and Ranging
LPI	Largest Patch Index
LSI	Landscape Shape Index
MR	Mixed residence
MS	Mining site *(for land use)*
MS	Multi-spectral
MSI	Multispectral Instrument
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NP	Number of Patches
OA	Overall accuracy
OLI	Operational Land Imager
OOB	Out-of-bag
PA	Producer's accuracy
PD	Patch Density
PLAND	Percentage of Landscape
PP	Perennial plants
QD	Quantity disagreement
QGIS	Quantum Geographic Information System
RF	Random forest

RG	Recreation and green space
ROC	Receiver operator characteristic
RS	Remote sensing
S-1	Sentinel-1
S-2	Sentinel-2
SAR	Synthetic aperture radar
SHDI	Shannon’s Diversity Index
SHEI	Shannon’s Evenness Index
SNAP	Sentinel Application Platform
SRTM	Shuttle Radar Topography Mission
UA	User’s accuracy
UL	Unused land
USGS	United States Geological Survey
VE	Vegetation
VH	Vertical transmit-horizontal receive
VV	Vertical transmit-vertical receive
WA	Water surface
WS	Water surface
WGS84	World Geodetic System 1984

1. Introduction

1.1. Background

1.1.1. Land cover and land use

Land cover and land use information plays an important role in monitoring the environment and natural resources as well as in urban management (Rimal et al. 2017; Arowolo et al. 2018; Grigoraș and Urițescu 2019). Therefore, the knowledge of the spatial distribution and pattern of them in a specific area is necessary.

Land cover is defined as "the observed (bio)physical cover on the earth's surface" (Di Gregorio 2005), ranging from natural objects, such as vegetation, water surface, bare rock, and bare soil, to artificial objects, such as buildings, roads, etc. Meanwhile, land use refers to "the arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it" (Di Gregorio 2005); in other words, land use is the way in which people use land cover types for one or more different purposes. For example, the forest is a land cover type. However, the way people use a forest determines its land use type. It can be used for logging, conservation, or recreation purposes. Similarly, a building (land cover type) can be used for residential, industrial, commercial, or entertainment activities (land use types), depending on the intention of its owner. In practice, a land cover type may be used for various purposes (like the examples above), while a land use type may also consist of one or many land cover types (Cihlar and Jansen 2001; Giri 2012), for example, an entertainment complex may include built-up, vegetation, and water surface. Furthermore, there is a connection between land cover and land use (Jansen and Di Gregorio 2003; Kim 2015). To some extent, the connection can help interpret land use information from land cover information and vice versa (Cihlar and Jansen 2001; Brown and Duh 2004).

Although they are defined differently and this issue has been discussed in previous studies (Cihlar and Jansen 2001; Brown and Duh 2004; Kim 2015), these two terms are still commonly used concurrently or interchangeably in many studies related to land cover and land use classification and mapping (Steinhausen et al. 2018; Carranza-García et al. 2019; L.H. Nguyen et al. 2020). This problem may sometimes cause ambiguity or confusion for readers or map users (Comber et al. 2008), as well as certain difficulties in using such maps, because land use information is often used for planning (Tapiador and Casanova 2003) and making policy (van Delden et al. 2011), while land cover information is often employed in environmental monitoring (Henits et al. 2017), modeling (Shooshtari and Gholamalifard 2015), and prediction (Rizeei et al. 2016).

1.1.2. Land use/land cover change and landscape pattern

Land use/land cover change is the conversion from a land use/land cover type to another type. In general, according to Giri (2012), the major types of conversion include (1) the

conversion of land cover types due to a land use change, for example, the conversion from vegetation to built-up due to the construction of residential areas on cultivation land; (2) the modification between land cover types without changing land use purpose, for example, the transition between crops and bare soil in agricultural activities; and (3) the conversion of land use types without changing land cover type, for example, a forest is converted from recreation to conservation purposes to conserve an endangered species.

Land use/land cover changes are caused by both natural and anthropogenic factors (Serra et al. 2008; Msofe et al. 2019). Natural factors include seasonal changes in weather, changes in water levels due to hydrometeorological cycles, alluvial accretion, and natural disasters such as hurricanes, floods, tsunamis, landslides, volcanic eruptions, wildfires, earthquakes, etc. In terms of anthropogenic factors, activities and policies relating to urban expansion, industrialization, agricultural development, and exploitation of natural resources strongly influence land change. Among them, urbanization and industrialization often lead to rapid, strong, and one-way transformation, especially in developing countries (Pham and Yamaguchi 2011; Kantakumar et al. 2016; Rimal et al. 2017; Fenta et al. 2017; Andrade-Núñez and Aide 2018; Cao et al. 2019; Sumari et al. 2020).

Land use/land cover changes and urban expansion have various impacts on the environment and human life, such as run-off characteristics (Sajikumar and Remya 2015), landscape pattern (Zhang et al. 2010; Dadashpoor et al. 2019), land surface temperature (Zhang and Sun 2019), soil erosion (Nampak et al. 2018), environmental quality (Kovács et al. 2019), as well as biodiversity and ecosystem services (Tolessa et al. 2017; Trisurat et al. 2019). Therefore, information on land use/land cover changes is crucial to resource and environmental monitoring as well as land management policymaking (Nampak et al. 2018). However, the availability of accurate information on spatiotemporal land use/land cover changes, urbanization status, urbanization rates, and their driving factors in localities is often untimely even though it is essential (Kantakumar et al. 2016).

As mentioned above, land use/land cover change affects landscape patterns and, as a result, ecosystem functions (Lin et al. 2013; Estoque and Murayama 2016; Tolessa et al. 2017; Kertész and Křeček 2019; Tang et al. 2020). Therefore, quantification of changes in landscape patterns, including shape, size, and spatial distribution, is also essential, especially where land use change is dramatic, such as in emerging urban areas. The quantification facilitates comparison and assessment of landscape change during past and future land use change. At the same time, it can also partly reveal the impact trend of land use changes on the structure and function of diverse types of landscapes and ecosystems. This information may also be useful for decisionmaking and land use planning toward efficient use of resources and sustainable development (Vaz et al. 2014; Abdolalizadeh et al. 2019). The landscape pattern change is often assessed through

landscape metrics at the three levels including patch, class, and landscape (Turner and Gardner 2015; Gergel and Turner 2017; Gudmann et al. 2020). Land use/land cover maps are often used as input to calculate landscape metrics on geographic information systems.

1.1.3. Remote sensing, geographic information system, and data fusion in land cover and land use study

Remote sensing (RS) is defined as "the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance (typically from satellite or aircraft). Special cameras collect remotely sensed images, which help researchers 'sense' things about the Earth" (USGS 2022). RS databases are increasingly diverse in quantity and quality, meeting different needs. With easy access and acquisition of images, such as MODIS, Landsat, and Sentinel, research related to the interpretation of RS imagery has become proactive, cost effective, and reproducible. Moreover, the development of image processing and classification techniques has increasingly improved the accuracy of results (Lu et al. 2011; Shao and Lunetta 2012; Noi and Kappas 2017; Toure et al. 2018; Quan et al. 2020). RS is a essential tool for land cover and land use mapping and monitoring due to its efficiency, economic benefits, and reliability (Toure et al. 2018; Cai et al. 2019).

A geographic information system (GIS) is defined as "a system that creates, manages, analyzes, and maps all types of data. GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). […] GIS helps users understand patterns, relationships, and geographic context" (ESRI 2022). GIS, and more broadly geographic information science – or GIScience for short, is considered a multidisciplinary science (Blaschke and Merschdorf 2014). It is related to geography, cartography, computer science, information science, geology, geodesy, RS, photogrammetry, ecology, statistics, urban planning, and others. Therefore, it supports working with and combining multidisciplinary tools and datasets. In the field of land use and land cover study, the combination of RS and spatial analysis techniques in GIS allows researchers to detect and to analyze land cover and land use change more easily and timely. This has been confirmed in many studies in the literature on a local (Wu et al. 2006; Rawat and Kumar 2015; Tadese et al. 2020), national (Sánchez-Cuervo et al. 2012; Schoeman et al. 2013; Xu et al. 2020), continental (Mertes et al. 2015; Netzel and Stepinski 2015) and global scale (Giri et al. 2013; X. Li et al. 2017). In addition, GIS also supports future land use simulation. There are many models developed for land-change simulation, such as CLUE-S, CLUMondo, Land Change Modeler (LCM), LucSim, DinamicaEGO, SLEUTH, etc. Each model has its own pros and cons, and the choice of model to use depends on the goals and the available data of the study (Camacho Olmedo et al. 2018). LCM is one of the popular applications used to assess and simulate land use change. The advantage of this application is that it is simple to use and easy to set up input parameters, has clear instructions, and is integrated

with many simulation algorithms. Many studies have used this application for land use change prediction for various purposes (Megahed et al. 2015; Nor et al. 2017; Islam et al. 2018; Mishra et al. 2018; Lennert et al. 2020).

Data fusion is defined as a technique that "combines data from multiple sensors, and related information from associated databases, to achieve improved accuracy and more specific inferences than could be achieved by the use of single sensor alone" (Hall and Llinas 1997). In the earth observation field, the rapid development of different kinds of sensors and data sources has made data fusion a vital research approach that aims to extract more detailed information from the RS imagery (Solberg 2006; Zhang 2010; Schmitt and Zhu 2016). By different fusion methods ranging from simple to complex, the extracted information can effectively serve various fields such as urban management (Guan et al. 2017; Shao, Cheng, et al. 2021; Shao, Sumari, et al. 2021), agriculture (Mfuka et al. 2020; Prins and Van Niekerk 2020), environmental monitoring (Xu and Ma 2021), etc. In general, RS data is fused at three common levels: pixel level, feature level, and decision level (Pohl and van Genderen 2016).

For land cover and land use classification and monitoring, optical and radar data are two types of RS data that are often used as the input for various fusion methods to achieve better mapping results. Some prominent recent studies can be mentioned as the fusion of Sentinel-1 (S-1) and Sentinel-2 (S-2) data at the pixel level (Tavares et al. 2019), S-1, S-2, multi-temporal Landsat-8 (L-8) and digital elevation model (DEM) (Liu et al. 2018), L-8 and Terra SAR-X textures images at the feature level (Tabib Mahmoudi et al. 2019), S-1 and Gaofen-1 images at the decision level (Shao et al. 2016), Quickbird multi-spectral and RADARSAT synthetic aperture radar (SAR) data at the decision level (Ban et al. 2010), light detection and ranging (LiDAR), S-2, and aerial imagery (Prins and Van Niekerk 2020), and Landsat images and Twitter's location-based social media data (Shao, Sumari, et al. 2021). In addition, there have also been attempts to use single or multiple RS data independently (Cihlar and Jansen 2001; Jansen and Di Gregorio 2003; Zhang and Wang 2003) or in conjunction with other ancillary data sources, such as census data (Hunt et al. 2001), land use inventory data (Bauer and Steinnocher 2001), social sensing data (Y. Zhang et al. 2017), and mobile-phone positioning data (Jia et al. 2018), to extract a land cover map and then translate it into a land use map, via a set of parameters and decision rules based on expert knowledge. These study results demonstrate that fusion data from various sources at the three fusion levels can improve accuracy in land cover and land use mapping.

However, there are various fusion techniques ranging from simple to very complex methods, and selecting which fusion method should be applied to deliver the best results is a challenge. In general, selecting a method for image classification depends on many factors. The factors comprise the purpose of study, the availability of data, the performance of the algorithm, the computational resources, and the analyst's experiences (Lu and Weng 2007). In addition, the performance of each method also

depends partly on the characteristics of the study area, the dataset used, and how the method works. A method can yield highly accurate results in one dataset and give poor results in others (Xie et al. 2019). Moreover, it is not necessary to employ a complicated technique when a simple one can solve the problem well.

1.1.4. Land use and land cover maps in Vietnam

In Vietnam, land use status maps are produced by the government at the local and national levels every five years. The basis for producing such maps consists of inventory data related to land changes, including land allocation, land lease, and change of land use purpose during the five-year inventory period (Minister of Natural Resources and Environment of Vietnam 2018). The land use categories for the maps are up to five levels and 56 classes, which are very detailed and complex. The first version of this type of map was generated in 2005 (since the Act on Land 2003 was passed and implemented); thus, there is a lack of spatial data on land use for previous years. In addition, the five-year interval to produce the maps is not suitable for real-time management and research. Furthermore, the maps are extracted on computer-aided design (CAD) file formats, which make it difficult and time-consuming to convert and use them for geographic information system analysis. Last but not least, it is difficult to access this data due to the government's relatively complicated administrative system.

Meanwhile, a land cover map has not yet been produced by the government of Vietnam. Only a few studies have been done independently by researchers or organizations at the national (ADPC 2020; JAXA EORC 2020) or local level (Linh et al. 2012; Disperati and Virdis 2015; H.T.T. Nguyen et al. 2020). However, depending on the purpose of each study, there are many dissimilarities in the land cover classification scheme as well as in the definitions of the classes in those schemes. In addition, a number of these studies also used the terms "land cover" and "land use" concurrently or interchangeably, which may cause a number of obstacles in using these maps, as noted. It should be highlighted that except for the SERVIR-Mekong (ADPC 2020) and Japan Aerospace Exploration Agency (JAXA EORC 2020) programs for land use and land cover mapping for the whole of Vietnam, there are no other studies related to land cover classification in the selected study area.

1.1.5. Study area

Binh Duong province is located in southeast Vietnam, between 10°51′46″ and 11°30′ N latitude and between 106°20′ and 106°58′ E longitude. The total area of the province is over 2,694.64 km^2, and its population is about 2.5 million people as of 2019 (Binh Duong Statistical Office 2020). Administratively, as of 2020, the province was divided into five urban districts (also known as cities and towns) including Thu Dau Mot, Di An, Thuan An, Tan Uyen, and Ben Cat, and four rural districts including Bac Tan Uyen, Bau Bang, Dau Tieng, and Phu Giao (Figure 1.1). Thu Dau Mot city is the administrative–economic–cultural center of the province.

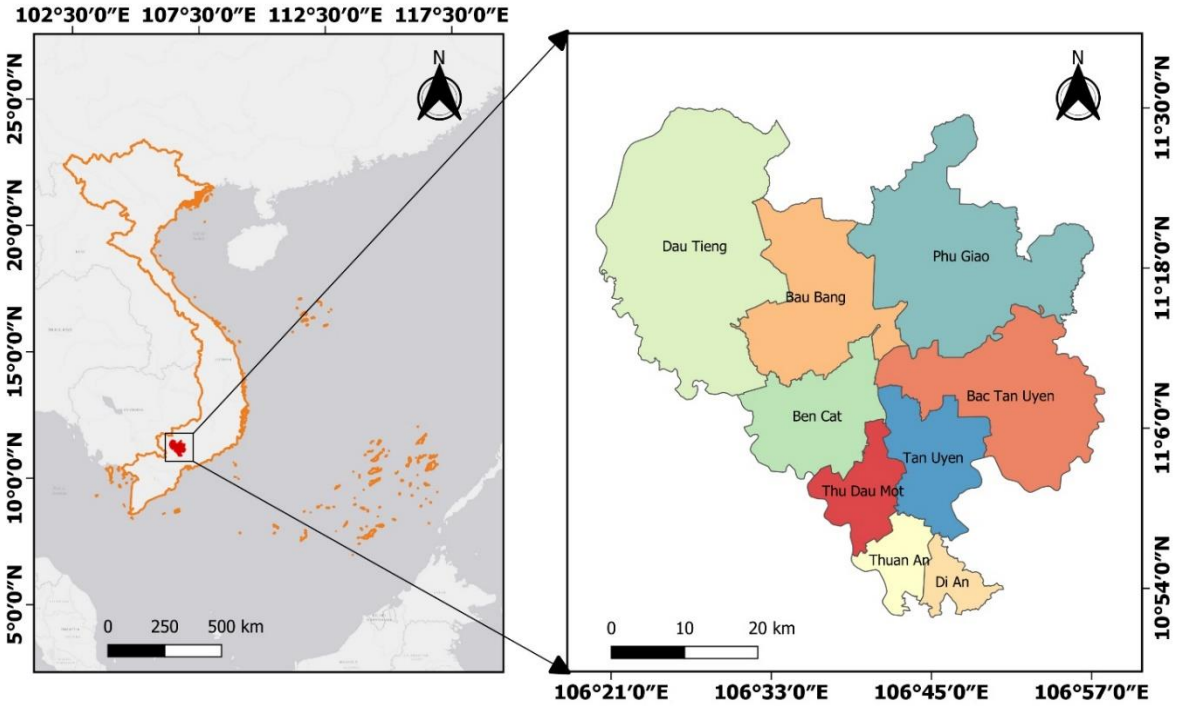


Figure 1.1. The study area

Located in the tropical monsoon region, there are two alternating seasons of Binh Duong's climate: a rainy season from May to November and a dry season from December to April of the following year. The climate is mostly warm all year round, the average temperature in the period of 2015–2019 is about 27.8°C. The difference in temperature between months is not too high, about 3°C–5°C. The average air humidity is from 70% to 96%. The average annual total rainfall is about 2,275.7 mm, of which the rainy season accounts for 90%. The province is surrounded by the Sai Gon-Dong Nai River system with many canals and small streams. The terrain has an average elevation of 20–25 m and is relatively flat with a slope of 3°–15°. More than 80% of the soil constituents in Binh Duong province are acrisols and ferralsols, which are favorable for perennial cropping (Binh Duong Statistical Office 2020; Department of Natural Resources and Environment of Binh Duong Province 2020).

In terms of socioeconomics, Binh Duong belongs to the Southern Key Economic Zone of Vietnam. Since its reestablishment in 1997, the urbanization and industrialization process of the province has been extremely rapid. In a period of ten years from 1995 to 2005, the urbanization rate only rose from 17.51% to 30.09%; however, in the following ten years (i.e., from 2005 to 2015), this rate grew rapidly from 30.09% to 76.72%. As of 2019, the rate reached 79.87% (General Statistics Office of Vietnam 2020). The first industrial park, i.e., Song Than 1, was established in 1995. As of 2019, Binh Duong has 29 industrial parks and 12 industrial clusters, with an average occupancy rate of over 70% in which more than 90% of many of them have been filled. Binh Duong is currently considered the "industrial capital" of Vietnam. The development of industry has considerably contributed to the economic development of the province. The gross regional domestic product at current prices increased from VND

3,915 billion in 1997 (industry and construction accounted for 50.4%) to VND 48,761 billion in 2010 (industry and construction accounted for 63%) and 360,797 billion in 2019 (industry and construction accounted for 66.77%) (Binh Duong Statistical Office 2016; Binh Duong Statistical Office 2020).

1.2. Problem statement

Owing to the expansion of the urban and industrial areas as well as other human activities, the land use in Binh Duong province have significantly changed from 1995 to 2020. The change may continue into the future. However, a study on spatiotemporal land use changes and urban expansion in Binh Duong is still a gap. Such study is necessary because it helps explore not only the pattern of land use change and urban expansion but also the factors influencing these processes. In addition, the effects of the change on the landscape pattern can also be revealed. From there, some practical experience can be learned for land use planning and policymaking in other areas not only in Vietnam but also in other countries.

To fill this gap, it is necessary to use land use maps or land cover maps at different times in the study period as the input for spatial analysis in GIS. However, as mentioned in Section 1.1.4, the land use status map of the province has only been released since 2005 by the government with very complex categories, and no land cover map has been released. Therefore, a prerequisite for this study is to generate such maps with a more generalized category system from 1995 to 2020 as well as to simulate the maps in the next decade.

With the availability of various satellite data sources and the development of new image processing and spatial analysis techniques, there is a potential for combining them in land use land cover mapping and prediction to get highly accurate maps for the need. Obviously, it is easier to observe and classify land cover types directly from aerial or satellite images than to do so with land use types (Zhang and Wang 2003; Giri 2012). However, because they may have a connection, land use types can be interpreted from land cover information once this relationship is clearly defined. Furthermore, it is essential to compare the effectiveness of different approaches to land use land cover mapping to choose the optimal one based on the data availability in the study area and the objective of the study.

1.3. Research objective and hypotheses

The main objective of this study is to use and to develop GIS and RS techniques for time-series land cover and land use monitoring and classification from 1995 to 2020 and prediction to 2030 for Binh Duong province of Vietnam. The hypotheses of this study are that:

(1) There is a connection between land cover and land use, and this connection can be measured and analyzed by geospatial information techniques in Binh Duong province.

(2) There are diverse effects of data sources, data structure, image processing, and fusion technique on land use land cover classification efficiency, and it is possible to select an optimal mapping approach given the data availability in the study area and the objective of the study.
(3) There is a significant change in land use patterns of the study area from 1995 to 2020.
(4) The urban expansion process in the study area varies both spatially and temporally during the study period.
(5) It is possible to predict future land use of the study area based on various natural and socioeconomic factors.
(6) Land use change and urban expansion cause significant changes in landscape patterns of the study area.

1.4. Data, methods and workflow

The overall workflow of this dissertation is illustrated in Figure 1.2. Specific descriptions of the data and methods used are detailed in the Materials and Methods section of Chapters 2, 3, 4, and 5. Below are just brief summaries to provide an overview.

For satellite imagery, the optical and SAR images of the study area acquired during the study period were investigated and collected. Landsat-5, -7, and -8 Collection 1 Level 2 surface reflectance images were ordered and downloaded from the United States Geological Survey (USGS) website (via the link https://earthexplorer.usgs.gov/). S-1 Level-1 Ground Range Detected (GRD) and S-2 Multispectral Instrument (MSI) Level-2A images were downloaded from the Copernicus Scientific Data Hub (via the link https://scihub.copernicus.eu/).

Ancillary data were collected from a variety of sources. The administrative boundary data were downloaded from the Database of Global Administrative Areas project website (via the link https://gadm.org/). The training and validation data were collected based on the field survey, Google Earth history images, and my personal experiences. Census data were collected from the provincial statistical yearbooks and from the website of the General Statistics Office of Vietnam (via the link https://www.gso.gov.vn/). In addition, the Shuttle Radar Topography Mission (SRTM) DEM was downloaded from the USGS website. Population density raster data were downloaded from the WorldPop website (via the link https://www.worldpop.org/). The road network map, land use status maps, and planning maps were collected from the provincial government. Other vector data were extracted from the OpenStreetMap project (via the link https://www.openstreetmap.org/) and downloaded from the GEOFABRIK website (via the link https://download.geofabrik.de/).

A field survey trip to the study area was conducted between 18 January and 18 February 2020 to collect ancillary data and gain a deeper understanding of land cover

and land use in the study area. The ArcGIS Collector application was used on this trip to take geotagged photos.

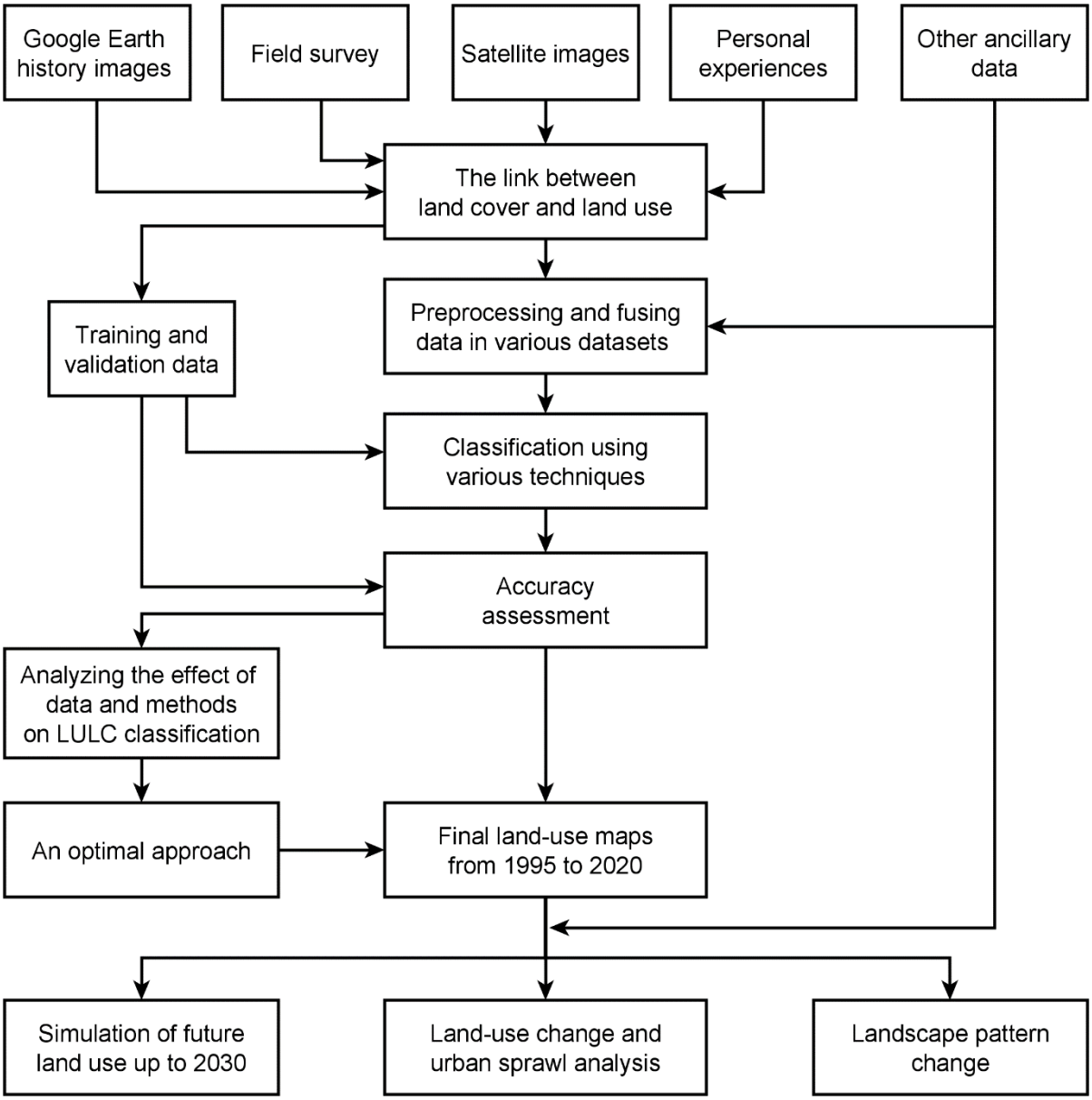


Figure 1.2. Overall workflow of the dissertation.

In terms of methods, in order to solve the research hypotheses and achieve the research objective, I used and developed a series of RS and GIS techniques in this dissertation. They consisted of (1) image processing techniques for preprocessing optical and SAR data, extracting spectral indices and gray-level co-occurrence matrix (GLCM) textures, and combining data at different levels, (2) land use land cover classification using pixel-based and object-based approaches, Dempster-Shafer (D-S) theory, spatial analysis, decision rules, and random forest classifier, (3) accuracy assessment based on visual assessment and confusion matrix, (4) change detection based on spatial and temporal analysis and statistics such as transition matrices, urban growth rate calculation, and district-based, ring-based, and sector-based analysis, (5) simulation of future land use based on the Markov chain and decision forest algorithm, and (6)

evaluation of landscape pattern change using landscape metrics. The ERDAS IMAGINE 2020, SNAP 8.0, QGIS 3, IDRISI TerrSet 2020, FRAGSTATS 4.2, and R 3.6 software, depending on the purpose, were used for these tasks.

1.5. Dissertation outline

This dissertation adopts the integrated dissertation format which is fully article-based. Four scientific papers that have been published in peer-reviewed journals become the backbone of the dissertation as follows.

1. Bui DH, Mucsi L. 2021. From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal Landsat data, a random forest classifier, and decision rules. Remote Sensing. 13(9):1700. https://doi.org/10.3390/rs13091700. *Journal subject: Scopus - Earth and Planetary Sciences (miscellaneous), Rank: Q1*. (Chapter 2)
2. Bui DH, Mucsi L. 2022. Comparison of layer-stacking and Dempster-Shafer theory-based methods using Sentinel-1 and Sentinel-2 data fusion in urban land cover mapping. Geo-spatial Information Science. 25(3):425-438. https://doi.org/10.1080/10095020.2022.2035656. *Journal subjects: Scopus - Computers in Earth Sciences and Geography, Planning and Development, Rank: Q1*. (Chapter 3)
3. Bui DH, Mucsi L. 2022. Land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020. Geocarto International. 37(27):17096–17118. https://doi.org/10.1080/10106049.2022.2123564. *Journal subject: Scopus - Geography, Planning and Development, Rank: Q1*. (Chapter 4)
4. Bui DH, Mucsi L. 2022. Predicting the future land-use change and evaluating the change in landscape pattern in Binh Duong province, Vietnam. Hungarian Geographical Bulletin. 71(4):349-364. https://doi.org/10.15201/hungeobull.71.4.3. *Journal subject: Scopus - Earth and Planetary Sciences (miscellaneous) and Geography, Planning and Development, Rank: Q2*. (Chapter 5)

To this end, the dissertation consists of six chapters as described below.

- **Chapter 1** provides an overview of the dissertation including a brief literature review on the main themes of the research, a description of the study area, the problem statement, the research objective, hypotheses, and a dissertation outline.
- **Chapter 2** first focuses on detecting the connection between land use and land cover in the study area. Then, a novel approach is developed to produce and convert a land cover map to a land use map based on this connection. This method combines pixel-based and object-based methods using multi-temporal Landsat data, random forest classifier, and decision rules. Discussions and comparisons on the effects of data integration (single and multi-temporal satellite images) and methods (pixel-based, object-based, and decision rules) are given.

- **Chapter 3** continues to focus on other aspects of data combination. It investigates the performance of fusing SAR and optical data for land cover mapping. S-1 and S-2 data are used as representations of each data type. The datasets are combined in diverse ways such as single- and multiple-sensor images with and without their extracted indices and textures. The layer-stacking method and D-S theory-based approach are applied to fuse data at the pixel level and decision level, respectively. The effect of data structure and fusion methods is analyzed and discussed.
- **Chapter 4** presents the application of the approach proposed in chapter 2, which is considered the optimal approach given the data availability in the study area and the objective of the study, to generate multi-temporal land use maps from 1995 to 2020 in the study area. Then, the spatial-temporal changes in land use and urban expansion are analyzed and discussed using GIS techniques.
- **Chapter 5** presents the simulation of the future land use up to 2030. A Markov chain and decision forest algorithm are used to determine the driving variables and to predict the quantity and location of future land use allocation. In addition, this chapter also provides an evaluation of the change in the landscape pattern of the study area. Landscape metrics at landscape and class levels are calculated and analyzed.
- **Chapter 6** summarizes the key findings, implications, limitations, and recommendations of the dissertation.

The style and format may vary and overlap between chapters in other to meet the specific requirements of the submitted journals. Generally, in chapters 2 to 5, the structure of each chapter is included an independent abstract, introduction, materials and methods, results and discussion, and conclusion. Because all the papers are closely related to the overall research objective and hypotheses, the dissertation cannot avoid repeating some contents through different chapters, such as the literature review, description of the study area, data, and methodology.

2. From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal Landsat data, a random forest classifier, and decision rules

This article is published in *Remote Sensing* as:

Bui DH, Mucsi L. 2021. From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal Landsat data, a random forest classifier, and decision rules. Remote Sensing. 13(9):1700. https://doi.org/10.3390/rs13091700

remote sensing

MDPI

Article

From Land Cover Map to Land Use Map: A Combined Pixel-Based and Object-Based Approach Using Multi-Temporal Landsat Data, a Random Forest Classifier, and Decision Rules

Dang Hung Bui [1,2,*] and László Mucsi [1]

[1] Department of Geoinformatics, Physical and Environmental Geography, University of Szeged, Egyetem utca 2, 6722 Szeged, Hungary; mucsi@geo.u-szeged.hu
[2] Institute for Environmental Science, Engineering and Management, Industrial University of Ho Chi Minh City, No. 12 Nguyen Van Bao Street, Go Vap District, Ho Chi Minh City 700000, Vietnam
* Correspondence: hungbui@geo.u-szeged.hu

Abstract: It is essential to produce land cover maps and land use maps separately for different purposes. This study was conducted to generate such maps in Binh Duong province, Vietnam, using a novel combination of pixel-based and object-based classification techniques and geographic information system (GIS) analysis on multi-temporal Landsat images. Firstly, the connection between land cover and land use was identified; thereafter, the land cover map and land use function regions were extracted with a random forest classifier. Finally, a land use map was generated by combining the land cover map and the land use function regions in a set of decision rules. The results showed that land cover and land use were linked by spectral, spatial, and temporal characteristics, and this helped effectively convert the land cover map into a land use map. The final land cover map attained an overall accuracy (OA) = 93.86%, with producer's accuracy (PA) and user's accuracy (UA) of its classes ranging from 73.91% to 100%. Meanwhile, the final land use map achieved OA = 93.45%, and the UA and PA ranged from 84% to 100%. The study demonstrated that it is possible to create high-accuracy maps based entirely on free multi-temporal satellite imagery that promote the reproducibility and proactivity of the research as well as cost-efficiency and time savings.

Keywords: land cover; land use; multi-temporal; pixel-based; object-based; segmentation; image classification; random forest; decision rules; Landsat-8

Citation: Bui, D.H.; Mucsi, L. From Land Cover Map to Land Use Map: A Combined Pixel-Based and Object-Based Approach Using Multi-Temporal Landsat Data, a Random Forest Classifier, and Decision Rules. *Remote Sens.* **2021**, *13*, 1700. https://doi.org/10.3390/rs13091700

Abstract

It is essential to produce land cover maps and land use maps separately for different purposes. This study was conducted to generate such maps in Binh Duong province, Vietnam, using a novel combination of pixel-based and object-based classification techniques and geographic information system (GIS) analysis on multi-temporal Landsat images. Firstly, the connection between land cover and land use was identified; thereafter, the land cover map and land use function regions were extracted with a random forest classifier. Finally, a land use map was generated by combining the land cover map and the land use function regions in a set of decision rules. The results showed that land cover and land use were linked by spectral, spatial, and temporal characteristics, and this helped effectively convert the land cover map into a land use map. The final land cover map attained an overall accuracy (OA) = 93.86%, with producer's accuracy (PA) and user's accuracy (UA) of its classes ranging from 73.91% to 100%. Meanwhile, the final land use map achieved OA = 93.45%, and the UA and PA ranged from 84% to 100%. The study demonstrated that it is possible to create high-accuracy maps based entirely on free multi-temporal satellite imagery that promote the reproducibility and proactivity of the research as well as cost-efficiency and time savings.

Keywords: land cover; land use; multi-temporal; pixel-based; object-based; segmentation; image classification; random forest; decision rules; Landsat-8

2.1. Introduction

Land cover is defined as "the observed (bio)physical cover on the earth's surface" (Di Gregorio 2005), including vegetation, water surface, bare rock, bare soil, buildings, and roads. Meanwhile, land use refers to "the arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it" (Di Gregorio 2005); in other words, land use is the way in which people use land cover types for one or more different purposes. Although they are defined differently and this issue has been discussed in previous studies (Cihlar and Jansen 2001; Brown and Duh 2004; Kim 2015), these two terms are still commonly used concurrently or interchangeably in many studies related to land cover and land use classification and mapping (Steinhausen et al. 2018; Carranza-García et al. 2019; L.H. Nguyen et al. 2020). This problem may cause ambiguity or confusion for readers or map users (Comber et al. 2008), as well as certain difficulties in using such maps, because land use information is often used for planning (Tapiador and Casanova 2003) and making policy (van Delden et al. 2011), while land cover information is often employed in environmental monitoring (Henits et al. 2017), modeling (Shooshtari and Gholamalifard 2015), and prediction (Rizeei et al. 2016).

Obviously, it is easier to observe and classify land cover types directly from aerial or satellite images than to do so with land use types (Zhang and Wang 2003). However, there is a strong connection between land cover and land use (Jansen and Di Gregorio

2003; Kim 2015), and once this relationship is clearly defined, land use types can be interpreted from land cover types. There have been attempts to use single or multiple remote sensing data independently (Cihlar and Jansen 2001; Jansen and Di Gregorio 2003; Zhang and Wang 2003) or in conjunction with other ancillary data sources, such as census data (Hunt et al. 2001), land use inventory data (Bauer and Steinnocher 2001), social sensing data (Y. Zhang et al. 2017), and mobile-phone positioning data (Jia et al. 2018), to extract a land cover map and then translate it into a land use map, via a set of parameters and decision rules based on expert knowledge. Such studies have shown potential in producing a land use map from a land cover map based on remote sensing data. However, the reproducibility of their methods is a critical matter of concern. Most of these studies have used either commercial high-resolution satellite images and/or ancillary data, much of which was only available in their study regions. This matter may cause a limitation in repeating the methods of those studies in other study areas. With the availability of completely free remote sensing data sources (e.g., Landsat satellite family or Sentinel satellite family) that cover most of the terrestrial area of the Earth's surface and are easy to access and download, it is necessary to have the research rely entirely on them to extract such maps. It not only enhances the reproducibility of the research methods but also their proactivity as well as cost-efficiency and time savings.

Vietnam is a developing country in Southeast Asia that has seen rapid urbanization and industrialization in recent years. According to the General Statistics Office of Vietnam (General Statistics Office of Vietnam 2020), the population in Vietnam's urban areas increased by an average of nearly 800,000 people each year in the 1999–2019 period. The rate of urbanization in the country, which was calculated as a percentage of the urban population per total population, reached 35.05% in 2019. The rate of urbanization is higher than the country average for megacities, such as Hanoi and Ho Chi Minh City, and also for provinces located in key economic zones, such as Binh Duong province. Specifically, in Binh Duong province, which is the area selected for this study, the rate of urbanization has increased rapidly since the 2000s up to now. In a period of ten years from 1995 to 2005, the urbanization rate only rose from 17.51% to 30.09%; however, in the following ten years (from 2005 to 2015), this rate grew rapidly from 30.09% to 76.72%. As of 2019, the rate reached 79.87%. As a result, in the process of urbanization and industrialization, the expansion of existing developed areas and the formation of new urban areas as well as new industrial and commercial regions have resulted in a rapid transformation of land use and land cover (Ha et al. 2020). Therefore, it is necessary to use land cover and land use maps in near real time for planning and management activities.

In Vietnam, land use status maps are produced by the government at the local and national levels every five years. The basis for producing such maps consists of inventory data related to land changes, including land allocation, land lease, and change of land use purpose during the five-year inventory period (Minister of Natural Resources

and Environment of Vietnam 2018). The land use categories for the maps are up to five levels and 56 classes, which are very detailed and complex. The first version of this type of map was generated in 2005 (since the Act on Land 2003 was passed and implemented); thus, there is a lack of spatial data on land use for previous years. In addition, the five-year interval to produce the maps is not suitable for real-time management and research. Furthermore, the maps are extracted on computer-aided design (CAD) file formats, which make it difficult and time-consuming to convert and use them for geographic information system (GIS) analysis. Last but not least, it is difficult to access this data due to the government's relatively complicated administrative system.

Meanwhile, a land cover map has not yet been produced by the government of Vietnam. Only a few studies have been done independently by researchers or organizations at the national (ADPC 2020; JAXA EORC 2020) or local level (Linh et al. 2012; Disperati and Virdis 2015; H.T.T. Nguyen et al. 2020). However, depending on the purpose of each study, there are many dissimilarities in the land cover classification scheme as well as in the definitions of the classes in those schemes. In addition, a number of these studies also used the terms "land cover" and "land use" concurrently or interchangeably, which may cause a number of obstacles in using these maps, as noted. It should be highlighted that except for the SERVIR-Mekong (ADPC 2020) and Japan Aerospace Exploration Agency (JAXA) (JAXA EORC 2020) programs for land use and land cover mapping for the whole of Vietnam, there are no other studies related to land cover classification in the selected study area.

With these issues in mind, this study was carried out in an effort to extract the land cover map and land use map of Binh Duong province separately. To achieve this objective, we proposed a novel combination of pixel-based and object-based classifications using random forest, decision rules, and free multi-temporal remote sensing data, specifically multi-temporal Landsat-8 imagery. The specific objectives of this study included:

- To identify the relationship between land cover and land use in Binh Duong province based on multi-temporal satellite images and field surveys.
- To test and assess the performance of the combination of pixel-based and object-based classification techniques and GIS analysis on multi-temporal Landsat images to generate a land cover map and a land use map separately.

2.2. Study area

Binh Duong province is located in southeast Vietnam, between 10°51′46″ and 11°30′ N latitude and between 106°20′ and 106°58′ E longitude (Figure 2.1). The total area of the province is over 2694.64 km^2, and its population is about 2.5 million people, of whom 79.87% live in urban areas as of 2019 (Binh Duong Statistical Office 2020).

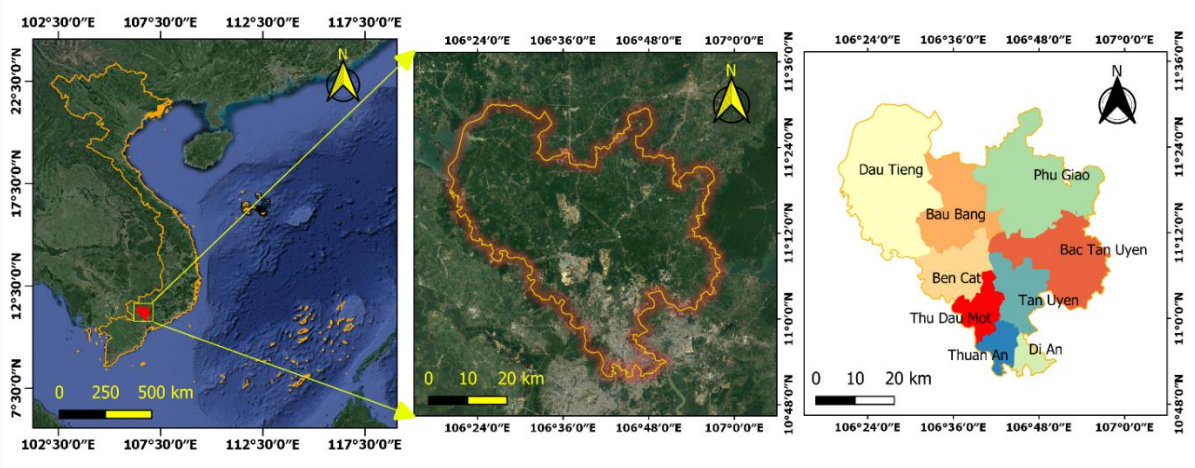


Figure 2.1. The study area.

Binh Duong's climate is divided into two separate seasons, which are characterized by the tropical monsoon and sub-equatorial climates, with a rainy season from May to November, and a dry season from December to April of the following year. The total annual precipitation is about 2275 mm, in which the rainy season accounts for over 90%. The province has a stable geology, relatively flat terrain featuring ancient alluvial hills with an average height of 20–25 m and a slope of 3–15°. There are two large rivers (the Dong Nai and the Sai Gon) with many canals and other small streams.

2.3. Materials and methods

The overall workflow is illustrated in Figure 2.2 and described in detail in the subsections.

2.3.1. The main land cover and land use classes in the study area

Based on the results of the field survey trip combined with observation of Landsat images and Google Earth history images, the main land cover types in Binh Duong province are defined below.

- Barren land: Totally bare soil areas without any cover or with very sparse vegetation or bare land areas partly covered with sunburned vegetation and/or very sparse fresh vegetation.
- Impervious surface with high albedo: Factories and commercial buildings whose material is often light-colored corrugated iron or concrete, or stone mining sites.
- Impervious surface with low albedo: Residences, small commercial and office buildings, roads whose material is often concrete, clay, corrugated iron, asphalt, or a mix of these materials, or stone mining sites.
- Grass: Fresh grass on cultivated grass farms, golf courses, and green spaces.
- Crops: Crops on farms and green spaces or plant nurseries with high density.
- Mature woody trees: Industrial trees, fruit trees, forests, and trees in green spaces which are of mature age with high coverage density.

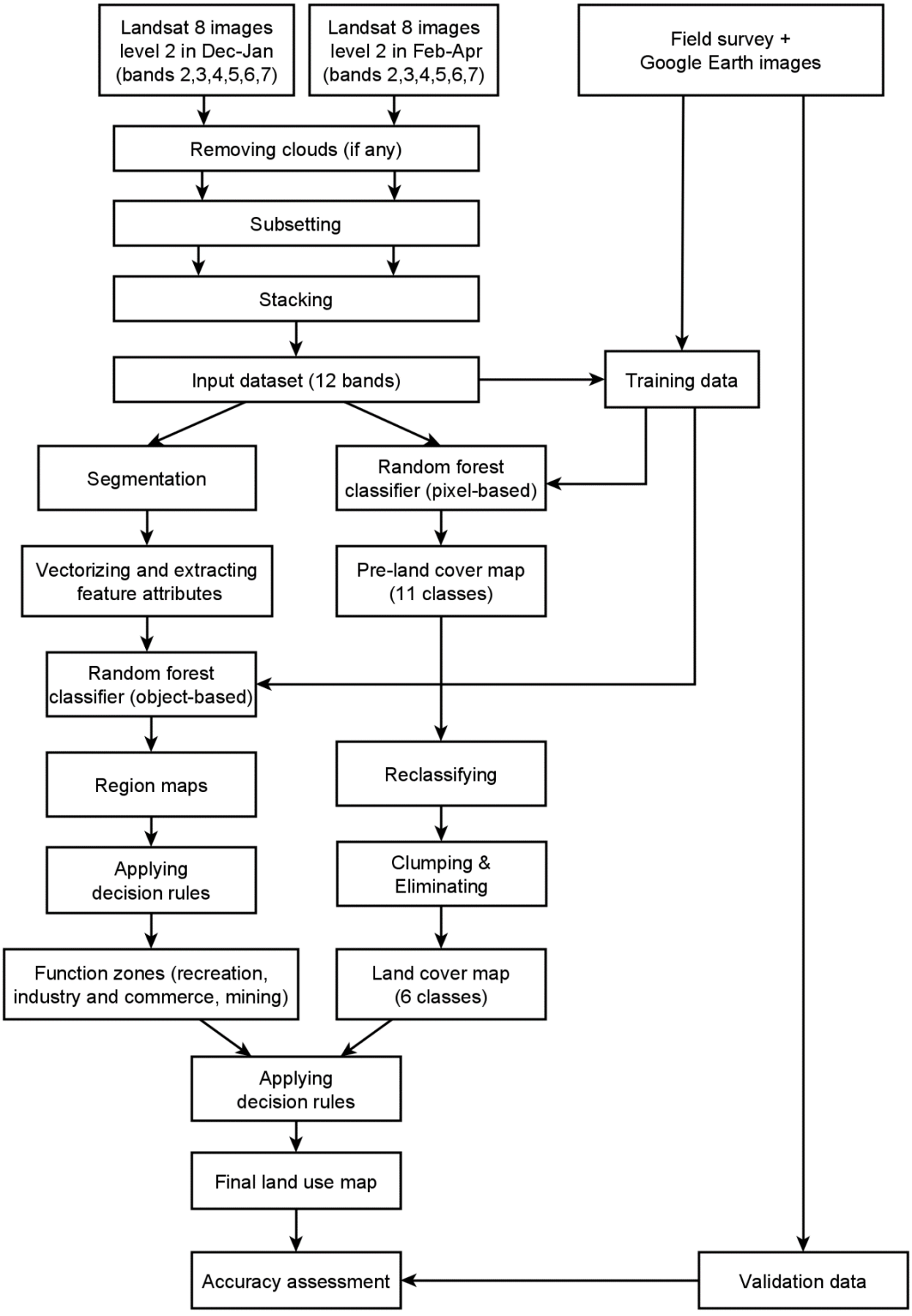


Figure 2.2. The overall workflow.

- Young woody trees: Industrial trees, fruit trees, forests, and trees in green spaces which are young in age with low coverage density because their canopies/crowns are still separate.
- Water: Rivers, canals, lakes, ponds, and pools.

Meanwhile, determining the main land use classes in the study area was relatively difficult. Although the Vietnamese government has issued a current land use classification system (Minister of Natural Resources and Environment of Vietnam 2018), there were many land use categories with similar or very ambiguous definitions (e.g., defense land versus security land, land for religious facilities versus land for belief facilities, etc.), which caused difficulties in finding the relationship between land cover and land use. Additionally, many categories did not exist in the study area or occupied only a very small part and were thus unrepresentative of the study area. Therefore, this study only used this land use classification system combined with the European Union's Coordination of Information on the Environment (CORINE) land cover system (Kosztra et al. 2017) and the Food and Agriculture Organisation's Land Cover Classification System (LCCS) (Di Gregorio 2005) as references. The final land use categories were defined and modified based on the actual characteristics of the study area. Thus, land use classes in this study were defined as follows.

- Unused land: Areas where there are temporarily no construction works or which are leveled, or agricultural land in the harvest stage or in an early stage of the cultivation season with very young trees.
- Industry and commerce: Factories, buildings, a road network, and other built-up areas for production activities and/or commerce and services.
- Recreation and green space: Areas for relaxation and recreation activities, or areas for landscaping or creating a microclimate.
- Mixed residence: Houses, apartments, a road network, and other built-up areas for living and daily life activities. It may also include some entertainment buildings intermingled within residential areas.
- Mining sites: Areas for mining, exploitation, processing, and storing construction stone.
- Agriculture with annual plants: Agricultural land used for growing plants with the growth period from planting to harvesting not exceeding one year, such as rice, maize, vegetables, cultivated grass, etc., or plant nurseries.
- Agriculture with perennial plants: Agricultural land used for growing plants with the growth period from planting to harvesting over one year, such as fruit trees, industrial trees, and forests.
- Water surface: Water body surface.

2.3.2. Collecting and pre-processing satellite images

Landsat-8 Operational Land Imager (OLI) level 2 surface reflectance images for Binh Duong province (projection: WGS 84/UTM Zone 48N, path/row: 125/52) were ordered and downloaded from the United States Geological Survey (USGS) website (USGS 2020). The images (bands 2, 3, 4, 5, 6, 7) were collected at two consecutive periods:

- Period T1: from the end of November 2019 to the end of January 2020, corresponding to the period from the late rainy season to the early dry season.
- Period T2: from the beginning of February to the end of April 2020, corresponding to the period from the middle to the end of the dry season.

There were two reasons for selecting Landsat images at these two periods. Firstly, most of the study area was covered by dense clouds during the rainy season, i.e., from May to November; therefore, it was almost impossible to choose or to mosaic an image that was completely free of clouds in this period. Secondly, due to human activities, seasonal hydrological activities, and the characteristics of each land cover type, there was a transformation of land cover at some locations between the T1 and T2 periods. These changes might be used to improve the land cover classification and to interpret land use types. This will be discussed in detail in Section 2.4.1.

As a result, at period T1, a completely cloud-free image acquired on 6 January 2020 was chosen. At period T2, images acquired on 23 February 2020 were selected. However, there was a small region in the study area covered by clouds on 23 February. Therefore, an image acquired on 7 February 2020, which was completely cloud-free in that small region, was additionally selected. Thus, the cloud region in the 23 February image was masked and replaced with the corresponding cloudless region in the 7 February image.

Then, selected images were subsetted to the study area and stacked together to create an input dataset with twelve bands (i.e., six bands in each period), which was ready for the next processing steps. The administrative boundary of the study area used for subsetting was downloaded from the Database of Global Administrative Areas (GADM) project website (GADM 2020).

2.3.3. Collecting training and validation data

A field survey trip to the study area was conducted from 18 January to 18 February 2020. From the results of the trip combined with the observations on the selected Landsat images at the T1 and T2 periods and the Google Earth History photos, a set of training and validation data was collected for classification and an accuracy assessment, respectively.

For the pixel-based classification, the two-stage sampling technique (De Gruijter et al. 2006) was used. A total of 390 polygons were collected in homogeneous areas and assigned to the respective land cover classes. Subsequently, 70% of the polygons in each

class were selected randomly, and all the pixels within them were extracted to create the training dataset (4909 pixels); in the remaining number of polygons, 30% of the pixels were extracted periodically to create the validation data (586 pixels). It should be highlighted that because mining sites often had heterogeneous surfaces, consisting of interspersed high-albedo and low-albedo impervious surfaces, it was difficult to obtain a homogeneous region. Therefore, there were no polygons or points collected in these areas in this step.

For training in object-based classification, based on the segmented result, there were 399 and 349 segments selected for training in the first and second rounds, respectively. It should be noted that since an extended rectangular boundary was used in the segmentation step (see Section 2.3.5), several segments that were outside the administrative boundary of the study area were selected to achieve the best result. Finally, to evaluate the accuracy of the final land use map, 586 points in the land cover validation dataset were used again, which were carefully assigned to the respective land use classes. In addition, 25 other points collected randomly at the mining sites were added. As a result, a total of 611 points were used at this stage.

The distribution of training and validation data depended on the proportion and spatial distribution of each class in the study area (Figure 2.3). The detailed number of polygons and points for each class at each stage is summarized in Appendix A.

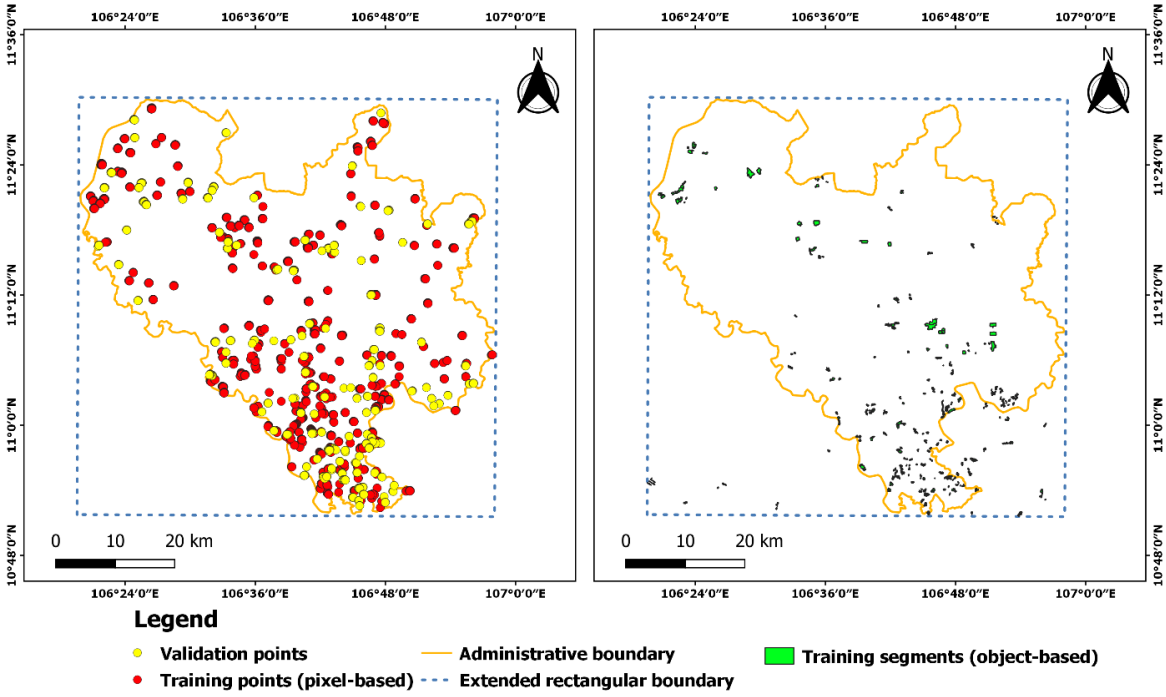


Figure 2.3. The spatial distribution of training and validation data.

2.3.4. Pixel-based classification

The pixel-based approach was used in this study to produce the land cover map by applying random forest classifier and some post-classification techniques. The pixel-

based approach was used due to its ability to generate a more detailed land cover map than one based on the object-based approach.

To pave the way for conversion from a land cover map to a land use map in the next steps, based on the fact that land cover at a location might be unchanged or changed from one type to another type between the two selected periods (see Section 2.4.1), an additional classification scheme with twelve classes was determined as in Table 2.1. It should be noted that despite the fact that there was a change from water to barren land at the coastal region of some water surfaces (e.g., reservoirs and lakes) between T1 and T2, that change only took place in a very small area. This issue made it difficult to obtain training and validation samples. Therefore, that type of change was ignored, and the classification process used only eleven classes.

Table 2.1. Pre-land cover and land cover classification scheme.

No.	Land Cover Class at T1	Land Cover Class at T2	Pre-Land Cover Class	Land Cover Class	Note
I. Unchanged classes					
1	Barren land	Barren land	(1) Barren land	(1) Barren land	
2	Crops	Crops	(2) Crops	(2) Annual plants	
3	Grass	Grass	(3) Grass	(3) Grass	
4	Young woody trees	Young woody trees	(4) Young woody trees	(4) Perennial plants	
5	Mature woody trees	Mature woody trees	(5) Mature woody trees	(4) Perennial plants	
6	IS with high albedo	IS with high albedo	(6) IS with high albedo	(5) Impervious surface	
7	IS with low albedo	IS with low albedo	(7) IS with low albedo	(5) Impervious surface	
8	Water	Water	(8) Water	(6) Water	
II. Changed classes					
9	Barren land	Crops/grass	(9) Barren land to crops/grass	(2) Annual plants	
10	Crops/grass	Barren land	(10) Crops/grass to barren land	(2) Annual plants	
11	Mature woody trees	Barren land	(11) Mature woody trees to barren land	(1) Barren land	
12	Water	Barren land	(12) Water to barren land	(6) Water	Ignored

Note: IS = Impervious surface.

A classified map (pre-land cover map) was produced using this classification scheme with the random forest algorithm (Breiman 2001), whose effectiveness in processing high-dimensional data has been proven to be fast and insensitive to overfitting (Belgiu and Drăguţ 2016). The “randomForest” package (Liaw and Wiener 2002) was used on R software (version 3.6.3) for the classification procedure. In this process, the number of variables randomly sampled as candidates at each split (mtry) was set at the default value, which was equal to the square root of the total number of

features. Meanwhile, the maximum number of trees (ntree) was decided based on the plot showing the relationship between ntree and the decrease of out-of-bag (OOB) error rates (Figure 2.4). As a result, ntree was set at 750 trees.

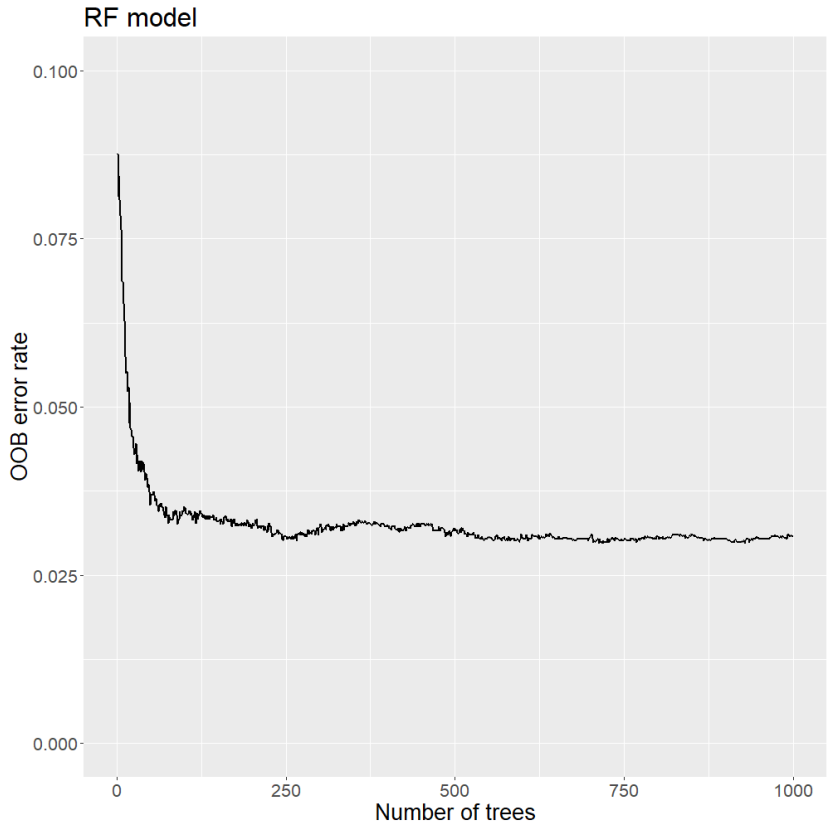


Figure 2.4. The relationship between out-of-bag (OOB) error rate and number of trees (ntree) in the random forest (RF) model for extracting the pre-land cover map.

Then, the pre-land cover map was re-classified from eleven to six classes to create the land cover map. The conversion of classes is also shown in Table 2.1. In addition, to remove "salt and pepper" noise on the land cover map, the Clump and Eliminate functions in ERDAS IMAGINE 2020 software were used, respectively. Clumps (i.e., contiguous groups of pixels in one thematic class) smaller than four pixels were eliminated and given the value of nearby larger clumps. The final land cover map with the six classes was produced following this process.

In addition, a similar classification procedure was also performed on single images at T1 and T2 independently to compare the results with the classification output from the multi-temporal image.

2.3.5. Object-based classification

The purpose of this step was to produce land use function regions (i.e., industrial and commercial regions, mining regions, and recreation regions), which would be used to combine with the final land cover map to produce the final land use map. The object-based classification approach was used to create these regions. The object-based approach is capable of overcoming the spectral and spatial limitations of single pixels.

An extended rectangular boundary (Figure 2.3) was used at this stage instead of the administrative boundary to ensure landscape continuity in areas surrounding the administrative boundary for more accurate segmentation. Firstly, the input raster dataset (12 bands) was segmented using the Full Lambda Schedule (FLS) Image Segmentation function in ERDAS software. In this study, after various experiments and visual analysis, the parameters were set as follows:

- Pixel:Segment Ratio: 50
- Relative Weights of Spectral: 0.5
- Relative Weights of Texture: 0.5
- Relative Weights of Size: 0.5
- Relative Weights of Shape: 0.5
- Size limits: minimum: 10; maximum: 1000

Subsequently, the segmentation image was converted to a polygon shapefile. From this shapefile and the input raster dataset, the following statistical and texture attributes of each raster band were extracted for each segment (i.e., each polygon) using a series of functions per feature in ERDAS (Table 2.2). As a result, 156 attributes of each segment were extracted.

Table 2.2. Extracted attributes per feature.

ERDAS Function	Extracted Attributes
Raster statistics per feature	Mean, Max, Min, Median, Standard Deviation
Kurtosis texture per feature	Mean, Standard Deviation
Variance texture per feature	Mean, Standard Deviation
Skew texture per feature	Mean, Standard Deviation
Mean Euclidean Distance texture per feature	Mean, Standard Deviation

Finally, land use function regions were created using the random forest classifier on the extracted feature shapefile and then using spatial analysis techniques. The process is shown in Figure 2.5. With the random forest algorithm, the ntree parameter was set to 650 trees for the first round and 600 trees for the second round (Figure 2.6); meanwhile, the mtry parameter was also set to the default value. The spatial analysis was conducted in QGIS software (version 3.4.15) and the criteria were chosen based on a trial-and-error process. There were two rounds of classification:

- The first round: to extract golf courses (recreation areas) and mining regions. The classification scheme included four potential classes: mining, recreation, industry and commerce, and other. The industrial and commercial region class was classified in this step in an effort to facilitate a better classification of the mining region.
- The second round: to extract industrial and commercial regions after subtracting recreation areas and mining areas in the first round. The classification scheme included two potential classes, industry and commerce, and other.

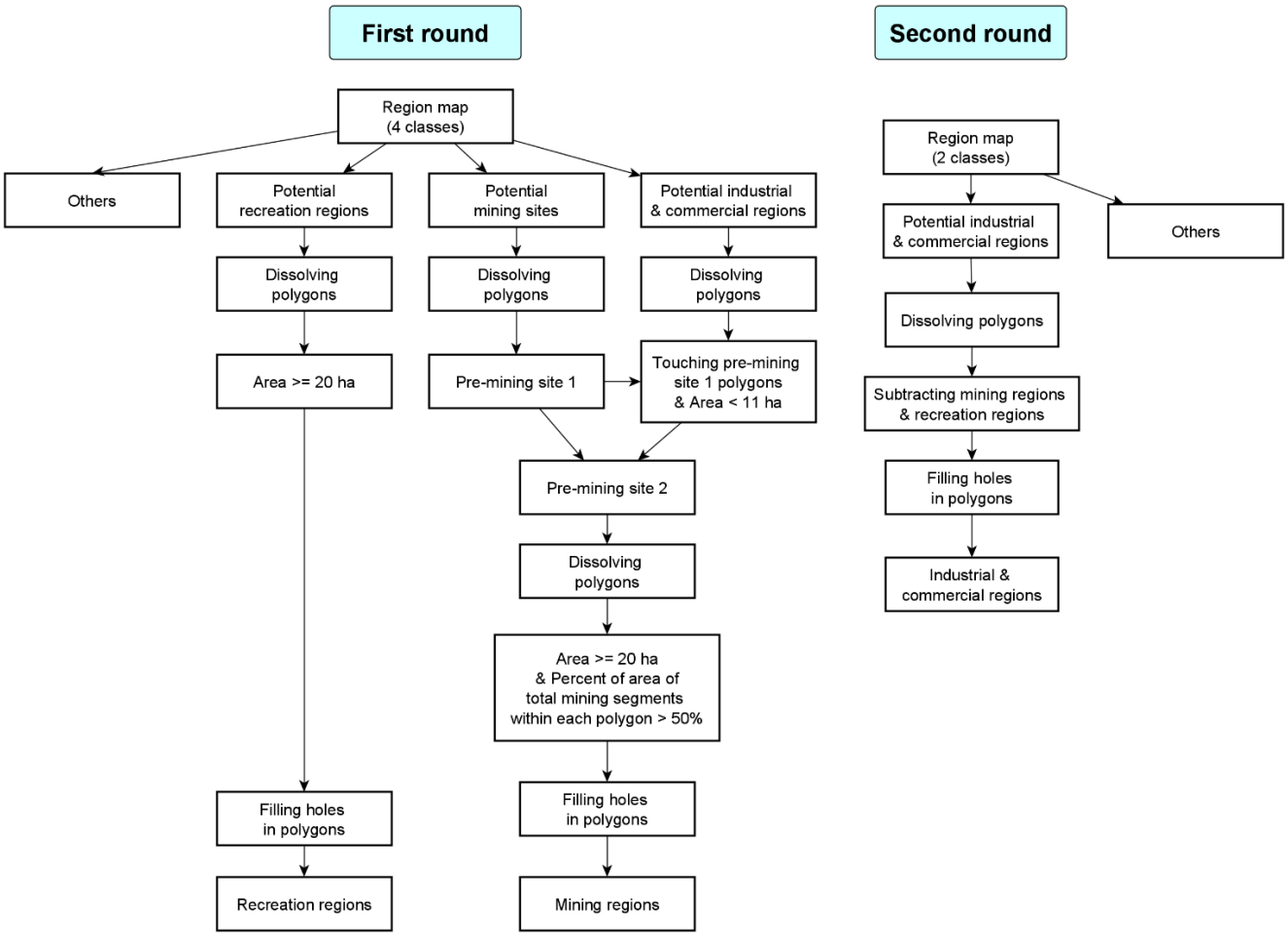


Figure 2.5. Process for extracting land use function regions.

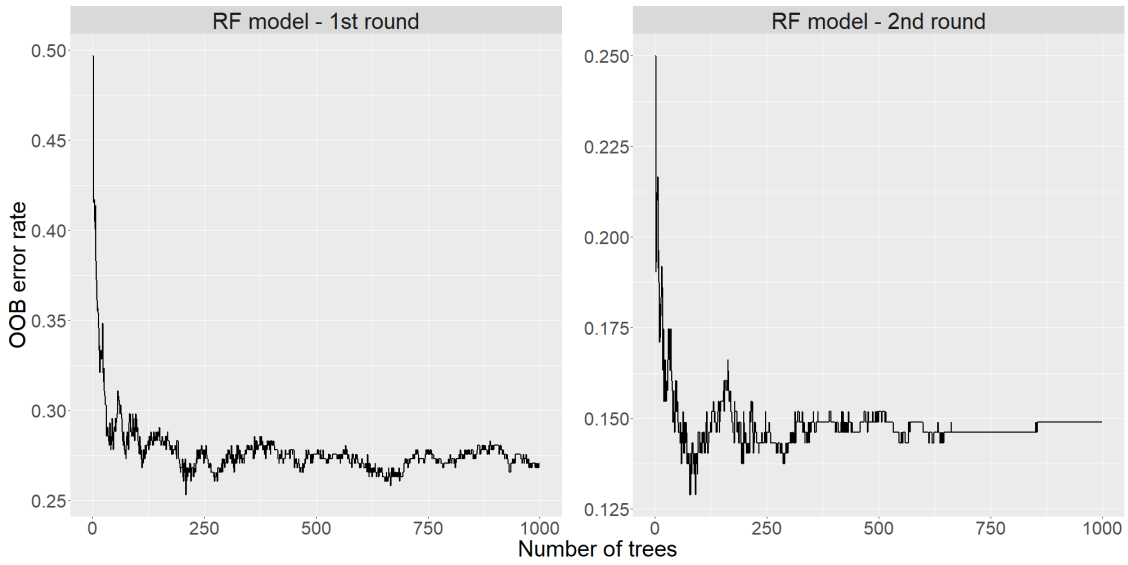


Figure 2.6. The relationship between OOB error rate and ntree in the RF models for extracting land use function regions.

2.3.6. Producing the land use map

The final land use map was created by combining the final land cover map and the land use function regions in a set of decision rules. The overall decision workflow is demonstrated in Figure 2.7.

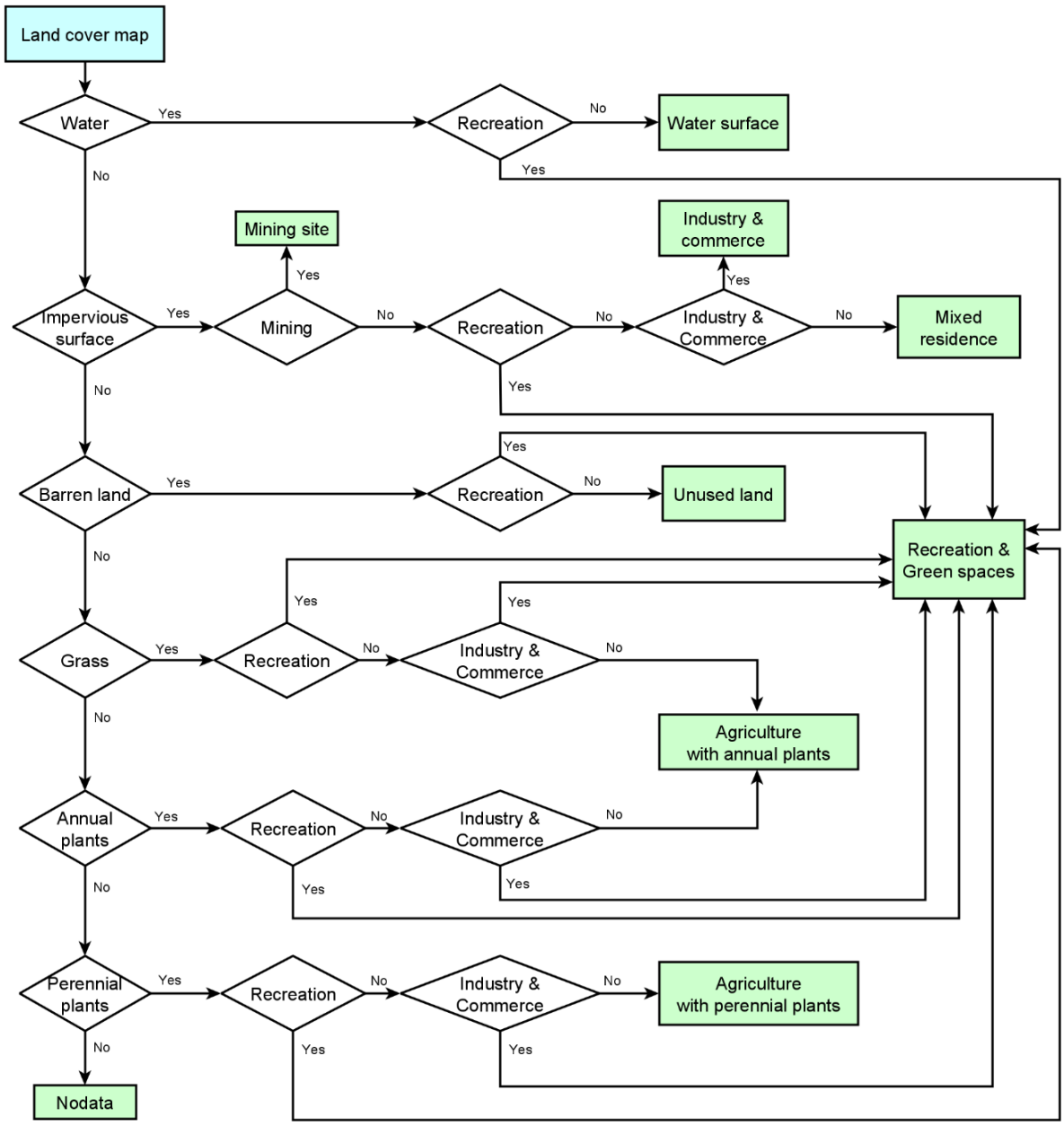


Figure 2.7. Decision rules for producing the land use map.

2.3.7. Accuracy assessment

The accuracy of the land cover maps, function regions, and final land use map produced at each processing step was assessed by both visual assessment and a confusion matrix with overall accuracy (OA), producer's accuracy (PA), and user's accuracy (UA) calculated. In addition, with the aim of assessing and comparing the accuracy of the generated maps in more detail, this study also computed two other parameters, quantity disagreement (QD) and allocation disagreement (AD), instead of the Kappa coefficient, which has been proven by Pontius and Millones to be redundant or misleading (Pontius and Millones 2011). Their study detailed the concept and method for calculating these two parameters.

2.4. Results

2.4.1. The link between land cover and land use types

After defining the main land cover and main land use classes in the study area, the link between land cover and land use was interpreted by analyzing the spectral, spatial, and temporal characteristics of each land cover and land use type, as well as human and seasonal hydrological activities. The connection helped form a set of decision rules (see Section 2.3.6) to efficiently convert the land cover map to a land use map. These connections are detailed below and illustrated in Figure 2.8.

- Barren land was often an area that was temporarily unused for any purpose. Its spectral signature overlapped partially with that of the impervious surfaces. However, the spectral signature of barren land regions fluctuated depending on the amount of vegetation scattered in the region. The less vegetation there was, the stronger the surface reflectance. Furthermore, the density and freshness of vegetation in barren land regions might slightly change with the seasons.
- Built-up areas and mining sites were all characterized by the domination of the impervious surface. The high density of the high-albedo impervious surface was usually characteristic of the industrial and commercial regions. These areas often consisted of large, corrugated iron buildings (usually larger than 1000 m^2/building) of a variety of colors, with each color marked by a different spectral signature. Thus, the spectral reflectance fluctuation in these regions was relatively wide in all bands. Meanwhile, low-albedo impervious surface often fell within mixed residential areas, including private houses, blocks of flats, transportation networks, or small commercial buildings and office blocks with the building size often less than 500 m^2. They were made up of different materials, such as corrugated iron, concrete, asphalt, brick, and clay tile. Hence, the value of each pixel in the Landsat image was a mixture of surface reflectance from these materials, and the spatial spectral variance was not too wide. With stone-mining sites, they included impervious surface (both high and low albedo) and water, and the area of existing quarries was often larger than 20 ha. Furthermore, in general, in terms of temporal change, impervious surface was almost unchanged in a short time.
- Regions relevant to the domination of vegetation included recreation and green space regions and agricultural regions. Golf courses were dominated by a large fresh grass area. Meanwhile, green spaces could include woody or herbaceous plants with a small area and located in developed regions. For agricultural regions, in agricultural and forestry activities in Vietnam, woody trees were considered as perennial plants, and crops/cultivated grass were considered as annual plants. Comparing spectral signatures between grass, crops, and mature woody trees, it can be seen that although their curve shapes were relatively similar, fresh grass had the highest spectral reflectance values in most bands, particularly in the NIR band.

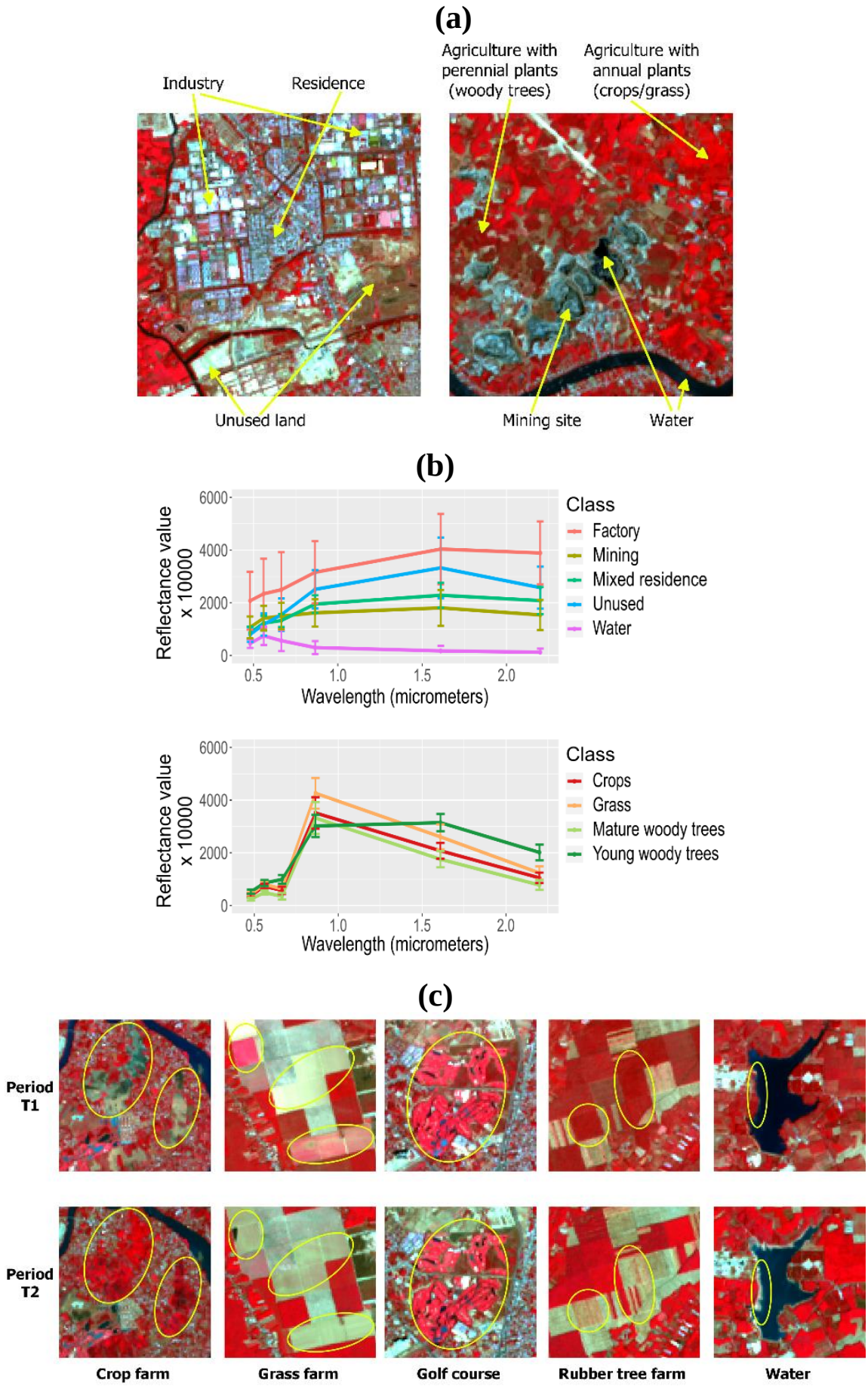


Figure 2.8. The characteristics of and connection between land cover and land use. (a) Spatial and visual characteristics; (b) Spectral characteristics; (c) Temporal characteristics.

Mature woody tree regions seemed to have the lowest values in all bands. With young woody trees, because their canopies have not intersected yet, it led to low coverage density, and the spectral reflectance of such regions was a mixture of the plants and the ground. This resulted in the spectral signature of this class being quite different from the other three vegetative classes. In addition, due to seasonal agricultural activities, cultivated grass/crops on farms might be changed to barren land and vice versa; meanwhile, grass on golf courses was unchanged.

- Mature woody tree regions could be changed to barren land after clear-cutting. This usually occurred on farms during the timber harvest (acacia, dipterocarps, etc.) or clear-cut poorly productive old trees for the new planting (with rubber, cashew, fruit trees, etc.). This also occurred in the area being leveled for construction activities in the future. The regions after such clear-cutting activities were considered as temporarily unused land.
- Water could be changed to barren land due to seasonal hydrological activities and vice versa. These semi-flooded areas were considered as falling within the water class.

2.4.2. Extracted maps and their accuracy

2.4.2.1. The pre-land cover classification result and the final land cover map

The OA, QD, and AD of the pre-land cover classification result achieved 89.76%, 5.68%, and 5.35%, respectively. The UA and PA of most classes achieved over 80%, in which the highest accuracy took place in the water class, which reached 100% for both PA and UA. In contrast, the crops class attained the poorest accuracy among eleven classes, with 53.33% of UA and 61.54% of PA. Many pixels of this class were misclassified in the mature woody tree class, whereas a number of pixels in the grass class were placed in this category. The misclassification from grass to crops also led to PA in the grass class also being low (65.22%).

A final land cover map was generated from the result of the pre-land cover classification. For this map, except for the grass class, which attained PA = 73.91%, UA and PA achieved over 88% in all the other classes (Table 2.3). Compared to the classification results from single images (Table 2.4), the UA and PA of most classes of the final land cover map were equal to or higher than those of the single-date land cover maps, in which the most significant differences could be observed in the classes of grass and annual plants. The water class also experienced a slight increase in accuracy when using the multi-temporal image. As a result, the OA of the classification result from multi-temporal images, which reached 93.86%, was higher than those of the single-date images, corresponding to OA = 89.59% for the land cover map at T1 and OA = 90.78% for the land cover map at T2. Similarly, the final land cover map also showed a lower disagreement in allocation compared to the land cover maps at T1 and T2, corresponding to AD of 1.51%, 2.16%, and 2.53%, respectively. The disagreement in terms of quantity

for the final land cover map (QD = 4.59%) was also less than that of the land cover map at T2 (QD = 6.48%); however, it was higher than that at T1 (QD = 1.98%). However, in general, the total disagreements for all three maps were low and acceptable.

The final land cover map is illustrated in Figure 2.9.

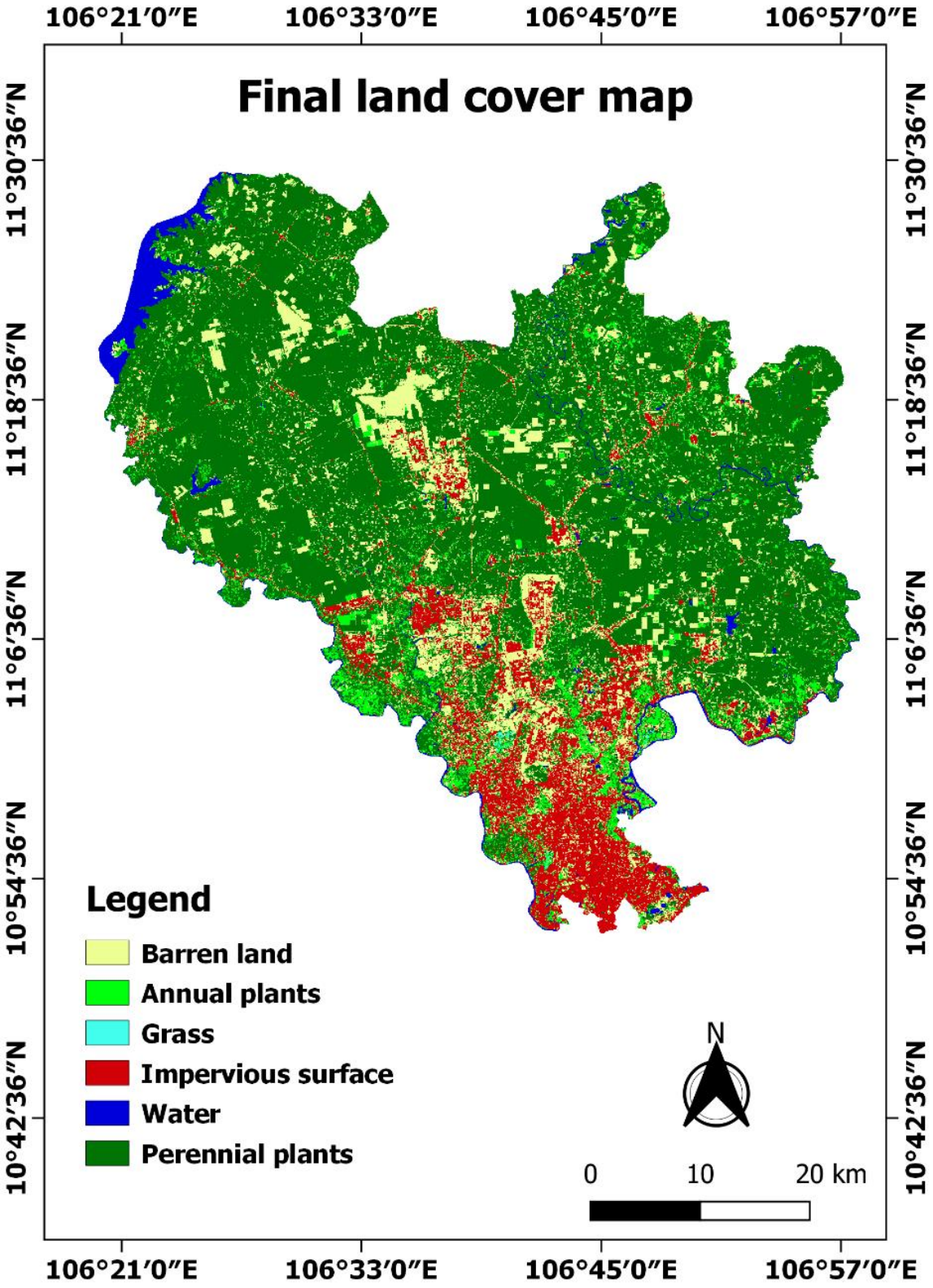


Figure 2.9. Final land cover map.

Table 2.3. Confusion matrix of final land cover map produced from the multi-temporal image.

Class		Classification						Total	PA (%)
		BL	AP	GR	IS	WA	PP		
Referenced	BL	150	0	0	9	0	0	159	94.34
	AP	5	77	0	0	0	5	87	88.51
	GR	0	6	17	0	0	0	23	73.91
	IS	11	0	0	184	0	0	195	94.36
	WA	0	0	0	0	39	0	39	100.00
	PP	0	0	0	0	0	83	83	100.00
Total		166	83	17	193	39	88	586	
UA (%)		90.36	92.77	100.00	95.34	100.00	94.32		
OA = 93.86%; QD = 4.59%; AD = 1.51%									

Note: BL = barren land; AP = annual plants; GR = grass; IS = impervious surface; WA = water; PP = perennial plants; PA = producer's accuracy; UA = user's accuracy; OA = overall accuracy; QD = quantity disagreement; AD = allocation disagreement.

Table 2.4. The accuracy of land cover maps produced from the single-date images.

Class	T1 Image		T2 Image	
	PA (%)	UA (%)	PA (%)	UA (%)
Barren land	92.70	90.16	95.45	87.91
Annual plants	55.77	80.56	75.00	77.14
Grass	69.57	42.11	42.86	88.24
Impervious surface	93.33	93.81	92.82	94.27
Water	87.18	97.14	94.87	100.00
Perennial plants	100.00	99.00	100.00	92.22
	OA = 89.59%; QD = 1.98%; AD = 2.16%		OA = 90.78%; QD = 6.48%; AD = 2.53%	

2.4.2.2. Function regions

Extracted function regions were assessed by visually comparing the study results with Landsat images and Google Earth History images. Some example regions are demonstrated in Figure 2.10. It can be seen that the shapes and sizes of the industrial parks, quarries, and golf courses compared relatively favorably to reality. Due to extracting from the remote sensing image, the boundaries of the image-based regions were not as smooth as they are in reality; however, in general, these image-based regions almost covered the entire land use areas. Although a few constructed buildings around quarries for processing and storing stone were classified under the mining function region, and some small, discrete stone storage yards were not included in this class, these miscategorizations were acceptable to some extent. Overall, it was appropriate to use the generated function regions to produce the land use maps.

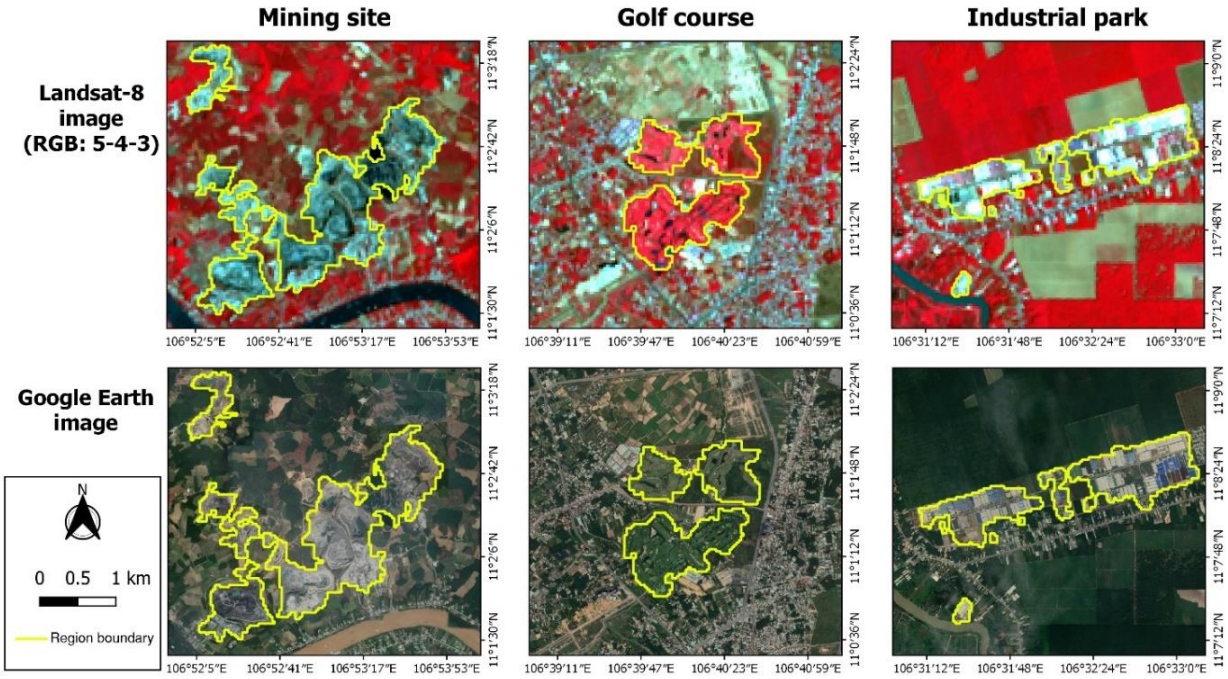


Figure 2.10. Examples of extracted function regions.

2.4.2.3. Land use map

The final land use map is illustrated in Figure 2.11, and its accuracy assessment is demonstrated in Table 2.5. The map achieved 93.45% of OA, 4.58% of QD, and 1.52% of AD, as well as the UA and PA of its classes ranged from 84% to 100%. With visual evaluation, some pixels at the boundary areas of the mining sites were considered as a residence area. This might come from the misclassification between barren land and impervious surface on the land cover map or generated mining function regions with inaccurate boundaries.

Table 2.5. Confusion matrix of the final land use map.

Class		Classification								Total	PA (%)
		UL	IC	RG	MR	MS	AA	AP	WA		
Referenced	UL	150	0	0	9	0	0	0	0	159	94.34
	IC	3	79	0	0	0	0	0	0	82	96.34
	RG	0	0	23	0	0	0	0	0	23	100.0
	MR	8	6	0	99	0	0	0	0	113	87.61
	MS	0	3	0	0	21	0	0	1	25	84.00
	AA	5	0	0	0	0	77	5	0	87	88.51
	AP	0	0	0	0	0	0	83	0	83	100.0
	WA	0	0	0	0	0	0	0	39	39	100.0
Total		166	88	23	108	21	77	88	40	611	
UA (%)		90.36	89.77	100.0	91.67	100.0	100.0	94.32	97.50		
OA = 93.45%; QD = 4.58%, AD = 1.52%											

Note: UL = unused land; IC = industry and commerce; RG = recreation and green space; MR = mixed residence; MS = mining site; AA = agriculture with annual plants; AP = agriculture with perennial plants; WA = water surface.

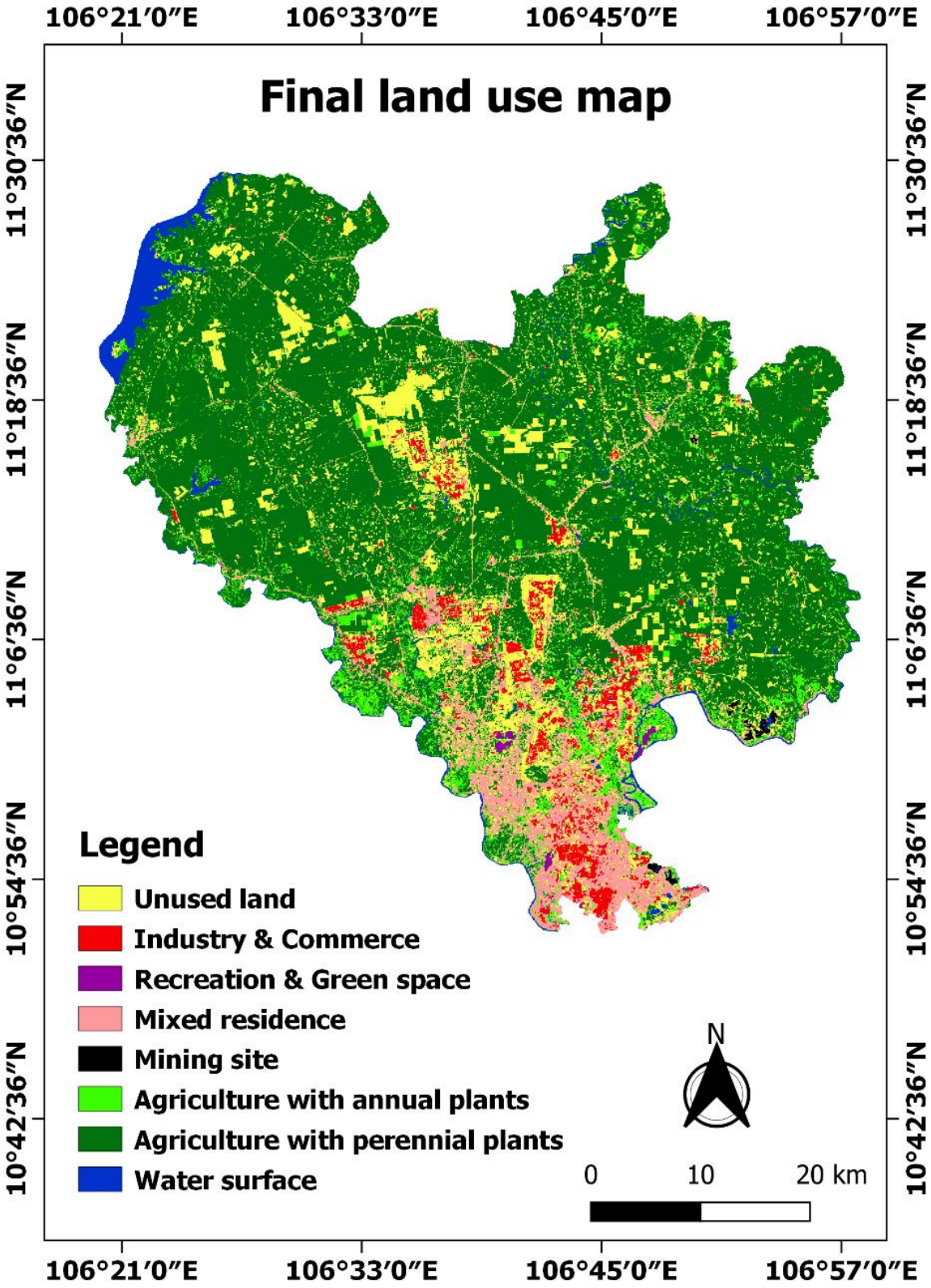


Figure 2.11. Final land use map.

2.5. Discussion

The results of this study showed that land cover and land use in Binh Duong province were not only linked by spatial distribution and spectral properties but also by temporal characteristics. On the one hand, each land use type has its own spatial pattern and structure characterized by the properties of the land cover classes within it, such as

composition, spatial distribution, spectral signature, and dominant class as well as the shape and size of objects. On the other hand, the change or non-change of land cover at a given site over different times of the year may also demonstrate the manner in which humans interact with the land, thereby showing type of land use. In this study, characterizing the properties and finding the relationship between land cover and land use are considered as a vital key to producing a land cover map and a land use map from satellite images. Once these relationships are clearly defined and the suitable classification schemes established, a set of decision rules can be established to demonstrate how to translate a land cover map into a land use map. Then, these maps can be efficiently extracted.

With our proposed procedure, the maps achieved high overall accuracy. It shows the potential of using a combination of pixel-based and object-based classification techniques and GIS techniques on free multi-temporal satellite images to effectively extract and translate a land cover map into a land use map. Many studies have also shown effectiveness in using a combination of these methods in the area of land cover and land use mapping (Shackelford and Davis 2003; Wang et al. 2004; Malinverni et al. 2011; Aguirre-Gutiérrez et al. 2012; Ceccarelli et al. 2013; Chen et al. 2018); however, the difference in this study is the production of a land cover map and land use map separately, which these previous studies have not done.

As the final land use map in our research was generated by combining the land cover map and the land use function regions in a set of decision rules, its accuracy depended on that of the combined components. The advantages to and limitations in how these components are produced are discussed in detail below.

First of all, the method for collecting training and validation data in our study is close to the stratified random sampling method. The distribution of the training and validation data depends on the spatial distribution and the proportion of each class in the study area. For example, in areas where many different types of land cover and land use classes are presented, e.g., in the southern part of the province, sampling density is higher than in other parts (Figure 2.3). In addition, the number of samples of the dominant classes, such as the woody tree or impervious surface classes, is also greater than those of others (Appendix A). Thus, it can be assumed that there is no bias in the research results, thereby ensuring the accuracy and reliability of the generated maps. Furthermore, in the case of a large study area, where more samples need to be collected, it is more appropriate to use the stratified random sampling method, because it ensures that rare classes are not ignored, and it also requires a smaller sample size than the simple random sampling method. This can help save time and effort.

In land cover mapping, in our study, the higher the overall accuracy, the higher accuracy within classes and the low total disagreement of the final land cover map have shown a certain efficiency when using multi-temporal images in a pixel-based

classification compared to using single-date images. These results are consistent with those of Feng et al. (2019), Henits et al. (2016), Yang et al. (2015), and Zoungrana et al. (2015). The pixel-based classification results using a single-date satellite image often encounter common misclassification problems, such as those between impervious surface and barren land (Shao et al. 2016), between dark impervious surface, object-cast shadow, and water (Zhang et al. 2012), between water boundary areas and barren land (Ji et al. 2015), between sparse vegetation and barren land (Shalaby and Tateishi 2007), and between different vegetation classes (Ghosh et al. 2014), due to the spectral similarity between the classes. However, using multi-date images can provide additional useful information to increase classification efficiency (Zoungrana et al. 2015). As noted in Section 2.4.1, there may be a change in land cover and/or its features between two or more time points, which leads to a change in the spectral signature in some areas with no change in others. Therefore, the spectral similarities of these classes may be reduced or removed when using multi-temporal images; thus, it is possible to reduce misclassification. In addition to an improvement in land cover mapping, using multi-temporal images could also facilitate the extraction of land use information, paving the way for land use mapping. For example, by detecting the change/non-change between grass/crops and barren land, it was possible to distinguish agricultural areas from recreation areas and agricultural areas from unused land.

However, taking a more detailed look at the pre-land cover classification result, a high misclassification between some classes, especially between unchanged vegetation classes, still occurred in our research. This can be explained as follows: since only two temporal Landsat-8 images in the dry season are used in this study, they are not able to cover all situations of land cover change at different time points in a year. Meanwhile, depending on the characteristics of crop and farming techniques, the timing of cultivation and harvest on different parcels may not be the same; therefore, there are some parcels covered by crops at both of the selected periods. This problem, combined with the spectral similarity of crops, grass, and woody trees, led to a high misclassification from grass to crops and from crops to mature woody trees. It should be noted that in addition to reducing the accuracy of the land cover map, the misclassification can also affect the extraction of wrong land use information in the steps that follow. Therefore, this limitation should be considered in further studies to improve the accuracy of extracted maps. Adding more remote sensing image data at different times should be considered for the classification process. In addition, other optical image data such as Sentinel-2 may be used in combination to address the problem of some cloud-covered areas in the dry season. In the rainy season, as noted, it is almost impossible to select or to mosaic an optical image that is completely free of clouds in the study area. Hence, it is suggested that a combination with synthetic aperture radar (SAR) images (e.g., Sentinel-1 imagery), which are unaffected by clouds (Gulácsi and Kovács 2020), should be experimented with. We will pursue this in further studies.

It is a fact that a land use type consists of many land cover types, not only one (Cihlar and Jansen 2001). Although industrial and commercial regions, for instance, are dominated by impervious surfaces with high albedo, they may also consist of impervious surfaces with low albedo, such as roads, yards, or small offices, or even grass and trees that are considered as landscaping. Similarly, mining site regions also comprise both high- and low-albedo impervious surfaces as well as water. Thus, if based only on a pixel-based classification—i.e., based solely on spectral characteristics—it is difficult to form boundaries between regions with different land use types. In this study, the segmentation technique and object-based classification have shown the suitability of creating such boundaries and then forming land use function regions. Firstly, segmentation techniques group adjacent pixels into a homogeneous object (i.e., a segment) with clear boundaries by taking into account not only spectral properties but also spatial information, such as texture, shape, and size. Then, the multiple statistical and texture variables are calculated from twelve bands of multi-temporal images based on the value distribution and spatial relationship of all the pixels within each segment; thus, they characterize the spatial distribution of land cover, in other words, the spatial pattern and structure of the land use type, within each segment. This captured information is the basis for classifying segments into land use classes. A comparison of value distribution of some derived attributes between different land use types is illustrated in Figure 2.12. Similar to pixel-based classification, adding more temporal images can help capture more spectral and spatial information, resulting in more efficient segmentation and classification. Despite a potential higher time cost in the processing, it is deemed acceptable.

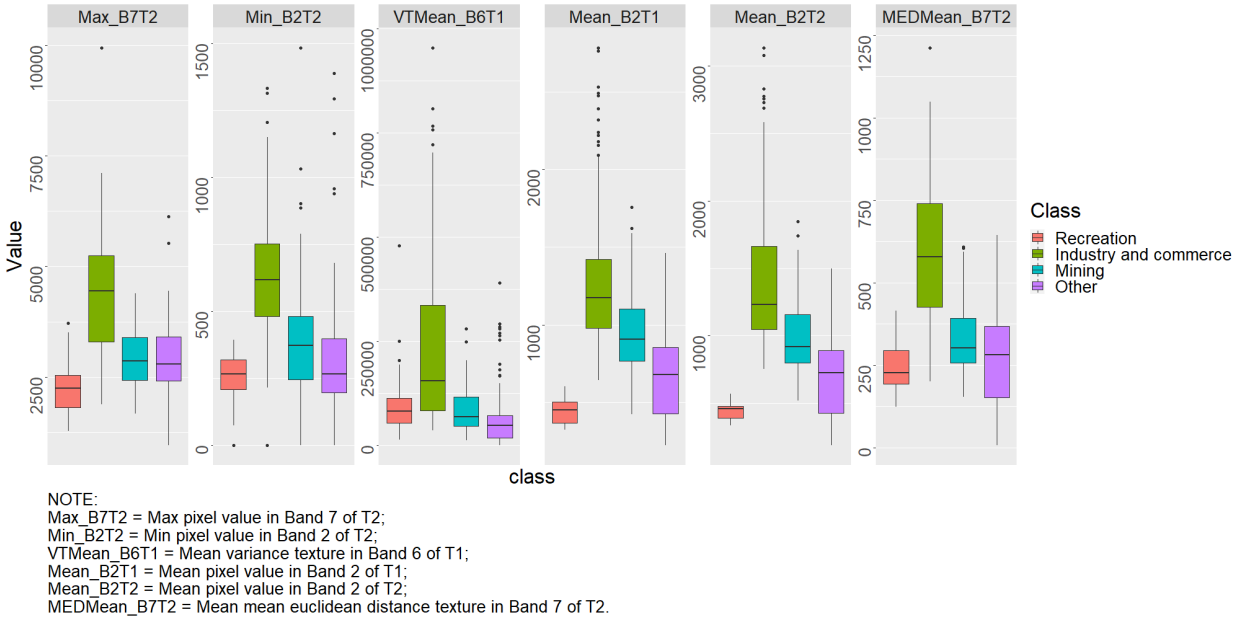


Figure 2.12. Value distribution of some derived attributes of classes.

It should be highlighted that at the segmentation stage, the setting of parameters will affect the performance of segmentation (Ruiz Hernandez and Shi 2018). The errors

of over-segmentation and under-segmentation (Liu and Xia 2010) may lead to incorrect boundaries. However, there are no perfect parameters for this stage (Mucsi et al. 2017). The performance of segmentation results is mostly based on user experience (Ruiz Hernandez and Shi 2018), and the setting also depends on satellite image resolutions and the spatial structure of study areas. Therefore, for similar studies, it is necessary to select and test the value of the parameters carefully to obtain the most accurate land use boundaries.

It is necessary to conduct spatial analysis at the post-classification stage to group adjacent segments with the aim of creating complete land use function regions. This workflow is transferable; however, the criteria used in this process were formed based on personal experience, visual observations, and experiments (trial and error). Although these values are applicable to neighboring areas, such as localities in Southern Vietnam with characteristics similar to those of the study area, it is suggested that these values need to be re-assessed and revised when applying them to other regions where the land cover and land use characteristics differ from those of our study area.

In addition, in this study, using the object-based approach just to form the land use function zones, not to directly generate the land use map, can help save time and effort. It may be a very time-consuming and labor-intensive process to directly extract the land use map for the entire study area using the object-based approach. Firstly, due to the variance in the shape, size, color, and other properties of objects of different classes or even within each class, using a universal scale parameter for segmentation is not often helpful to extract all types of land cover (Johnson and Jozdani 2018) and/or land use. It is necessary to use a multi-scale approach to achieve the best result; however, this approach may require a great deal of time and effort to determine appropriate parameters. Secondly, it can be seen in our study that land use function zones are not created immediately after the classification step. Further spatial analysis steps are conducted at the post-classification stage to form them, and this stage also requires trial-and-error attempts as well as expert knowledge. Meanwhile, following our proposed workflow could help save time and effort while still creating a highly detailed land use map.

Another limitation of this study is that entertainment complexes (e.g., Dai Nam Wonderland) were placed within the mixed residence class. This is because in Binh Duong province in particular as well as in Vietnam in general, these areas often consist of a mixture of low-albedo construction areas and green spaces, which have a spatial pattern and structure similar to new high-end residential areas in medium-resolution images. The integration of landscape metrics into classification stages, which is similar to studies by Zheng et al. (2016) and Gudmann et al. (2020), will probably help in this case. This approach should be considered in future research.

2.6. Conclusions

In short, land cover and land use in Binh Duong province were interconnected and characterized by a combination of spectral, spatial, and temporal properties. By analyzing that relationship, classification schemes and decision rules for converting a land cover map to a land use map have been effectively defined.

The high overall accuracy of the final land cover map (OA = 93.86%) and the final land use map (OA = 93.45%) produced in this study proved the suitability and effectiveness of a combination of pixel-based and object-based classifications using a random forest classifier and decision rules on free multi-temporal remote sensing data. Using multi-temporal images in a pixel-based classification confirmed their effectiveness for improving the accuracy of the generated land cover map compared to those using single-date images (OA = 89.59–90.78%). Furthermore, by capturing both spectral and spatial information, the segmentation technique and object-based classification created boundaries between regions with different land use types and then relatively precisely formed land use function regions, which paved the way for producing the final land use map.

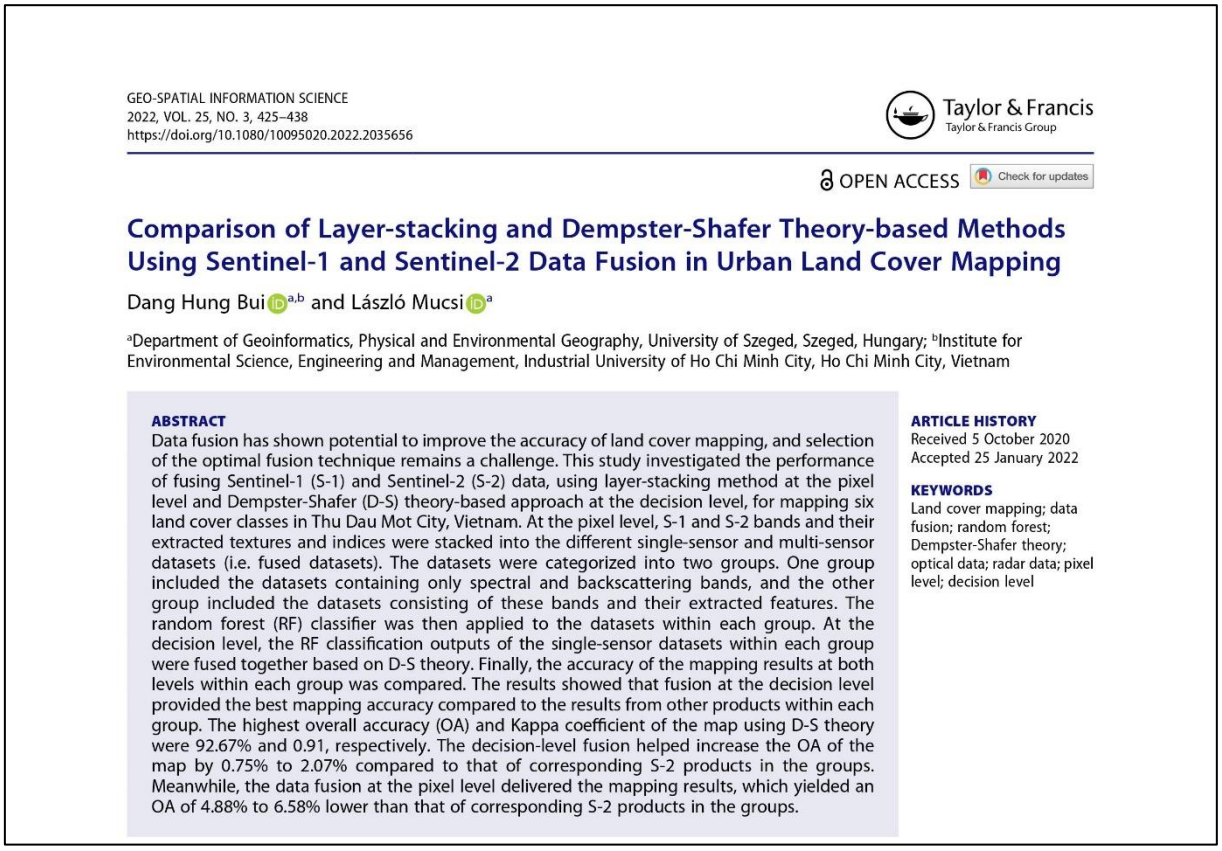
Overall, our research results have shown the potential for separately producing a land cover map and a land use map effectively, which was based entirely on free multi-temporal remote sensing data using the proposed method. These tasks offer many advantages in terms of saving time and cost, increasing the reproducibility and proactivity of the research, and ease of comparison between areas at different times. However, a combination with a field survey and/or expert knowledge of the study area is indispensable, since it promotes a deeper understanding of land cover and land use to define a more accurate classification scheme and reasonable decision rules.

However, despite the effective efforts, there are still some limitations, as discussed, in our study that need to be remedied in future work to further improve work efficiency.

3. Comparison of layer-stacking and Dempster-Shafer theory-based methods using Sentinel-1 and Sentinel-2 data fusion in urban land cover mapping

This article is published in *Geo-spatial Information Science* as:

Bui DH, Mucsi L. 2022. Comparison of layer-stacking and Dempster-Shafer theory-based methods using Sentinel-1 and Sentinel-2 data fusion in urban land cover mapping. Geo-spatial Information Science. 25(3):425-438. https://doi.org/10.1080/10095020.2022.2035656



GEO-SPATIAL INFORMATION SCIENCE
2022, VOL. 25, NO. 3, 425–438
https://doi.org/10.1080/10095020.2022.2035656

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Comparison of Layer-stacking and Dempster-Shafer Theory-based Methods Using Sentinel-1 and Sentinel-2 Data Fusion in Urban Land Cover Mapping

Dang Hung Bui [a,b] and László Mucsi [a]

[a]Department of Geoinformatics, Physical and Environmental Geography, University of Szeged, Szeged, Hungary; [b]Institute for Environmental Science, Engineering and Management, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

ABSTRACT

Data fusion has shown potential to improve the accuracy of land cover mapping, and selection of the optimal fusion technique remains a challenge. This study investigated the performance of fusing Sentinel-1 (S-1) and Sentinel-2 (S-2) data, using layer-stacking method at the pixel level and Dempster-Shafer (D-S) theory-based approach at the decision level, for mapping six land cover classes in Thu Dau Mot City, Vietnam. At the pixel level, S-1 and S-2 bands and their extracted textures and indices were stacked into the different single-sensor and multi-sensor datasets (i.e. fused datasets). The datasets were categorized into two groups. One group included the datasets containing only spectral and backscattering bands, and the other group included the datasets consisting of these bands and their extracted features. The random forest (RF) classifier was then applied to the datasets within each group. At the decision level, the RF classification outputs of the single-sensor datasets within each group were fused together based on D-S theory. Finally, the accuracy of the mapping results at both levels within each group was compared. The results showed that fusion at the decision level provided the best mapping accuracy compared to the results from other products within each group. The highest overall accuracy (OA) and Kappa coefficient of the map using D-S theory were 92.67% and 0.91, respectively. The decision-level fusion helped increase the OA of the map by 0.75% to 2.07% compared to that of corresponding S-2 products in the groups. Meanwhile, the data fusion at the pixel level delivered the mapping results, which yielded an OA of 4.88% to 6.58% lower than that of corresponding S-2 products in the groups.

ARTICLE HISTORY
Received 5 October 2020
Accepted 25 January 2022

KEYWORDS
Land cover mapping; data fusion; random forest; Dempster-Shafer theory; optical data; radar data; pixel level; decision level

Abstract

Data fusion has shown potential to improve the accuracy of land cover mapping, and selection of the optimal fusion technique remains a challenge. This study investigated the performance of fusing Sentinel-1 (S-1) and Sentinel-2 (S-2) data, using layer-stacking method at the pixel level and Dempster-Shafer (D-S) theory-based approach at the decision level, for mapping six land cover classes in Thu Dau Mot City, Vietnam. At the pixel level, S-1 and S-2 bands and their extracted textures and indices were stacked into the different single-sensor and multi-sensor datasets (i.e. fused datasets). The datasets were categorized into two groups. One group included the datasets containing only spectral and backscattering bands, and the other group included the datasets consisting of these bands and their extracted features. The random forest (RF) classifier was then applied to the datasets within each group. At the decision level, the RF classification outputs of the single-sensor datasets within each group were fused together based on D-S theory. Finally, the accuracy of the mapping results at both levels within each group was compared. The results showed that fusion at the decision level provided the best mapping accuracy compared to the results from other products within each group. The highest overall accuracy (OA) and Kappa coefficient of the map using D-S theory were 92.67% and 0.91, respectively. The decision-level fusion helped increase the OA of the map by 0.75% to 2.07% compared to that of corresponding S-2 products in the groups. Meanwhile, the data fusion at the pixel level delivered the mapping results, which yielded an OA of 4.88% to 6.58% lower than that of corresponding S-2 products in the groups.

Keywords: Land cover mapping; data fusion; random forest; Dempster-Shafer theory; optical data; radar data; pixel level; decision level

3.1. Introduction

Land cover information plays an important role in monitoring the environment and natural resources as well as in urban management (Rimal et al. 2017; Arowolo et al. 2018; Grigoraș and Urițescu 2019). Therefore, the knowledge of the spatial distribution and pattern of land cover in a specific area is necessary. Among the various sources for delivering land cover information and producing land cover maps, remote sensing is considered as an essential one due to its efficiency, economic benefits, and reliability (Cai et al. 2019).

Data fusion is defined as a technique that "combines data from multiple sensors, and related information from associated databases, to achieve improved accuracy and more specific inferences than could be achieved by the use of single sensor alone" (Hall and Llinas 1997). In the earth observation field, the rapid development of different kinds of sensors and data sources has made data fusion a vital research approach that aims to extract more detailed information from the remote sensing imagery (Solberg 2006; Zhang 2010; Schmitt and Zhu 2016). By different fusion methods ranging from simple

to complex, the extracted information can effectively serve various fields such as urban management (Guan et al. 2017; Shao, Cheng, et al. 2021; Shao, Sumari, et al. 2021), agriculture (Mfuka et al. 2020; Prins and Van Niekerk 2020), environmental monitoring (Xu and Ma 2021), etc. In general, remote sensing data is fused at three common levels: pixel level, feature level, and decision level (Pohl and van Genderen 2016).

For land cover classification and monitoring, optical and radar data are two types of remote sensing data that are often used as the input for various fusion methods to achieve better mapping results. For instance, Tavares et al. (2019) combined Sentinel-1 (S-1) and Sentinel-2 (S-2) data at the pixel level for urban land cover mapping in Belem, Eastern Brazilian Amazon. The authors used the simple method of layer stacking for fusing data and applied the random forest (RF) algorithm as a classifier. Their results showed that, in comparison to other combinations, the integration of all spectral and backscattering bands achieved the best mapping result with overall accuracy (OA) reached 91.07%. Liu et al. (2018) combined S-1, S-2, Multi-Temporal Landsat 8 and digital elevation model (DEM) data for mapping eight forest types in Wuhan city, China. The authors derived various spectral indices and textures and compositing the data in various scenarios. Afterward, they applied a complex hierarchical strategy, including multi-scale segmentation, threshold analysis, and the RF algorithm. Their results showed that the fusion of imagery, terrain, and multi-temporal data reached the highest classification accuracy (OA = 82.78%) among the scenarios. Tabib Mahmoudi, Arabsaeedi, and Alavipanah (2019) classified urban land cover by fusing Landsat-8 and Terra SAR-X textures images at the feature level. They used the multi-resolution segmentation technique and knowledge-based classification based on thresholds and decision rules to fuse the data. The accuracy of the fusion result was not too high, as the OA and Kappa coefficient were 50.53% and 0.37, respectively. However, they improved by 2.48% and 0.06, respectively, compared to that of Landsat-8 imagery. Shao et al. (2016) fused S-1 and Gaofen-1 images at the decision level based on Dempster-Shafer (D-S) theory to map the urban impervious surfaces in the metropolitan area of Wuhan city in China. Their results indicated that fusion at the decision level achieved an OA ranging from 93.37% to 95.33%, which is better than those from single-sensor data. Ban, Hu, and Rangel (2010) fused Quickbird multi-spectral (MS) and RADARSAT synthetic aperture radar (SAR) data at the decision level for mapping 16 urban land cover classes at the rural–urban fringe of the Greater Toronto Area, Ontario, Canada. Complex hierarchical object-based and rule-based approaches were applied in both single-sensor data and their fused outputs. The study results revealed that decision-level fusion helped improve the accuracy of some vegetation classes by a range from 17% to 25%. In addition, some emerging data sources, such as LiDAR or social data, can also be used in conjunction with conventional data sources. For example, Prins and Van Niekerk (2020) investigated the effectiveness of combining LiDAR, Sentinel-2, and aerial imagery for classifying five crop types. The data were combined in various ways,

and 10 machine learning algorithms were used. Their results showed that the highest OA of 94.4% was achieved when applying the RF algorithm on the combination of all three data sources. Shao et al. (2021) combined Landsat images and Twitter's location-based social media data to classify urban land use/land cover and analyze urban sprawl in the Morogoro urban municipality, Africa. Their results proved the potential of combining remote sensing, social sensing, and population data for classifying urban land use/land cover and evaluating the expansion of urban areas and the status of access to urban services and infrastructure.

These study results demonstrate that fusion data from various sources at the three fusion levels can improve accuracy in land cover mapping. In these studies, various fusion techniques were used, ranging from simple to very complex methods. However, selecting which fusion method should be applied to deliver the best results is a challenge. In general, selecting a method for image classification depends on many factors. The factors comprise the purpose of study, the availability of data, the performance of the algorithm, the computational resources, and the analyst's experiences (Lu and Weng 2007). In addition, the performance of each method also depends partly on the characteristics of the study area, the dataset used, and how the method works. A method can yield highly accurate results in one dataset and give poor results in others (Xie et al. 2019). Moreover, it is not necessary to employ a complicated technique when a simple one can solve the problem well. Therefore, for studies related to land cover mapping, it is essential to compare the performance of different methods to choose the optimal one that gives the most accurate results.

Since being launched into space in 2014 under the Copernicus program (The European Space Agency 2021), Sentinel-1 and Sentinel-2 missions provide a high-quality satellite imagery source for earth observation. The Sentinel-1 mission comprises a two-satellite constellation: Sentinel-1A (S-1A) and Sentinel-1B (S-1B). The mission provides C-band SAR images with a 10-m spatial resolution and a 6-day temporal resolution. Meanwhile, the Sentinel-2 mission also consists of a two-satellite constellation: Sentinel-2A (S-2A) and Sentinel-2B (S-2B). S-2A/B data together have a revisit time of 5 days, and they deliver the multi-spectral products with a spatial resolution ranging from 10 m to 60 m. The advantages of the Sentinel data are a high spatial resolution and a short revisit time, and S-2 are multi-spectral, while S-1 are unaffected by cloud and acquiring time. Furthermore, they are free and easy to access and download. Combining these data can help enhance the efficiency of monitoring land cover information, and as mentioned, selection of the optimal combination method is needed. To the extent of the authors' knowledge from the literature review, no study to date has compared the efficiency of the fusion of S-1 and S-2 data at the pixel level and decision level for land cover mapping.

With these issues in mind, the purpose of this paper is to evaluate and compare the performance of fusing S-1 and S-2 data at the pixel level and decision level for land

cover mapping in a case study of Thu Dau Mot City, Binh Duong province, Vietnam. To achieve this objective, our proposed procedure is briefly highlighted as follows:

- Pre-processing data and deriving textures and indices.
- Stacking the obtained products into different datasets.
- Applying the RF algorithm on the datasets to produce land cover maps at pixel level.
- Fusing the RF results of single-sensor datasets based on D-S theory to produce land cover maps at decision level.
- Comparing the accuracy of the mapping results at both levels.

3.2. Study area

Thu Dau Mot City is the administrative, economic, and cultural center of Binh Duong province, Vietnam. The city is located in the southwest of the province, between 10°56′22″ to 11°06′41″ N latitude and 106°35′42″ to 106°44′00″ E longitude (Figure 3.1). It belongs to the tropical monsoon climate, which has the rainy season from May to November and the dry season from December to April of the following year. Its annual mean temperature is 27.8°C; its annual rainfall ranges from 2104 mm to 2484 mm; and its annual mean air humidity varies from 70 to 96% (Binh Duong Statistical Office 2019). The mean elevation of the city is from 6 to 40 m, and it increases from west to east and from south to north. However, the terrain surface is relatively flat, and the majority of the city has a slope of 7 degrees or less. The total area of the city is about 118.91 km^2, and its population was 306,564 in 2018 (Binh Duong Statistical Office 2019).

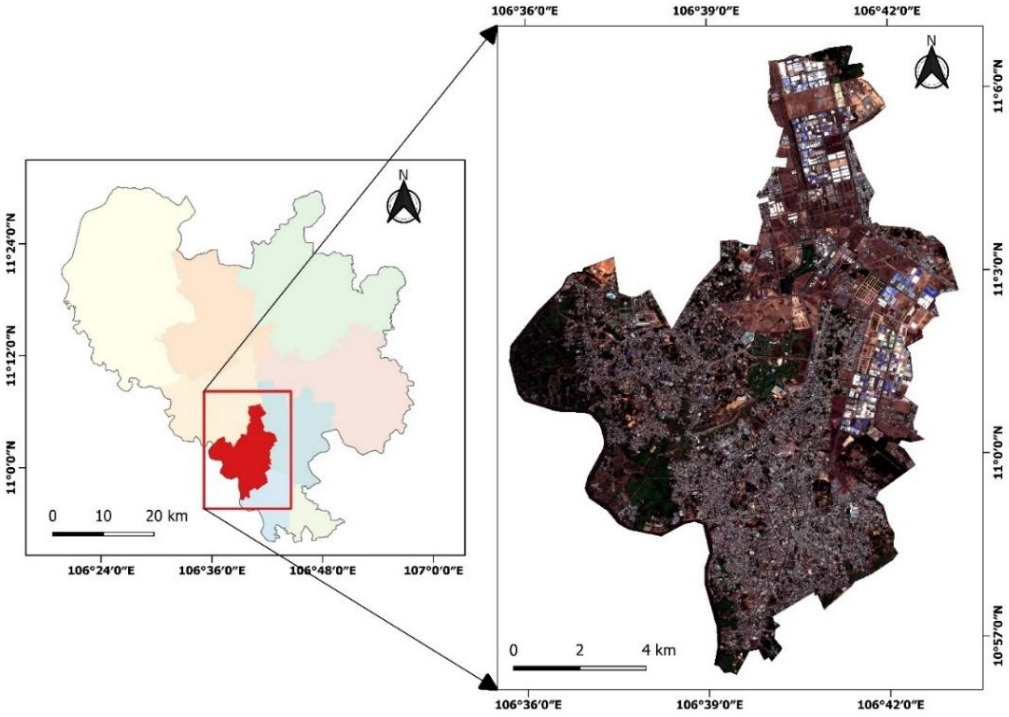


Figure 3.1. Study area.

The main types of land cover in the city are built up, vegetation, bare land, and water surface. Based on a field survey trip in January 2020 and careful consideration of the characteristics of each land cover subject, the land cover in the study area was categorized into the following classes (Figure 3.2):

i. Bare land with high albedo (BL_H): including totally bare soil areas without any cover or very little vegetation.
ii. Bare land with low albedo (BL_L): including bare land areas partly covered with sunburned vegetation and/or little fresh vegetation.
iii. Built-up with high albedo (BU_H): mainly including factories and industrial buildings that are often light-colored corrugated iron or concrete.
iv. Built-up with low albedo (BU_L): mainly including residences, commercial and office buildings, and roads that are often concrete, clay, tole, asphalt, or a mix of these materials.
v. Vegetation (VE): including crops, fruit trees, industrial trees, mature trees for landscaping, and fresh grass.
vi. Open water surface (WA): including rivers, canals, lakes, ponds, and pools.

(a) Bare land with high albedo
(b) Bare land with low albedo
(c) Built up with high albedo
(d) Built up with low albedo
(e) Vegetation
(f) Open water surface

Figure 3.2. Land cover classes in the study area.

3.3. Materials and methods

3.3.1. Data

3.3.1.1. Satellite images

One free-cloud tile of S-2A Multispectral Instrument (MSI) Level-2A and one tile of S-1A Ground Range Detected (GRD), which cover the study area, were downloaded from the Copernicus Scientific Data Hub (https://scihub.copernicus.eu/).

The S-2A MSI Level-2A product provides the bottom of atmosphere (BOA) reflectance images. The product includes four bands of 10 m (2, 3, 4, 8), six bands of 20 m (5, 6, 7, 8A, 11, 12), and two bands of 60 m (1, 9). The cirrus band 10 was omitted as it does not contain surface information. The product's band wavelength ranges from about 493 nm to 2190 nm, and its radiometric resolution is 12 bits.

The S-1A GRD product provides the C-band SAR data, which had been detected, multi-looked and projected to ground range using an Earth ellipsoid model. The acquired imagery was collected in the Interferometric Wide Swath (IW) mode with high resolution (a pixel spacing of 10 m and a spatial resolution of approximately 20 m × 22 m) in dual-polarization mode: vertical transmit-vertical receive (VV) and vertical transmit-horizontal receive (VH).

Due to its climatic characteristics, the study area is often covered by clouds during the rainy season (i.e. from May to November). Therefore, in this study, the optical product was collected in the dry season. One free-cloud tile of S-2, acquired on 22 February 2020 was selected. Meanwhile, although the radar product is not affected by cloud coverage, the selected tile of S-1 was acquired on 25 February 2020 to minimize the change in the land cover.

3.3.1.2. Vector data

The administrative boundary of the study area was downloaded from the Database of Global Administrative Areas (GADM) project website (https://gadm.org/). It was used for subsetting and masking the satellite images.

The training dataset for the six land cover classes was collected based on the results of the field trip in January 2020 combined with Google Earth images. The validation data were collected based on a stratified random sampling strategy. Based on the classification result of the S-2 dataset, the proportion of each land cover class was roughly estimated by visual observation. Based on the proportion, 70 points of BL_H, 150 points of BL_L, 90 points of BU_H, 150 points of BU_L, 140 points of VE, and 50 points of WA were randomly selected. Thus, a total of 650 points were generated. These points were visually interpreted by the S-2 image, Google Earth image, and the authors' personal knowledge. Some points being on mixed pixels, which could not be interpreted correctly, were discarded. As a result, only 532 points could be used for validation,

including 56 points of BL_H, 86 points of BL_L, 89 points of BU_H, 135 points of BU_L, 115 points of VE, and 51 points of WA.

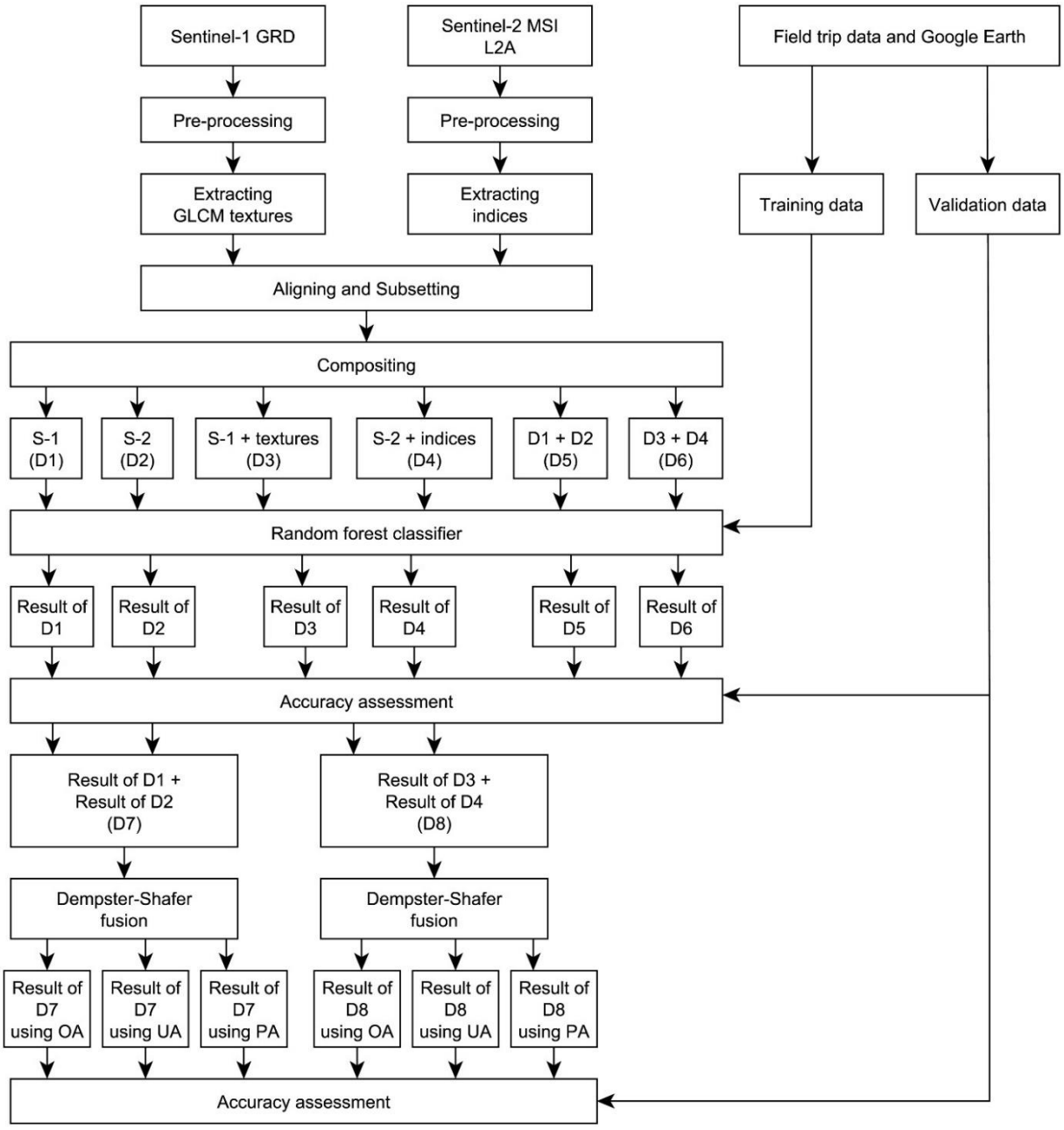
3.3.2. Methods

Five main steps were carried out to achieve the study goals. First, the downloaded S-1 and S-2 data were pre-processed, and their textures and indices were extracted. In the second step, the products obtained were stacked into different datasets, including the datasets from single sensors and the fused datasets from multiple sensors. The datasets were categorized into two groups based on whether they included textures and indices or not. In the third step, the RF classifier was then applied to each dataset, and the accuracy of their results was assessed. In the fourth step, the classification results of the single-sensor datasets within each group were used as the inputs for the decision-level fusion based on D-S theory. Finally, the accuracy of classification results at the decision level was assessed and compared to those at the pixel level. The overall process followed in this study is presented in Figure 3.3 and described in detail below.

3.3.2.1. Pre-processing and extracting indices and textures

The S-2 tile was downloaded as a Level 2 product in WGS 84/UTM Zone 48 N projection, which has already applied geometric and atmospheric correction and is ready to use for classification. Bands 2, 3, 4, 8 (10 m) 5, 6, 7, 8A, 11, 12 (20 m) were used in this study. The 20-m bands were resampled to the 10-m ones using the nearest neighbor method to ensure the preservation of original values. Then, the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) were extracted. These two indices were included in this study because they have been widely used and have shown the potential to improve land cover classification results (Shao et al. 2016; Tian et al. 2016).

Several common pre-processing steps were applied with the downloaded S-1 GRD tile. They included apply orbit file, thermal noise removal, calibration, speckle filtering, range-Doppler terrain correction using WGS 84/UTM Zone 48 N projection and 30 m Shuttle Radar Topography Mission (SRTM), and conversion to dB (sigma0 dB) for both VH and VV. The pre-processed products had a resolution of 10 m. Speckle filtering was used for reducing noise to improve image quality (Filipponi 2019); however, it also can lead to a massive loss of information when extracting texture features (Hu et al. 2018). Therefore, there were two sets of products in this step: VH and VV with speckle filtering were used as input data for classifiers, and the ones without speckle filtering were used for extracting textures. Afterward, eight gray-level co-occurrence matrix (GLCM) textures were derived for both VH and VV by using a 9 × 9 window size in all directions. The derived textures included mean, correlation, variance, homogeneity, contrast, dissimilarity, entropy, and angular second moment. As a result, 16 texture products were generated.



Note: S-1 = Sentinel-1; S-2 =Sentinel-2; GLCM = grey-level co-occurrence matrix; OA = overall accuracy; UA = user's accuracy; PA = producer's accuracy.

Figure 3.3. Process flowchart.

Because there was a small shift in pixels between the optical and SAR products, the resulting products were aligned using band 2 of S-2 as a reference image to make them fit together. Finally, all products were subset to the study area. These pre-processing steps were conducted on the Sentinel Application Platform (SNAP) and Quantum Geographic Information System (QGIS) software.

3.3.2.2. Combination, classification, and accuracy assessment

After pre-processing, the products were stacked into different datasets, including the datasets from single sensors (D1, D2, D3, and D4) and the fused datasets from multiple sensors (D5 and D6). This study applied the common combination method of layer

stacking to fuse the data from S-1 and S-2 together at the pixel level. The datasets were then categorized into two groups: a group of datasets containing only spectral and backscattering bands (group 1) and a group of datasets consisting of these bands and their extracted textures and indices (group 2). Table 3.1 summarizes the information of all datasets.

Table 3.1. Summary of the input datasets.

Dataset	Data sources	Variables	Note
D1	S-1 only	VH, VV	Group 1
D2	S-2 only	2, 3, 4, 5, 6, 7, 8, 8A, 11, 12	Group 1
D3	S-1 with GLCM textures	VH, VV, and textures of mean, correlation, variance, homogeneity, contrast, dissimilarity, entropy, and angular second moment of VH and VV	Group 2
D4	S-2 with indices	2, 3, 4, 5, 6, 7, 8, 8A, 11, 12, NDVI, NDWI	Group 2
D5	D1 and D2	All variables of D1 and D2	Group 1, pixel-level fusion
D6	D3 and D4	All variables of D3 and D4	Group 2, pixel-level fusion
D7	Random forest results of D1 and D2	Probability of each land cover class, and OA, or UA, or PA of each result	Group 1, decision-level fusion
D8	Random forest results of D3 and D4	Probability of each land cover class, and OA, or UA, or PA of each result	Group 2, decision-level fusion

Note: GLCM = gray-level co-occurrence matrix; VV = vertical transmit-vertical receive; VH = vertical transmit-horizontal receive; NDVI = Normalized Difference Vegetation Index; NDWI = Normalized Difference Water Index; OA = overall accuracy; UA = user's accuracy; PA = producer's accuracy.

In this study, the RF algorithm, developed by Breiman (2001), was selected as the classifier for land cover classification at the pixel level. A random forest consists of a set of decision trees, each of which is generated by randomly drawing a subset from the training dataset. From the results of the trees, a majority vote is conducted to determine the final output (Xie et al. 2019). RF is easy to use, highly efficient, fast to process, and suitable for remote sensing applications (Belgiu and Drăguţ 2016; Gudmann et al. 2020). Since its results come from voting, RF has the ability to produce classification output as probabilities of each class, which was used as the input for fusion at the decision level. The classification process was implemented on R software, using the "randomForest" package (Liaw and Wiener 2002). Two important parameters affecting the classification performance of the RF model are the maximum number of trees (ntree) and the number of variables randomly sampled as candidates at each split (mtry). The mtry parameter was set at the default value, which is equal to the square root of the total number of features. After testing the relationship between the ntree and the decrease in out-of-bag error rates, the ntree was set at 300 trees as out-of-bag error

rates were relatively stable after this point. The composited datasets were used as inputs for the classification process. As a result, six land cover maps were generated at the pixel level, and their accuracy was then assessed. In addition, four classification results of single-sensor datasets, in the form of probabilities of each land cover class, were also produced to use in the next stages.

At the decision level, the probability-form classification results were fused within each group. The classification result of D1 was fused with that of D2 (D7), while the results of D3 and D4 were combined (D8). This study applied the data fusion method based on the D-S evidence theory (Dempster 1967; Shafer 1976) using the dst package (Boivin and Stat.ASSQ 2020) in R software. D-S evidence theory, which is often described as a generalization of the Bayesian theory, is based on belief functions and plausible reasoning. The advantages the theory offer in data classification include: (i) flexible construction of the mass function and the data organization; (ii) no requirement regarding the prior knowledge or conditional probabilities, which makes it suitable for handling data with unseen labels; and (iii) possibility to provide the uncertainty of the result (Chen et al. 2014). Theoretical calculation steps were carried out according to the detailed description of Shao et al. (2016). The Basic Probability Assignment (BPA – or mass function) of each pixel, which is a prerequisite for fusion according to D-S theory, was calculated as follows:

$$m_i\,(A) = p_v \times p_i$$

in which $m_i\,(A)$ is the mass function value of the calculated pixel in class A of data source i, p_v is the probability of belonging to each land cover class of the calculated pixel, and p_i is the probability of correct classification of data source i. In this study, the OA, user's accuracy (UA) and producer's accuracy (PA) were used in turn to measure the probability of correct classification for the calculation.

As a result, six land cover maps (two by using OA, two by using UA, and two by using PA) were generated at this decision level, and their accuracy was then assessed. Finally, the accuracy of all classification results at both pixel and decision levels was compared by both visual assessment and OA, PA, UA, and Kappa coefficients.

3.4. Results and discussion

The accuracy assessments of all classification results are presented in Table 3.2. The land cover maps of the two groups are also presented in Figures 3.4 and 3.5. The fusion results using PA were chosen as a representation of the decision level in these figures.

In group 1, the fusion method using D-S theory provided the most accurate results, in which OA ranged from 90.23% to 90.60% and the Kappa coefficient was 0.88. The best result in this group occurred in the fusion of D7, based on the OA. Similarly, results from the decision-level fusion in group 2 also gave the highest accuracy, with the OA ranging from 91.35% to 92.67% and the Kappa coefficient varying from 0.89 to 0.91. The fusion of D8 using UA produced the best result in this

Table 3.2. Comparison of the overall accuracy and Kappa coefficient of the classification result of all datasets.

Dataset	Overall Accuracy (%)	Kappa coefficient
Group 1: datasets without textures and indices		
D7 using OA	90.60	0.88
D7 using PA	90.23	0.88
D7 using UA	90.23	0.88
D2	89.47	0.87
D5	84.59	0.81
D1	42.86	0.29
Group 2: datasets with textures and indices		
D8 using UA	92.67	0.91
D8 using PA	91.92	0.90
D8 using OA	91.35	0.89
D4	90.60	0.88
D6	84.02	0.80
D3	52.07	0.41

Note: OA = overall accuracy; UA = user's accuracy; PA = producer's accuracy.

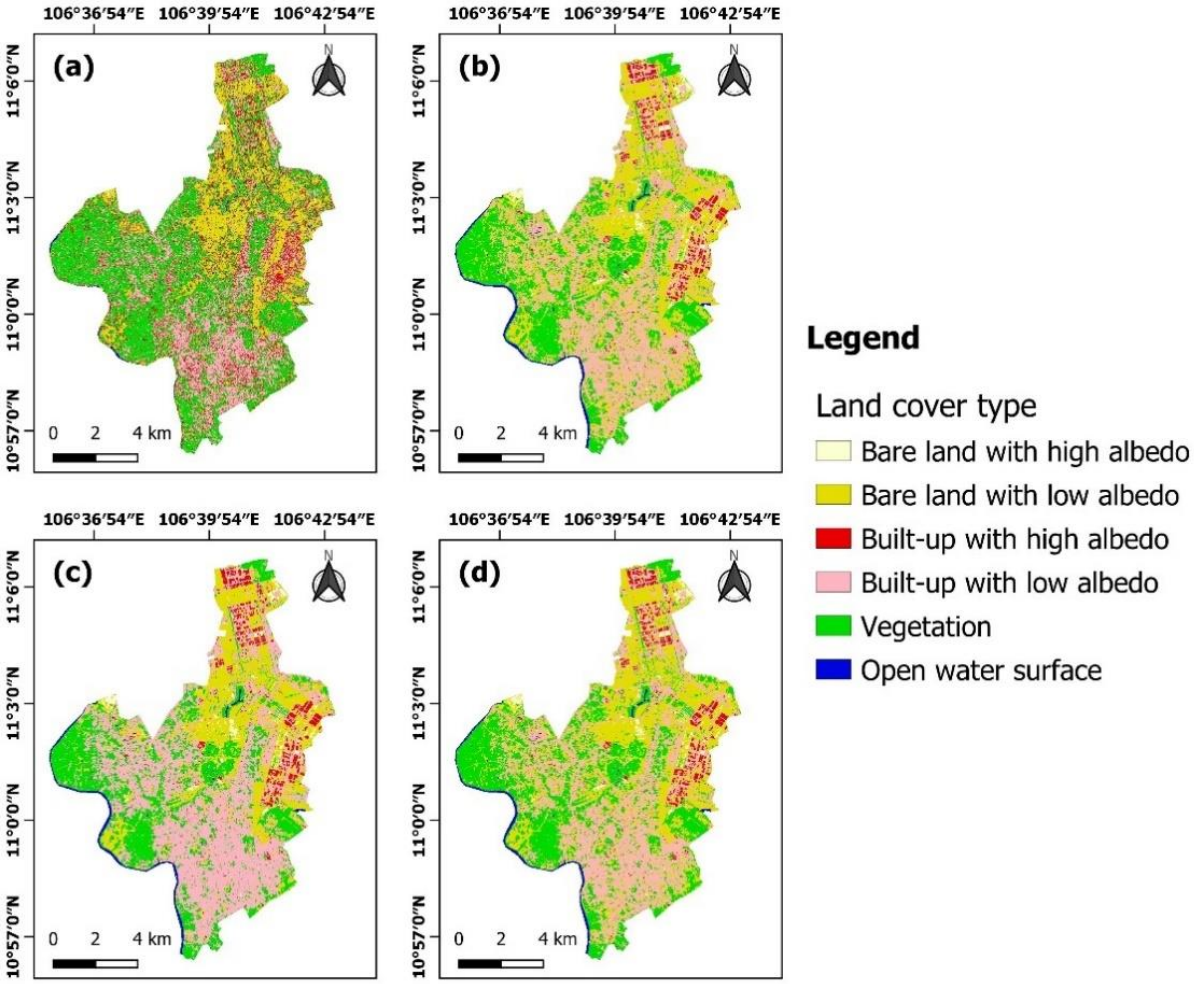


Figure 3.4. Land cover maps from the datasets without textures and indices: (a) dataset D1; (b) dataset D2; (c) dataset D5; (d) dataset D7 using PA.

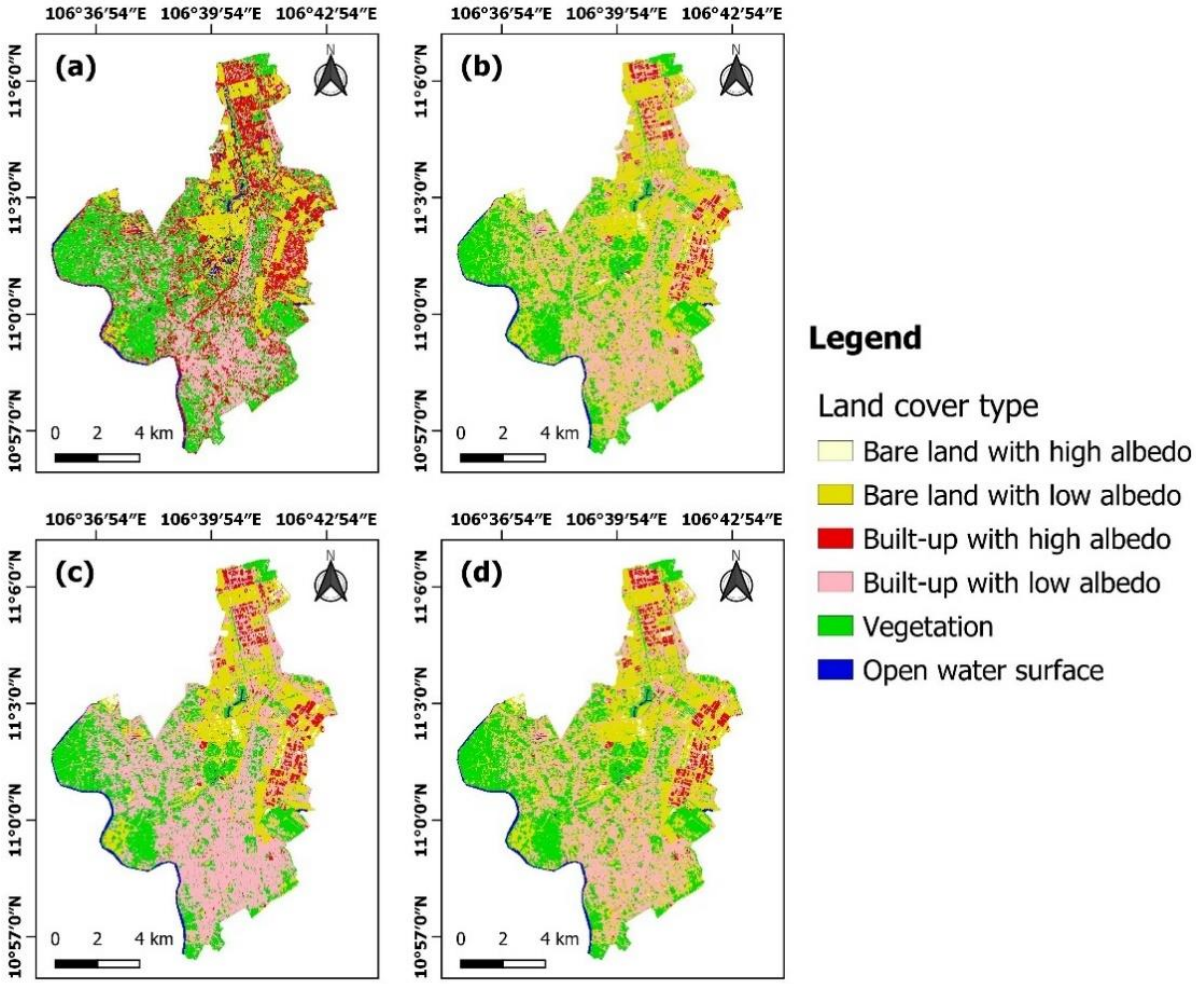


Figure 3.5. Land cover maps from the datasets with textures and indices: (a) dataset D3; (b) dataset D4; (c) dataset D6; (d) dataset D8 using PA.

group with an OA of 92.67% and a Kappa coefficient of 0.91. It was also the product with the most accuracy in all datasets. Therefore, the highest accuracy was found in the results of fusion at the decision level in both groups, whether using OA, UA, or PA for mass function construction. In contrast, the poorest results occurred in S-1 only (OA = 42.86%, Kappa = 0.29) in group 1 and in S-1 with its texture variables (OA = 52.07%, Kappa = 0.41) in group 2. In general, both groups followed a similar trend in the accuracy of mapping results from datasets and decreased in the following order: decision-level fusion dataset, single optical dataset, pixel-level fusion dataset, and single SAR dataset.

As a result, the fusion results from S-1 and S-2 products at the decision level increased mapping accuracy by a range of 0.75% to 2.07% in comparison to the results of corresponding S-2 products in the two groups. D-S theory considered each land cover class from different inputs as independent evidence. Evidential probability was constructed entirely based on the results of the classification algorithm at the pixel level, without taking into account the input of that algorithm. Therefore, this evidence theory could reduce the impact of noise data and feature selection in land cover classification (Shao et al. 2016). By that advantage, the use of D-S theory at the decision level in this

study produced mapping results with a higher level of accuracy. This finding is consistent with many previous studies (Ran et al. 2012; Shao et al. 2016; Mezaal et al. 2018). It is clear that the result of the D-S fusion depends on how the mass function is constructed. Mezaal, Pradhan, and Rizeei (2018) converted the posterior probabilities of the classification results to the form of mass function directly; Ran et al. (2012) identified the parameter for the mass function construction from a literature review and expert knowledge; The p_v parameter of the mass function in Shao et al. (2016) was similar to our study, and the p_i was based on the PA of each class. This study is distinguished by testing the construction of mass function using the OA, UA, and PA in turn for the parameter of p_i to get a more comprehensive assessment of the effectiveness of the D-S theory-based fusion. As mentioned, each of the three construction methods yielded better results than that of single-sensor and fused datasets at the pixel level. The results show that whether using the OA, UA, or PA for mass function construction, applying the D-S fusion method on S-1 and S-2 data provides a better result for land cover mapping. In addition, the results suggest that such a method is applicable for high-accuracy mapping in other urban areas.

However, the fusion data from different sensors using the layer-stacking technique at the pixel level did not improve classification efficiency. It reduced the accuracy of classification by a range of 4.88% to 6.58% compared to the results of corresponding optical products in the two groups. Although most studies in the literature reported the ability to improve the overall accuracy when fusing various data sources at the pixel level compared to using a single data source, some studies have shown the opposite (de Furtado et al. 2015; Fonteh et al. 2016). Zhang and Xu (2018) found that whether the combination of optical and SAR data could improve the accuracy of urban land cover mapping or not depended on the fusion levels and the fusion methods. Therefore, in our study, the extraction and selection of variables as well as the choice of combination technique and classification algorithm may have influenced the outcome of the classification. To improve mapping performance at the pixel level, further studies are needed to determine the optimal variable selection for data integration and to test other fusion techniques, such as the component substitution methods or the multi-scale decomposition methods (Kulkarni and Rege 2020).

When comparing the results from group 1 and group 2, the accuracy of most of the datasets containing indices and textures was higher than that of the corresponding datasets without these extracted variables, except for the pair of datasets D5 and D6. The most significant increase took place in the pair of datasets D1 and D3, where the addition of the GLCM textures along with VH and VV raised the OA by 9.21%. The accuracy of the remaining pairs also increased by a range of 1.12% to 2.07% when including these extracted variables in the datasets. This finding confirms that the GLCM textures can provide additional useful information to improve classification results (Lu et al. 2014; Zakeri et al. 2017; Tavares et al. 2019); however, the effectiveness of

spectral indices is still controversial. Our results showed the spectral indices were effective in land cover classification to some extent. While many studies have included some common spectral indices (e.g. NDVI, NDWI, and Normalized Difference Built-up Index) in the input dataset and enhanced the accuracy of mapping results (Shao et al. 2016; Tian et al. 2016; Abdi 2020), other studies have indicated the opposite results (Tavares et al. 2019; Adepoju and Adelabu 2020). This discrepancy may result from differences in land cover characteristics of the study areas and the selection of indices included in the dataset. Therefore, these indices should be used with caution in future studies.

A detailed comparison of PA and UA in each class of each classification result is presented in Tables 3.3 and 3.4. In addition, three example regions from classification maps in group 2 are presented in Figure 3.6 to provide a visual comparison. The Google Earth images were captured on 16 April 2020 using the historical imagery function on Google Earth Pro software. As seen in Tables 3.3 and 3.4, while S-1 only and S-1 with GLCM texture classification results yielded relatively low accuracy, the majority of PA and UA of all classes from other classifications were high (over 85%). BL_L was the class that had the most misclassifications, which resulted in the lowest accuracy in most cases.

At the pixel level, the fusion data from different sources significantly reduced the PA of BL_L and the UA of BU_L when compared to the corresponding S-2 products in both fusion cases. The former was reduced by 31.39% in the datasets without derived products and by 27.91% in the datasets with derived products. Meanwhile, the latter was decreased by 19.79% in the datasets of group 1 and by 22.05% in the datasets of group 2. The misclassification between these two classes could be clearly seen in the three

Table 3.3. The producer's accuracy and user's accuracy of the classification result of the datasets without textures and indices.

Dataset	Accuracy index	Class					
		BL_H	BL_L	BU_H	BU_L	VE	WA
D1	PA (%)	10.71	46.51	41.57	43.70	64.35	23.53
	UA (%)	60.00	31.50	37.37	51.30	49.01	40.00
D2	PA (%)	91.07	93.02	87.64	83.70	91.30	96.08
	UA (%)	89.47	73.39	92.86	91.13	97.22	98.00
D5	PA (%)	92.86	61.63	92.13	86.67	86.96	90.20
	UA (%)	94.55	69.74	95.35	71.34	97.09	95.83
D7 using OA	PA (%)	91.07	94.19	86.52	88.15	92.17	94.12
	UA (%)	89.47	78.64	95.06	89.47	97.25	97.96
D7 using UA	PA (%)	91.07	88.37	86.52	89.63	92.17	96.08
	UA (%)	89.47	83.52	92.77	86.43	96.36	96.08
D7 using PA	PA (%)	91.07	95.35	86.52	85.93	92.17	94.12
	UA (%)	89.47	77.36	95.06	89.92	96.36	97.96

Note: BL_H = Bare land with high albedo; BL_L = Bare land with low albedo; BU_H = Built-up with high albedo; BU_L = Built-up with low albedo; VE = Vegetation; WA = Open water surface; OA = overall accuracy; UA = user's accuracy; PA = producer's accuracy.

Table 3.4. The producer's accuracy and user's accuracy of the classification result of the datasets with textures and indices.

Dataset	Accuracy index	Class					
		BL_H	BL_L	BU_H	BU_L	VE	WA
D3	PA (%)	7.14	47.67	78.65	45.93	60.00	60.78
	UA (%)	50.00	44.57	46.67	50.82	58.47	73.81
D4	PA (%)	96.43	91.86	92.13	83.70	95.65	86.27
	UA (%)	94.74	73.83	94.25	91.87	96.49	100.00
D6	PA (%)	83.93	63.95	91.01	87.41	87.83	88.24
	UA (%)	95.92	74.32	88.04	69.82	98.06	100.00
D8 using OA	PA (%)	91.07	90.70	95.51	87.41	95.65	86.27
	UA (%)	98.08	75.00	97.70	90.77	95.65	100.00
D8 using UA	PA (%)	94.64	87.21	95.51	91.11	94.78	94.12
	UA (%)	98.15	83.33	98.84	89.13	95.61	96.00
D8 using PA	PA (%)	92.86	91.86	95.51	86.67	95.65	90.20
	UA (%)	98.11	75.96	97.70	92.13	95.65	100.00

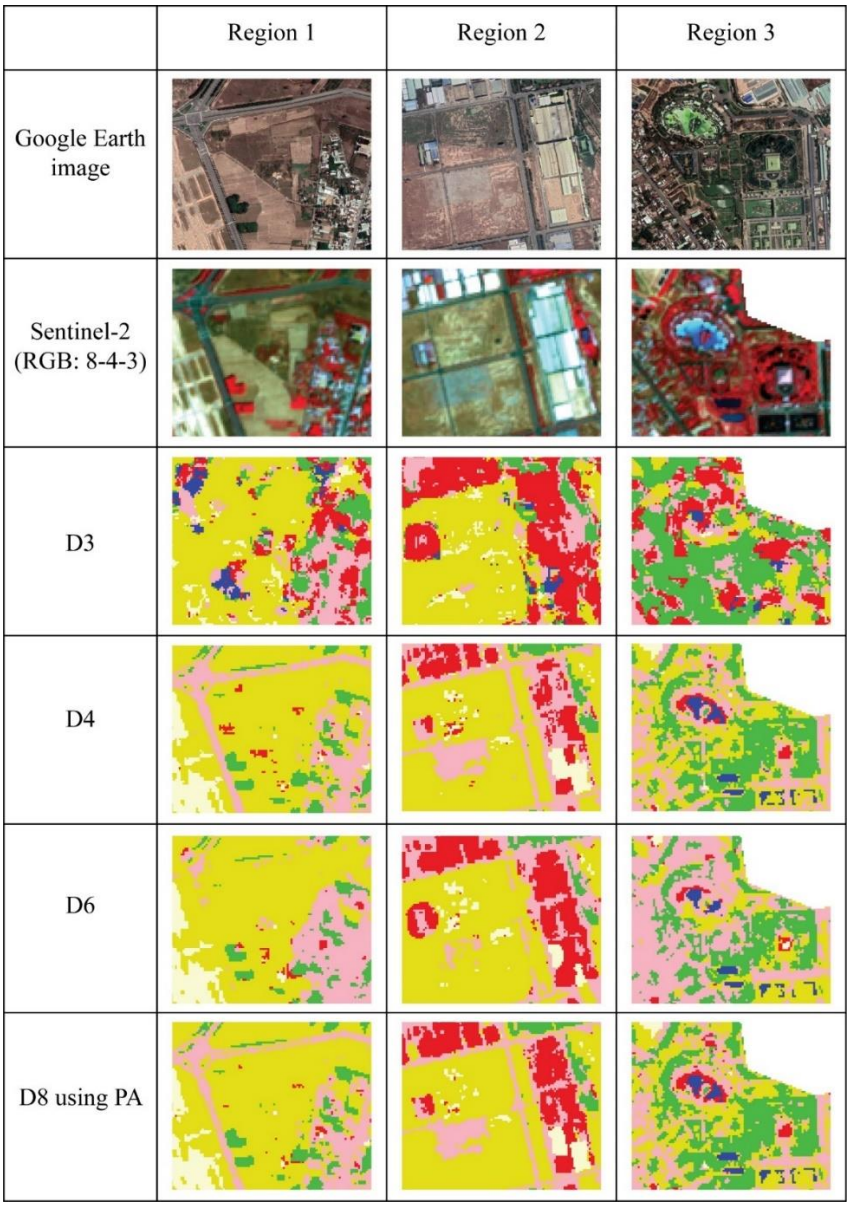


Figure 3.6. Comparison of the classification results from the datasets with textures and indices in three example regions.

sample regions, in which the BU_L areas, especially roads, were misclassified as bare land. Moreover, with the BU_H class, the misclassification from bare land areas to factories and from factories to low albedo built-up areas decreased, but the misclassification between factories and totally bare soil areas increased. Therefore, the UA and PA of classes increased or decreased unevenly, but overall, the total reduction was greater than the total increase in both fusion cases.

On the contrary, at the decision level, although the UA and PA of classes also increased or decreased unevenly, the total reduction was lower than the total increase in both fusion cases. By visual assessment, the greatest improvement was found in the classes BU_H, BU_L, and BL_L. In these classes, the misclassification from high-albedo build-up to bare soil and to low-albedo built-up was significantly reduced, contributing to the increase in the OA of the mapping result. However, because the BU_H class only took a small proportion of the study area (about 5% of the total area), the reduced misclassification only resulted in a slight increase in the OA compared to the maps from the optical datasets.

In general, in most cases of both single-sensor datasets and integrated datasets, the BU_L and BL_L had the highest rate of misclassification among all classes, which may be due to the similarity in their spectral characteristics. The study results of Chen et al. (2019), Li et al. (2017), Shao et al. (2016), and Wei et al. (2020) and many others have also shown this issue. Meanwhile, although the UA of water class achieved up to 100%, some water areas were misclassified as high albedo built-up area by visual assessment in all datasets at the nearshore of an artificial swimming pool in example region 3. The misclassification from WA to BU_H in this region may be explained by a few factors. First, the pool is in the Dai Nam Wonderland water park, and in fact, it is an artificial sea with saline water, not a freshwater swimming pool. The depth of this artificial sea gradually rises from the nearshore to the offshore, where the shallower water leads to higher reflectance contribution from the floor material of the water area (Chuvieco and Huete 2016); Second, the floor of this artificial sea is made of light-colored concrete, which belongs to BU_H class. These factors combined may have caused the misclassification from water to high-albedo built-up area at the nearshore area of the sea. For the vegetation class, the difference in the accuracy was not significant between the fused datasets and corresponding optical datasets.

3.5. Conclusions

In summary, the fusion of S-1 and S-2 data based on D-S theory at the decision level yielded better mapping results compared to others. It comes from the advantages of the D-S theory-based technique in reducing the impact of noise data and feature selection in land cover classification. The most obvious improvement was found in the classes of barren land and built up. As a result, the datasets fused at the decision level increased the OA by a range of 0.75% to 2.07% compared to the S-2 datasets. The fusion of S-1

and S-2 data with their derived textures and indices at the decision level using D-S theory brought the best results in this study, achieving an OA and Kappa coefficient of 92.67% and 0.91, respectively.

Moreover, the integration of SAR and optical products using the layer-stacking technique at the pixel level did not give more power to the classification process. It reduced the accuracy of the mapping result by 4.88% to 6.58% compared to that of the optical datasets. These findings may be influenced by the processing and selection of features, fusion technique, and classifier. Further studies on this issue are needed.

Furthermore, the inclusion of GLCM textures and spectral indices in the datasets helped improve the mapping results. However, while the effectiveness of the textures is clear, the contribution of the indices needs to be studied further.

In general, the results of this study show that using the D-S fusion method for high-accuracy mapping in other urbanized areas holds great potential. This study represents an initial step, and it paves the way for further research on land cover mapping using additional available data from the active and passive sensors for performance improvement.

4. Land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020

This article is published in *Geocarto International* as:

Bui DH, Mucsi L. 2022. Land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020. Geocarto International. 37(27):17096–17118. https://doi.org/10.1080/10106049.2022.2123564

GEOCARTO INTERNATIONAL
2022, VOL. 37, NO. 27, 17096–17118
https://doi.org/10.1080/10106049.2022.2123564

Taylor & Francis
Taylor & Francis Group

Check for updates

Land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020

Dang Hung Bui[a,b] and László Mucsi[a]

[a]Department of Geoinformatics, Physical and Environmental Geography, University of Szeged, Szeged, Hungary; [b]Institute for Environmental Science, Engineering and Management, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

ABSTRACT

This study aims to analyze land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020. Multitemporal Landsat images were used to develop land-use maps. Area statistics and transition matrices were employed to explore the land-use change; meanwhile, annual expansion rate (AER), expansion contribution rate (ECR), and district-, ring-, and sector-based analyses were employed to analyze the urban expansion. The results showed that there was a large transition from agricultural and unused lands to other uses. This resulted in an expansion of developed areas, recreational regions, mining sites, and water surfaces, a drastic decline of agricultural land for annual crops, and a fluctuation of perennial cropland and unused land. The study also indicated that the urban area has expanded 65 times within 25 years at an increasing rate. The AER and ECR were uneven between subregions, and there was a gradual expansion and shift from south to north of the province. The factors affecting the changes comprise natural conditions, development histories, policies and practices for urbanization, industrialization, and agricultural development, and product price fluctuations in the market. Practical lessons learned from this study could be useful for land planning and policymaking in other localities.

ARTICLE HISTORY
Received 11 May 2022
Accepted 6 September 2022

KEYWORDS
Land-use change; urban expansion; change detection; spatiotemporal analysis; remote sensing

Abstract

This study aims to analyze land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020. Multitemporal Landsat images were used to develop land-use maps. Area statistics and transition matrices were employed to explore the land-use change; meanwhile, annual expansion rate (AER), expansion contribution rate (ECR), and district-, ring-, and sector-based analyses were employed to analyze the urban expansion. The results showed that there was a large transition from agricultural and unused lands to other uses. This resulted in an expansion of developed areas, recreational regions, mining sites, and water surfaces, a drastic decline of agricultural land for annual crops, and a fluctuation of perennial cropland and unused land. The study also indicated that the urban area has expanded 65 times within 25 years at an increasing rate. The AER and ECR were uneven between subregions, and there was a gradual expansion and shift from south to north of the province. The factors affecting the changes comprise natural conditions, development histories, policies and practices for urbanization, industrialization, and agricultural development, and product price fluctuations in the market. Practical lessons learned from this study could be useful for land planning and policymaking in other localities.

Keywords: Land-use change; urban expansion; change detection; spatiotemporal analysis; remote sensing

4.1. Introduction

Land-use change has various impacts on the environment and human life, such as run-off characteristics (Sajikumar and Remya 2015), landscape pattern (Zhang et al. 2010; Dadashpoor et al. 2019), land surface temperature (Zhang and Sun 2019), soil erosion (Nampak et al. 2018), as well as biodiversity and ecosystem services (Tolessa et al. 2017; Trisurat et al. 2019). Therefore, studies on land-use change are crucial to resource and environmental monitoring as well as land management policymaking (Nampak et al. 2018). Land-use changes are caused by both natural and anthropogenic factors (Serra et al. 2008; Msofe et al. 2019). In terms of anthropogenic factors, activities and policies relating to urban expansion, industrialization, agricultural development, and exploitation of natural resources strongly influence land-use change. Among them, urbanization and industrialization often lead to rapid, strong, and one-way transformation, especially in developing countries (Pham and Yamaguchi 2011; Kantakumar et al. 2016; Rimal et al. 2017; Fenta et al. 2017; Andrade-Núñez and Aide 2018; Cao et al. 2019; Sumari et al. 2020). However, the availability of accurate information on spatiotemporal land-use changes, urbanization status, urbanization rates, and their driving factors in localities is often untimely even though it is essential (Kantakumar et al. 2016).

Remote sensing (RS) is a reliable tool for land cover and land-use monitoring (Toure et al. 2018). RS databases are increasingly diverse in quantity and quality, meeting different needs. With easy access and acquisition of images, such as MODIS, Landsat, and Sentinel, research related to the interpretation of RS imagery has become proactive and cost effective. Moreover, the development of image processing and classification techniques has increasingly improved the accuracy of results (Lu et al. 2011; Shao and Lunetta 2012; Noi and Kappas 2017; Toure et al. 2018; Quan et al. 2020). The combination of RS and spatial analysis techniques in geographic information systems allows researchers to detect land cover and land-use change more easily and timely. This has been confirmed in many studies in the literature on a local (Wu et al. 2006; Rawat and Kumar 2015; Tadese et al. 2020), national (Sánchez-Cuervo et al. 2012; Schoeman et al. 2013; Xu et al. 2020), continental (Mertes et al. 2015; Netzel and Stepinski 2015) and global scale (Giri et al. 2013; X. Li et al. 2017).

Binh Duong province is a province located in the Southern Key Economic Zone of Vietnam. Over the past 25 years, Binh Duong has emerged as a typical area of rapid urbanization and industrialization. As a consequence, the land-use change took place dramatically. However, a study on spatiotemporal land-use change and urban expansion in Binh Duong is still a gap. Thus, it is necessary to study these issues in this area. Such study helps explore not only the pattern of land-use change and urban expansion but also the factors influencing these processes. From there, some practical experience can be learned for land-use planning and policymaking in other areas not only in Vietnam but also in other countries.

Therefore, this study aims to (1) explore the spatiotemporal dynamics of land-use in Binh Duong province from 1995 to 2020, (2) analyze the urban expansion and its orientation over the past 25 years, and (3) analyze the factors affecting the land-use change and urban expansion of Binh Duong province.

The rest of this paper is organized as follows. Section 4.2 introduces the study area. Section 4.3 describes the data and methods used. The results are reported in Section 4.4. In Section 4.5, a discussion and recommendation are given. Finally, the conclusions are presented in Section 4.6.

4.2. Study area

Binh Duong province is located in the Southeast region of Vietnam, covering approximately 2,694.64 km^2, with a total population of about 2.5 million as of 2019 (Binh Duong Statistical Office 2020). Administratively, as of 2020, the province was divided into five urban districts (also known as cities and towns) and four rural districts (Figure 4.1). Thu Dau Mot city is the administrative–economic–cultural centre of Binh Duong province.

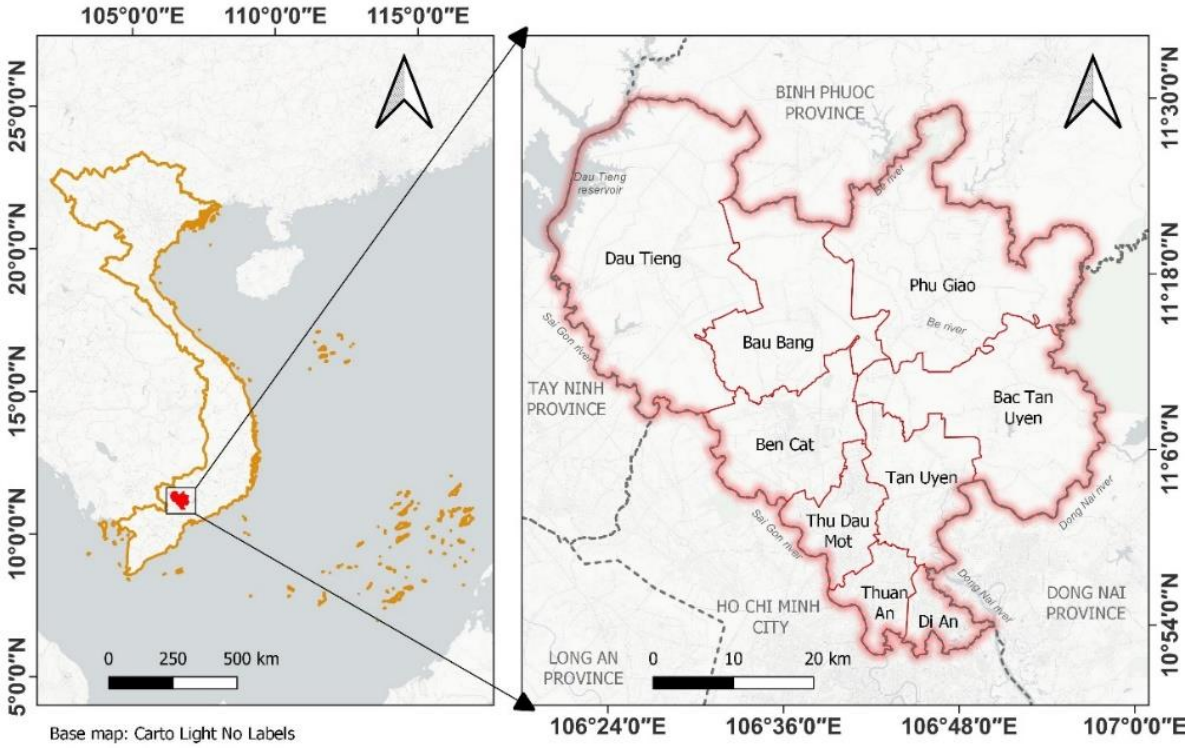


Figure 4.1. Study area.

Since its reestablishment in 1997, the urbanization and industrialization process of the province has been extremely rapid. In 1995, the urbanization rate, which was calculated as a percentage of the urban population per total population, accounted for only 17.51%; in 2019, the rate reached 79.86% (General Statistics Office of Vietnam 2020). The first industrial park, i.e. Song Than 1, was established in 1995. As of 2019, Binh Duong has 29 industrial parks and 12 industrial clusters, with an average occupancy rate of over 70% in which more than 90% of many of them have been filled. Binh Duong is currently considered the 'industrial capital' of Vietnam. The development of industry has considerably contributed to the economic development of the province. The gross regional domestic product at current prices increased from VND 3,915 billion in 1997 (industry and construction accounted for 50.4%) to VND 48,761 billion in 2010 (industry and construction accounted for 63%) and 360,797 billion in 2019 (industry and construction accounted for 66.77%) (Binh Duong Statistical Office 2016; Binh Duong Statistical Office 2020). Owing to the expansion of the urban and industrial areas as well as other human activities, the land cover, and land-use in the province have significantly changed over the past 25 years.

4.3. Material and methods

The overall workflow consisted of selection of time points and satellite images, preprocessing, generation of land use maps according to the method of Bui and Mucsi (2021), accuracy assessment, change detection, and urban expansion analysis. The classification steps for generating land-use maps were performed in ERDAS IMAGINE 2020 and R software; meanwhile, other analysis steps were performed in QGIS 3.10

software. The workflow is illustrated in Figure 4.2 and some highlights of the process followed are described below.

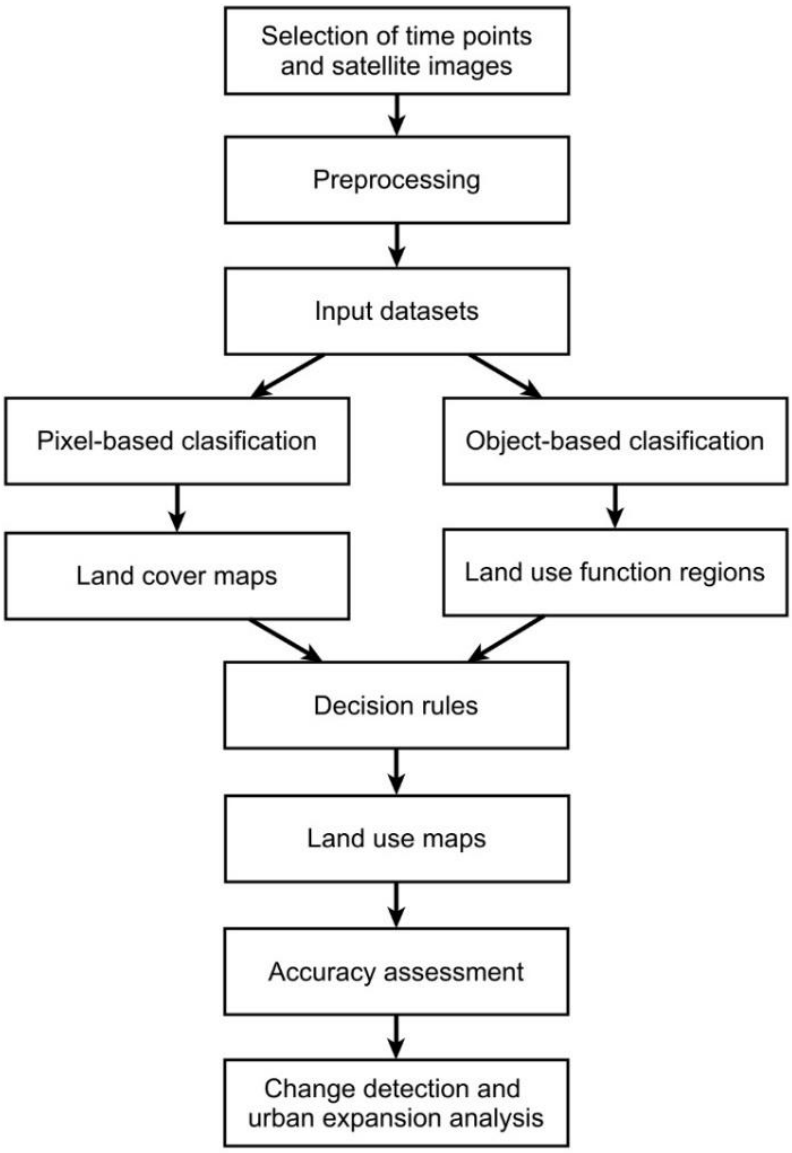


Figure 4.2. Overall workflow.

4.3.1. Selection of time points and satellite images

The Landsat level-2 surface reflectance images (projection: WGS 84/UTM Zone 48 N, path/row: 125/52) were employed in this study. The images were ordered and downloaded from the United States Geological Survey website (via the link: https://earthexplorer.usgs.gov/). A survey of the availability and quality of Landsat images from 1994 to 2020 (hereinafter, referenced years) in the study area was conducted to select suitable images and time points for the classification process. To generate the land-use map of each year, the method required at least two cloud-free images as the input for classification. Therefore, the criteria for selecting the time points and images included the following: (1) The images belong to Landsat-5, −7, or −8 sensors. (2) There is no error of scan-line corrector for the Landsat-7 images. (3) It is possible to collect or mosaic to create two cloud-free images in the same year or in two

consecutive years. As a result, six referenced years were identified to create land-use maps: 1995, 2001, 2005, 2010, 2015, and 2020. The images used at each referenced year are listed in Table 4.1. No Landsat-7 image was selected because the three mentioned criteria were not met. In addition, except for 2001, most selected images were in the dry season.

Table 4.1. Summary of Landsat images used.

Time point	Acquired date	Sensor
1995	14-Nov-94	Landsat-5
	02-Feb-95	Landsat-5
2001	27-Sep-00	Landsat-5
	06-Nov-00	Landsat-5
	14-Nov-00	Landsat-5
	15-Apr-01	Landsat-5
	09-May-01	Landsat-5
2005	11-Dec-04	Landsat-5
	11-Feb-04	Landsat-5
	13-Feb-05	Landsat-5
2010	09-Dec-09	Landsat-5
	11-Feb-10	Landsat-5
2015	24-Jan-15	Landsat-8
	29-Mar-15	Landsat-8
	09-Feb-15	Landsat-8
2020	06-Jan-20	Landsat-8
	23-Feb-20	Landsat-8
	07-Feb-20	Landsat-8

4.3.2. Preprocessing

The Landsat images collected were level 2 images, which have been geocorrected, projected to WGS84 Zone 48 N, and converted to surface reflectance, so they were ready for use. In this study, bands 2, 3, 4, 5, 6, and 7 for Landsat 8 images and bands 1, 2, 3, 4, 5, and 7 for Landsat 5 images were used. The preprocessing steps included masking cloud, mosaicking, subsetting, and stacking. The administrative boundary data downloaded from the Database of Global Administrative Areas project website (via the link: https://gadm.org/) were used for the subsetting step. Each year, two subsetted images were stacked to create a multitemporal image with 12 bands. Then, these multitemporal images were used as input for the classification.

4.3.3. Generation of land-use maps

In this study, we employed a classification scheme and method proposed by Bui and Mucsi (2021) to generate land-use maps of Binh Duong province from 1995 to 2020. This is a hybrid approach, combining pixel-based and object-based classification with spatial analysis and decision rules, to generate a land-use map from multi-temporal Landsat images. This method has proven to be effective in generating highly accurate land-use maps in the study area.

For classification scheme, the main land-use types in the study can be classified into eight classes: unused land, industry and commerce, recreation and green space, mixed residence, mining sites, agriculture with annual plants, agriculture with perennial plants, and water surface. For the procedure, there are three main steps to produce a land-use map including pixel-based classification to generate land cover map (Step 1), object-based classification and spatial analysis to extract land-use function regions (Step 2), and the combination of the land cover map and function region in a set of decision rules to produce the final land-use map (Step 3). A detailed description can be found in the mentioned study.

During Steps 1 and 2, the random forest classifier was applied to both pixel- and objected-based classifications. The training data were collected based on the field survey between January and February 2020, Google Earth history images, and our personal experiences. For consistency, the parameters of mtry and ntree were set to default values at all classification rounds, i.e. the mtry was equal to the square root of the total number of variables, and the ntree was 500 trees. Further, a 3 × 3 majority filter was applied to reduce the 'salt-and-pepper' noise instead of the clump and eliminate in Step 1. Other parameters in this stage were set similar to the study of Bui and Mucsi (2021).

In addition, to increase the accuracy of results, several additional works were performed manually as follows: (1) Some land-use function regions were added manually after Step 2. They included the regions that could not be automatically formed due to some reasons, such as only accounting for a very small area insufficient for training and classification (e.g. industry and commerce regions in 1995, golf courses in 1995, 2001, 2005), being excluded after spatial analysis within Step 2 (e.g. small and discrete patches of recreation area and mining site), or being entertainment complexes (e.g. the Dai Nam wonderland and new city park). (2) A 30-m buffer was applied to the mining function regions in Step 2 to make their boundary precise. (3) In the land cover maps generated after Step 1, some easily observable areas where barren lands were misclassified into impervious surfaces were corrected. Consequently, six land-use maps for the six referenced years were developed.

4.3.4. Accuracy assessment

The accuracy of the extracted land-use maps was assessed based on the overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA). To collect the validation data, we applied the stratified random sampling method with a compromise approach proposed by Congalton and Green (2019). In each map, 1,000 points were generated randomly. Among them, the minimum points for each class were set to 50 points, and the other points were allocated proportionally to the area of each class. The advantage of this method is that it can ensure that the distribution of validation points was proportional to the area of each land-use type, whereas the PA and UA of the rare classes were statistically significant (Congalton and Green 2019). The generated points were

then visually interpreted and labelled based on the Landsat images, Google Earth history images, and our personal experiences. The allocation of validation points in each map is listed in Table 4.2.

Table 4.2. Allocation of validation points (unit: points).

Year Class	1995	2001	2005	2010	2015	2020
Unused land	90	66	89	106	111	105
Industry & Commerce	50	51	53	59	65	75
Recreation & Green space	50	50	50	50	52	53
Mixed residence	51	55	62	70	80	100
Mining site	50	50	51	50	50	50
Agriculture with annual plants	171	158	129	83	72	70
Agriculture with perennial plants	474	504	502	514	502	480
Water surface	64	66	64	68	68	67
TOTAL	1000	1000	1000	1000	1000	1000

There is no common threshold for the acceptable accuracy rate in the literature because it depends on the purpose of the research (Congalton and Green 2019). In this study, we adopted the values given by Thomlinson et al. (1999) where OA was at least equal to 85% and PA and UA of classes were equal to or greater than 70%.

4.3.5. Change detection and urban sprawl analysis

Besides statistics on the area accounted for by each land-use type in each year, transition matrices were employed to assess the detailed 'from-to' change between land-use classes in different years. It is a common approach in land-use change studies to compare maps between different time points (B. Zhang et al. 2017). The rows of the transition matrix represent the land-use classes of the former time point (T1), whereas the columns represent the ones of the later time point (T2). The main diagonal elements indicate the landscape area that shows the persistence of class i. Off diagonal entries indicate a transition from class i in T1 to a different class j in T2.

Urban sprawl is defined as 'the spreading of urban developments (such as houses and shopping centres) on undeveloped land near a city' (Merriam-Webster 2022). Although some studies distinguished the terms 'urban expansion' and 'urban sprawl' (Amponsah et al. 2022; Pratama et al. 2022), most studies in the literature used the two terms with the same meaning. In this paper, we also used them interchangeably. To analyze the spatial and temporal processes of the urban expansion, the land-use maps were first reclassified into two: urban and nonurban. In this study, the urban area was considered as both mixed residential regions as well as industrial and commercial zones; meanwhile, other land-use types belonged to the nonurban area. Afterward, the annual expansion rate (AER) and expansion contribution rate (ECR) were calculated to measure the characteristics of the urban sprawl of each administrative unit of Binh Duong province. The metrics were defined by Kantakumar et al. (2016) as follows:

$$AER = \frac{B_{(i,t_2)} - B_{(i,t_1)}}{t_2 - t_1} \quad (1)$$

$$ECR = \frac{B_{(i,t_2)} - B_{(i,t_1)}}{B_{t_2} - B_{t_1}} \times 100\% \quad (2)$$

where *AER* is the annual urban expansion rate during the period (in $km^2.year^{-1}$), *ECR* is the percentage share of urban expansion of an individual administrative unit in the study area (in percent); $B_{(i,tj)}$ is the built-up area within administrative unit *i* at time t_j (in km^2), and B_{tj} is the total built-up area in the study area at time t_j.

In addition, we employed the ring- and sector-based analyses. A ring-based analysis is a spatial analysis based on concentric circles separated by a certain distance, meanwhile, a sector-based analysis is based on sectors (i.e. fan-shaped areas) that have the same central vertex but in different orientations. These approaches have proven their effectiveness in exploring the spatial distribution of urban areas in terms of distance and orientation relative to a predefined urban centre point (Yin et al. 2011; Jiao 2015; Peng et al. 2015; Rimal et al. 2017; Acheampong et al. 2018; Cao et al. 2019). Analyzing urban expansion in terms of distance and orientation is important because it can reveal the pattern of urbanization, and it also can reflect the impact of land-use planning and policies on the urbanization process. In these analyses, defining a place as the urban centre point is an essential requirement. We considered the Thu Dau Mot market as the urban centre. From this point, 32 buffer zones with a distance of 2 km and 16 sector fans with an angle of 22.5° were generated to measure the distance and orientation, respectively (Figure 4.3). After that, the buffer zones and sectors were overlaid with land-use maps. The urban area was then summarized in each buffer zone and sector.

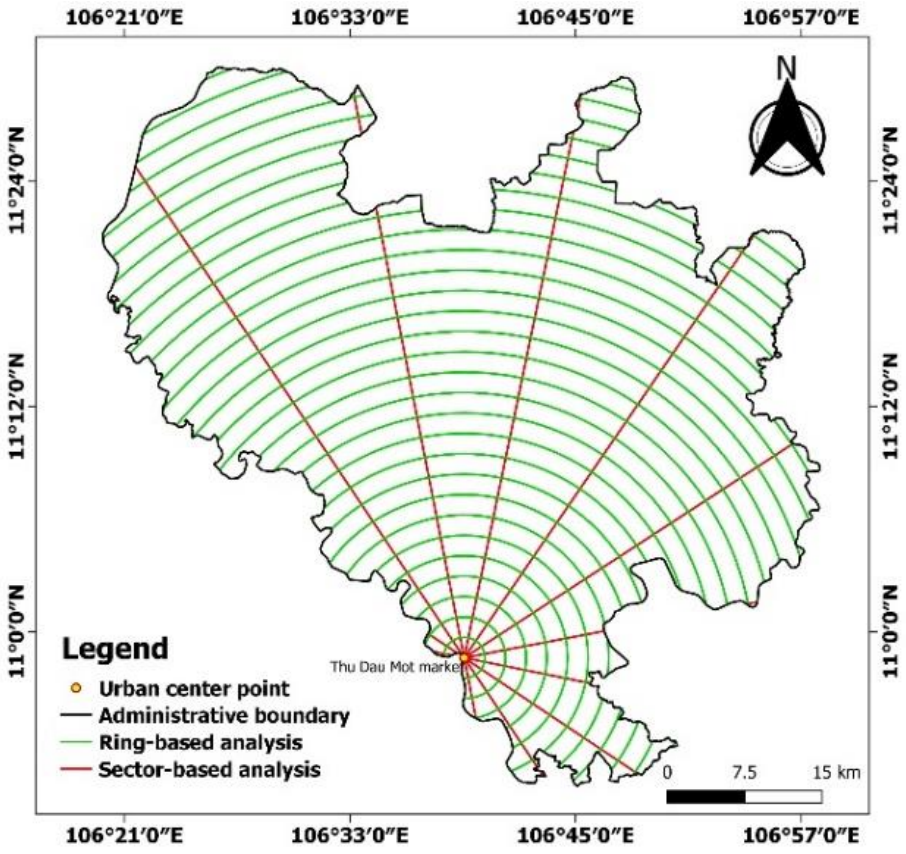


Figure 4.3. Ring- and sector-based analyses.

4.4. Results

4.4.1. Accuracy of extracted land-use maps

The OA of the extracted land-use maps in 1995, 2001, 2005, 2010, 2015, and 2020 were 89.2%, 88.9%, 89.6%, 90.8%, 93.0%, and 90.1%, respectively. The PA of the maps ranged from 70.8% to 100%, whereas the UA ranged from 70.9% to 100% (Table 4.3). These results showed that the maps are suitable for land-use change analyses.

Table 4.3. Accuracy of extracted land-use maps.

Year	1995		2001		2005		2010		2015		2020	
Class	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
UL	70.8	75.6	80.0	78.8	73.3	83.1	82.7	81.1	87.5	88.3	80.2	88.6
IC	88.2	90.0	71.0	86.3	78.8	98.1	82.8	89.8	87.5	86.2	87.7	85.3
RG	95.9	94.0	97.3	72.0	91.7	88.0	95.9	94.0	98.1	98.1	90.2	86.8
MR	81.4	94.1	73.6	70.9	90.7	79.0	75.0	81.4	80.5	87.5	82.4	75.0
MS	92.6	100	94.2	98.0	96.2	98.0	96.0	96.0	100	98.0	98.0	96.0
AA	83.5	80.1	83.7	81.0	79.4	77.5	84.0	75.9	81.7	80.6	77.8	80.0
AP	93.7	91.4	93.9	94.2	95.1	92.4	95.3	94.9	96.8	96.2	94.0	94.6
WS	98.5	100	91.7	100	96.9	98.4	94.3	97.1	98.5	95.6	100	97.0
	OA = 89.2%		OA = 88.9%		OA = 89.6%		OA = 90.8%		OA = 93.0%		OA = 90.1%	

Note: OA = overall accuracy; PA = producer's accuracy; UA = user's accuracy; UL = Unused land; IC = Industry & Commerce; RG = Recreation & Green space; MR = Mixed residence; MS = Mining site; AA = Agriculture with annual plants; AP = Agriculture with perennial plants; WS = Water surface.

4.4.2. Land-use dynamics

The land-use maps in Binh Duong province at six time points are illustrated in Figure 4.4, and the detailed dynamics of each class from 1995 to 2020 are presented in Figure 4.5 and Table 4.4. In addition, a transition matrix of land-use types between 1995 and 2020 was developed (Table 4.5).

In terms of spatial distribution, mixed residential areas, industrial and commercial zones, recreation areas, and quarries were concentrated in the south of the province, and the north was mainly agricultural land for perennial plants; meanwhile, the fields of annual plants were mainly distributed along rivers, canals, and streams of the Sai Gon-Dong Nai river system. Agriculture with perennial plants was always the dominant class in the study area, with a proportion of 70.5% or more. In general, it is clear that from 1995 to 2020, a large amount of the agricultural land and unused land was converted to other uses. This conversion led to a strong change in the proportion between each land-use class over the past 25 years.

Between 1995 and 2020, there was a dramatic continuous upward trend in the mixed residential areas as well as industrial and commercial zones. In 1995, the area of mixed residence only accounted for 4.9 km^2. It rose to 21.9 km^2 in 2001, 133.5 km^2 in 2015, and 222.8 km^2 in 2020. Similarly, from occupying an area of only 0.2 km^2 in 1995,

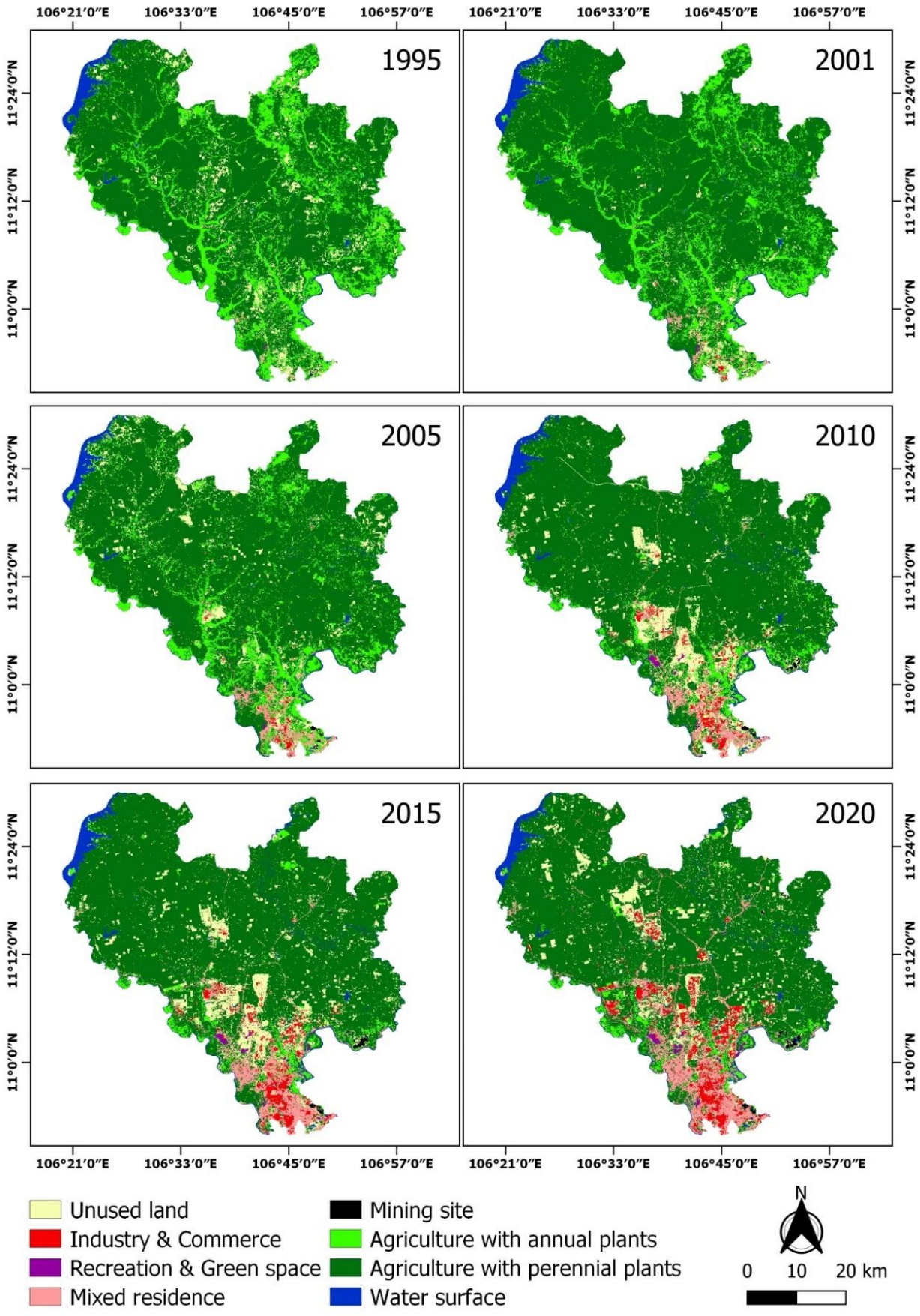


Figure 4.4. Land-use maps of Binh Duong province in the referenced years.

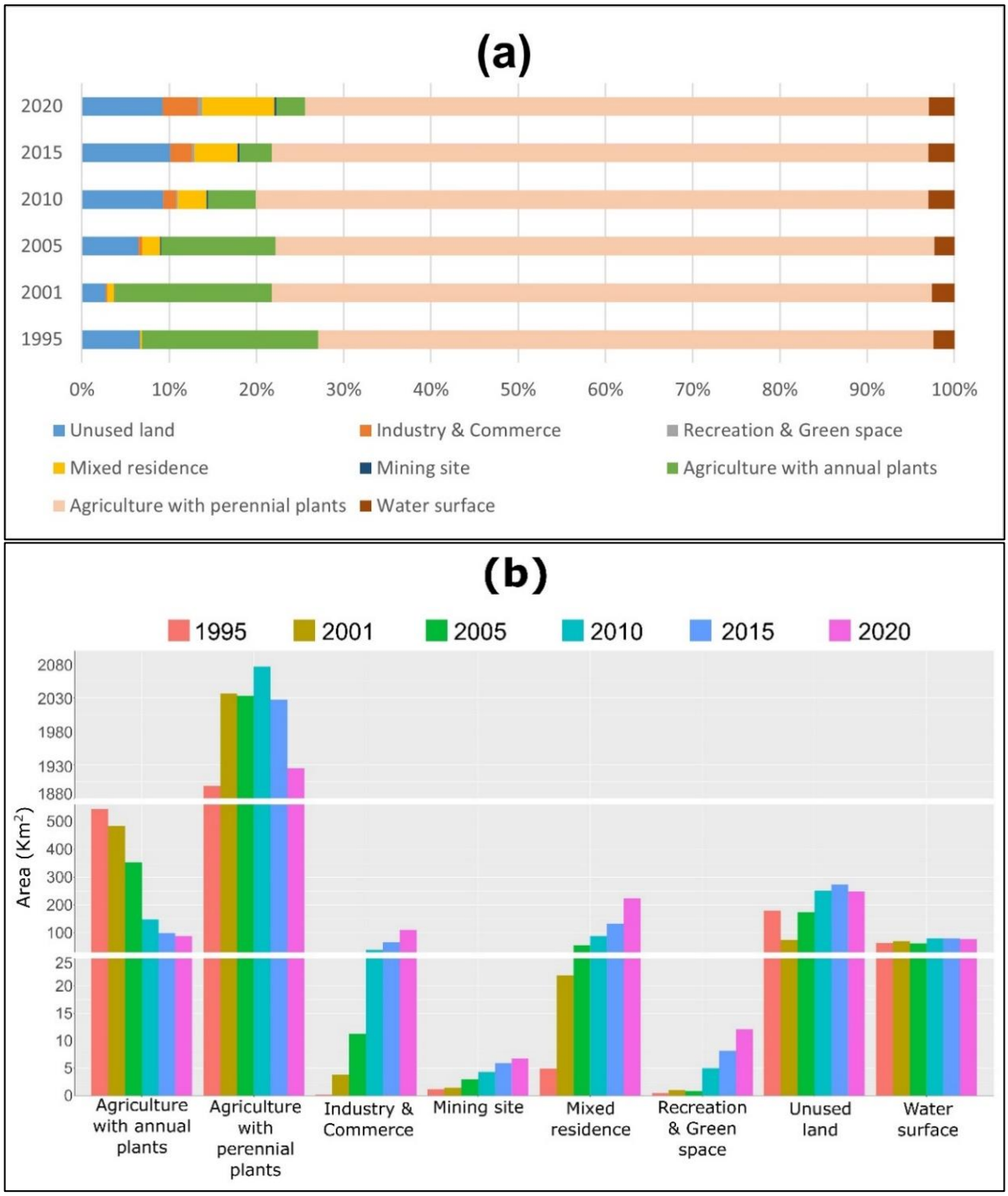


Figure 4.5. Dynamics of land-use in (a) proportion and (b) area.

Table 4.4. The annual change rate of each land-use type in each period (in $km^2.year^{-1}$).

Land-use type	1995–2001	2001–2005	2005–2010	2010–2015	2015–2020
Unused land	−17.6	25.0	15.4	4.4	−4.9
Industry & Commerce	0.6	1.9	5.7	5.2	8.9
Recreation & Green space	0.1	−0.1	0.8	0.6	0.8
Mixed residence	2.8	8.5	6.6	8.9	17.9
Mining site	0.0	0.4	0.2	0.3	0.2
Agriculture with annual plants	−10.0	−32.9	−40.9	−9.7	−2.2
Agriculture with perennial plants	23.0	−0.9	8.7	−9.9	−20.3
Water surface	1.0	−2.0	3.5	0.1	−0.3

Table 4.5. Transition between land-use classes from 1995 to 2020 (in km2).

	Transition								Summary		
2020 1995	UL	IC	RG	MR	MS	AA	AP	WS	Loss	Gain	Total loss/gain
UL	26.1	21.1	1.3	22.4	0.4	2.6	103.8	1.1	152.7	222.2	69.5
IC	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.1	109.9	109.7
RG	0.0	0.0	0.3	0.1	0.0	0.0	0.0	0.0	0.1	11.8	11.7
MR	0.1	0.6	0.0	4.0	0.0	0.0	0.1	0.0	0.9	218.8	217.9
MS	0.1	0.0	0.0	0.1	0.3	0.1	0.3	0.3	0.9	6.5	5.6
AA	40.1	17.9	2.6	48.1	3.4	59.1	357.9	14.7	484.8	29.5	−455.3
AP	181.5	69.9	7.9	147.8	2.7	25.9	1459.0	4.1	439.8	466.5	26.7
WS	0.3	0.2	0.1	0.2	0.0	1.0	4.3	58.1	6.1	20.2	14.1

Note: UL = Unused land; IC = Industry & Commerce; RG = Recreation & Green space; MR = Mixed residence; MS = Mining site; AA = Agriculture with annual plants; AP = Agriculture with perennial plants; WS = Water surface.

the industrial and commercial zones expanded to 11.3 km^2 in 2005 and 110.0 km^2 in 2020. As a result, the mixed residential areas as well as industrial and commercial zones expanded by a total of 217.9 and 109.7 km^2, respectively, within the 25 years. These expansions took about 217.7, 66.0, and 43.5 km^2 from perennial cropland, annual cropland, and unused land, respectively.

Although not as dramatic as the two built-up classes, the recreation, green space, and mining sites also had an uptrend in its area. From 1995 to 2005, the area of recreation and green space increased slightly from 0.5 to 0.8 km^2. After that, it increased rapidly to 12.1 km^2 in 2020. To tradeoff for this expansion, the areas of perennial plants, annual crops, and unused land reduced a total of 7.9, 2.6, and 1.3 km^2, respectively. Meanwhile, the mining sites experienced continuous growth from 1.2 km^2 in 1995 to 6.8 km^2 in 2020 for the 25 years. This expansion took a total of 6.5 km^2 from agricultural and unused land, whereas, since 2005, some quarries in Di An district have been closed permanently and converted to other land uses.

During the study period, 6.1 km^2 of water surface were occupied for other purposes, mainly for agricultural activities. Meanwhile, 20.2 km^2 of water surface were added, mainly from agricultural land, unused land, and mining sites. Consequently, although there was a fluctuation over time, the water surface area increased from 64.2 km^2 in 1995 to 78.3 km^2 in 2020.

Contrariwise, although supplemented by the conversion from other types, the area of agriculture with annual plants continuously decreased significantly over the 25 years. The area of annual plants reduced by 83.7%, from 543.9 km^2 in 1995 to 88.6 km^2 in 2020. From occupying more than 20% of the province's area, its proportion reduced to only 3.3%. In which, the nine-year period from 2001 to 2010 was the period that experienced the most dramatic decline. Meanwhile, the change in agricultural land for perennial plants can be categorized into two main trends: upward from 1995 to 2010 and downward from 2010 to 2020. The area increased from 1,898.8 km^2 in 1995 to

2,076.5 km^2 in 2010 before reducing to 1,925.5 km^2 in 2020. In addition to the conversion to other types of land-use, there was a transition between annual and perennial croplands over the studied period. About 357.9 km^2 of annual cropland was converted into perennial cropland from 1995 to 2020. Besides, about 25.9 km^2 of perennial croplands have been converted into annual croplands.

In terms of unused land, there was a fluctuation over the 25 years. Its area decreased from 1995 to 2001 and from 2015 to 2020; meanwhile, it increased from 2001 to 2015. In the 25 years, the total unused land area converted to other purposes was 152.7 km^2, mainly for re-cultivation and construction. In addition, 222.2 km^2 of other classes, mainly from agricultural land, were temporarily converted to unused land. Therefore, from accounting for 178.8 km^2 in 1995, Binh Duong province had approximately 248.3 km^2 of temporarily unused land in 2020.

4.4.3. Urban expansion analysis

The pattern of urban expansion in Binh Duong province is illustrated in Figure 4.6. The study result revealed that the developed area in Binh Duong province expanded rapidly, nearly 65 times, from 5.1 km^2 in 1995 to 332.8 km^2 in 2020. In addition, the gradual increase in the slope of segments in the line graph in Figure 4.6 indicates that the urban sprawl rate of the following period was always higher than that of the previous period. The AER increased from 3.4 $km^2.year^{-1}$ in 1995–2001 to 6.9, 10.2, 11.8, and 22.3 $km^2.year^{-1}$ in the periods of 2001–2005, 2005–2010, 2010–2015, and 2015–2020, respectively.

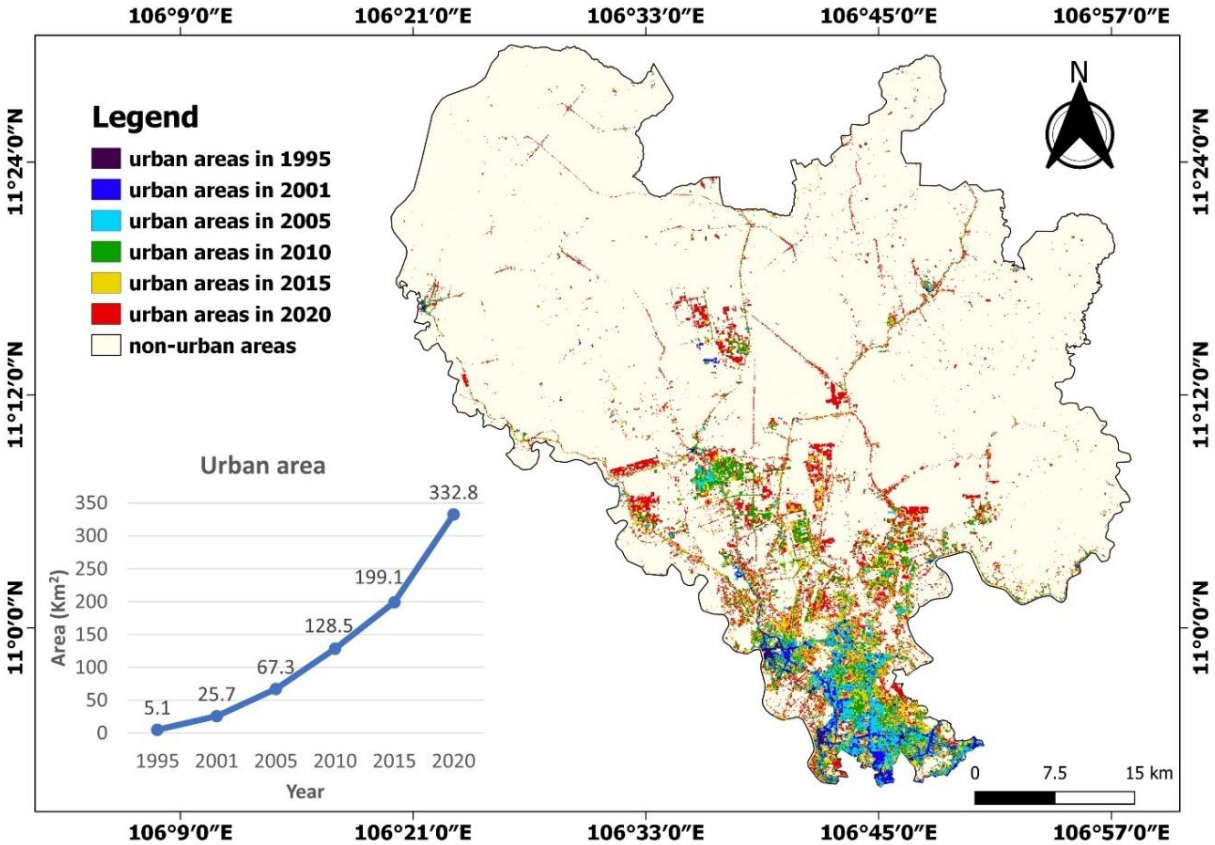


Figure 4.6. Urban expansion in Binh Duong province from 1995 to 2020.

However, the analysis result showed that the growth rate and contribution to the urban expansion among districts were not equal (Table 4.6) in each period. For urban districts, Di An and Thuan An districts had the highest two urbanization rates in the province from 1995 to 2005; their rates reached 3.1 and 4.1 $km^2.year^{-1}$, respectively. Thus, they were the main contributors to the urban expansion of the entire province in that period, with the ECR of each district ranging from about 30%–40%. However, in later periods, the urbanization rate of these two districts tended to decrease. Meanwhile, the AER of Tan Uyen, Ben Cat, and Thu Dau Mot gradually increased, and they shared, in turn, the top positions of main contributors from 2005 to 2020. In particular, the period of 2015–2020 witnessed an extremely rapid urban expansion of Tan Uyen and Ben Cat at an AER of approximately 6.0 $km^2.year^{-1}$, whereas the AERs of Di An and Thuan An decreased to within 1.0–2.0 $km^2.year^{-1}$. For rural districts, urban areas have only expanded significantly since 2015. The most significant was Bau Bang achieving an AER of 2.8 $km^2.year^{-1}$ and an ECR of 10.5%, whereas other districts had an AER between 1.5 and 2.2 $km^2.year^{-1}$.

Table 4.6. Annual expansion rate (AER in $km^2.y^{-1}$) and expansion contribution rate (ECR in percent) of districts.

District	Type	1995–2001		2001–2005		2005–2010		2010–2015		2015–2020	
		AER	ECR	AER	ECR	AER	ECR	AER	ECR	AER	ECR
Thu Dau Mot	Urban	0.6	16.2	0.9	8.6	2.5	20.2	2.9	20.8	3.5	13.3
Thuan An	Urban	1.1	33.0	4.1	39.2	2.1	16.8	3.0	21.4	2.0	7.5
Di An	Urban	1.3	38.1	3.1	29.9	1.6	13.4	2.3	16.0	1.0	3.7
Tan Uyen	Urban	0.2	5.9	1.2	11.1	1.7	14.3	3.6	25.3	6.0	22.4
Ben Cat	Urban	0.1	3.3	0.8	7.9	2.7	21.8	1.4	9.8	5.7	21.5
Phu Giao	Rural	0.0	0.2	0.2	1.5	0.3	2.2	0.3	2.2	2.2	8.1
Dau Tieng	Rural	0.0	0.7	0.1	0.7	0.3	2.8	0.1	0.5	1.5	5.5
Bau Bang	Rural	0.1	2.3	0.1	0.5	0.5	4.4	0.1	0.8	2.8	10.5
Bac Tan Uyen	Rural	0.0	0.3	0.1	0.7	0.5	4.1	0.5	3.3	2.0	7.6

The results of the sector- and ring-based analysis were illustrated in Figures 4.7 and 4.8, respectively. Within 25 years, there has been a shift in the direction of urban expansion in the study area (Figure 4.7). From 1995 to 2015, the developed areas mainly expanded in the Southeast and East–Southeast directions from the urban centre point. However, since 2005, besides these two main directions, the urban area has been gradually expanded in the north and east directions, including Northwest, North–Northwest, North, North–Northeast, Northeast, East–Northeast, and East. Besides, from 2015 to 2020, the expansion to the north direction was stronger, whereas the expansion in the Southeast, East–Southeast, and East directions decreased significantly. Because this study was based on the administrative boundary, and Thu Dau Mot market—the predefined urban center—was located on the east bank of the Saigon River, whereas the west bank was the Ho Chi Minh City, there was almost no urban development in the directions from South–Southwest to West–Northwest (in the clockwise direction).

In addition, urbanization occurred strongly within 22 km from the urban centre over the 25 years (Figure 4.8). This zone encompasses the centre area of Thu Dau Mot city within 4 km and the Thuan An, Di An, Tan Uyen, and Ben Cat districts within a distance of 6–22 km from the urban centre. The outward movement of peak value from the distance of 10 km to 14 km indicated that the urban areas expanded strongly in the direction of gradually moving away from the urban centre. Meanwhile, from the distance of 22 km onward, urbanization seemed to have only begun to accelerate since 2015, mainly in some central areas of the rural districts, yielding small peaks at distances of 26, 30, 40, and 46 km from the urban centre.

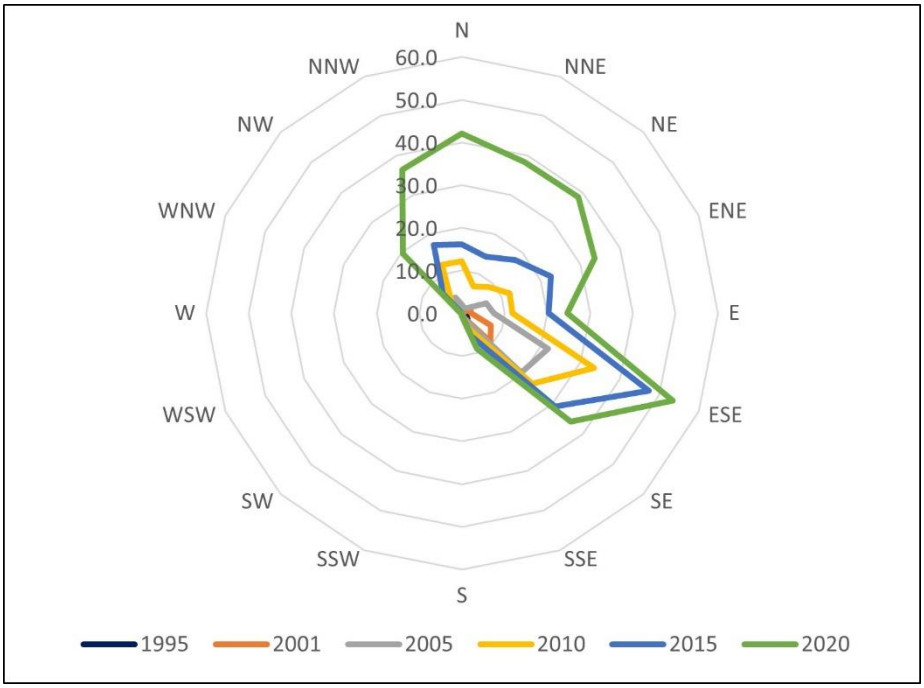


Figure 4.7. Spatial orientation of urban area from 1995 to 2020 (Units: km^2).

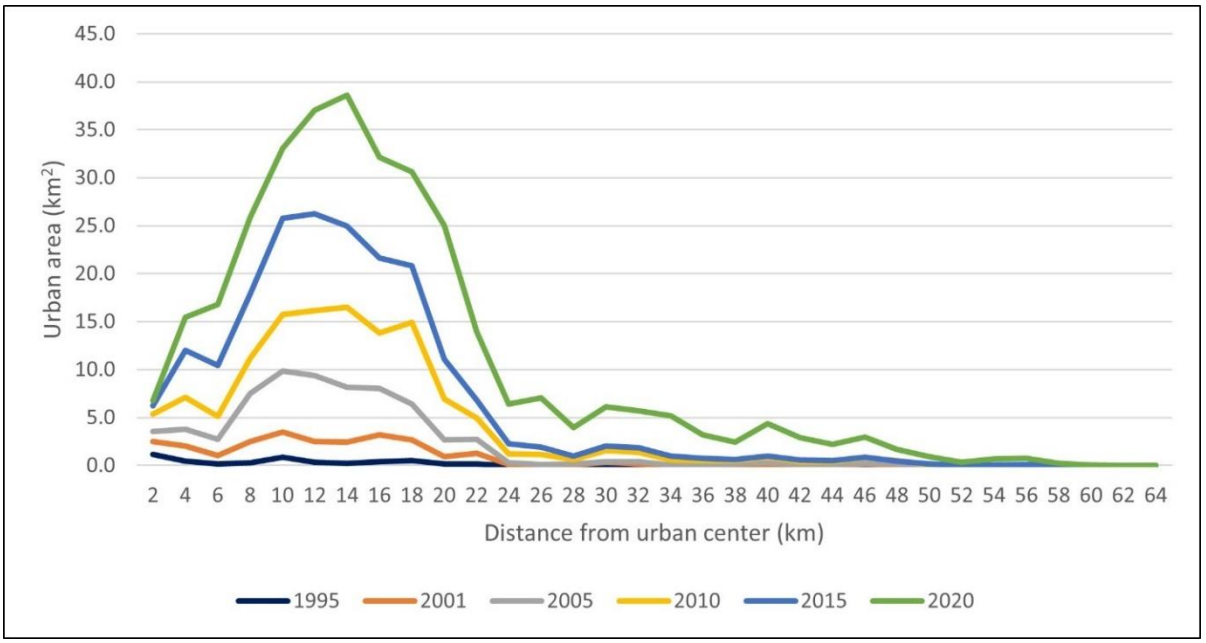


Figure 4.8. Variation in urban area by distance from urban centre from 1995 to 2020.

Overall, the results of the district-, sector-, and ring-based analyses clearly showed that urbanization gradually expanded and shifted from the south to the north of the province during the 25 years. Before 2005, urban expansion mainly took place in Thuan An, Di An, and southern Thu Dau Mot. In the periods after 2005, the urban expansion rate in the southern districts decreased because the industrial activities in these localities had stabilized, the districts were almost filled by mixed residential areas and industrial zones. Afterward, industrial zones and mixed residential areas rapidly expanded to Tan Uyen, Ben Cat, and northern Thu Dau Mot as well as spread to rural districts at an increasingly rapid rate.

4.5. Discussions

4.5.1. Factors affecting land-use change from 1995 to 2020

In summary, the common trends in land-use change over the 25 years in Binh Duong province is the conversion from agricultural land and unused land to other types of land-use. These changes are attributable to urbanization and industrialization, agricultural development policies and practices, natural conditions, and rubber price fluctuation. This section mainly discusses the factors causing the transition of agricultural land, unused land, water surface, mining activity, and recreational areas. The factors leading to rapid urban expansion in the province are discussed in Section 4.5.2.

Urbanization and industrialization are the primary causes of change. They lead not only to the inevitable expansion of built-up areas but also to other changes. These processes increase the demand for building materials for housing, industrial infrastructure, transportation systems, etc. (Schiller et al. 2020). Products of quarries in Binh Duong province serve the needs of not only the province but also neighbouring provinces. Consequently, although the mining activity may cause damages to the landscape, ecosystem, environmental quality, and health of people living around these areas (Bui et al. 2020; Vandana et al. 2020), the expansion of stone quarries occurred continuously during the 25 years. In addition, urban development includes not only the construction of residential, commercial, and industrial regions but also the establishment of amenity, relaxation, and leisure areas (Ty et al. 2014) as well as green space. In the period 1995–2005, there was only one recreation area, i.e. the Song Be golf course, and a few industrial parks. From 2005, many other recreation areas were formed, including the Dai Nam wonderland (2008), new city park (2009), Twin Dove golf course (2010), Harmonie golf course (2018), and Mekong golf course (2019). In addition, many new industrial parks were established and put into operation. The emergence of these areas caused the area of recreation and green space to increase rapidly. Overall, it can be observed that the majority area of the mentioned transition during the process of urbanization and industrialization is from agricultural land and unused land.

The changes in water surface area mainly came from the operation of the Dau Tieng reservoir, agricultural and aquacultural activities, hydrological regime, and the

accumulation of rainfall water in exquarries. In addition, the construction of the Phuoc Hoa irrigation dam on Be River in the northern part of the province contributed to an increase in the water surface area since the 2010s. Phuoc Hoa Dam locates between Binh Duong and Binh Phuoc provinces and is connected to the Dau Tieng reservoir. This project has significant roles in supplying domestic and industrial uses and irrigation of more than 58,360 ha of agricultural land in Binh Duong, Binh Phuoc, Ho Chi Minh City, Tay Ninh, and Long An. Besides, the dam helps improve the environment and water quality in the downstream areas of the Sai Gon and Vam Co Dong rivers (Vu Van and Nguyen Hai 2015).

In addition to the conversion to other types of land-use, there was a transition between annual and perennial croplands over the studied period. In fact, the conversion from annual cropland into perennial cropland was mainly from low-yielding crops to higher-value perennial trees, such as rubber, pepper, cashew, and fruit trees, which were more favourable with the climate and soil of Binh Duong. Especially before 2011, owing to the continuous growth in demand for rubber leading to its price increase (Fox and Castella 2013; Hurni and Fox 2018), the conversion to rubber plantations was common in the northern part of the province. The expansion of rubber farms was also a general trend occurring in many other localities in Vietnam and Southeast Asia during this period. The rubber plantation area in Vietnam increased from 3,950 km^2 in 1999 to 5,500 km^2 in 2007 (Fox and Castella 2013), mainly in the central highlands and southern part. Meanwhile, in the entire Mainland Southeast Asia, 74,960 km^2 of land was converted to rubber farms from 2003 to 2014, of which 30% was conversions from low vegetations (mainly annual crops) (Hurni and Fox 2018). In addition, since 2010, the province has had an agricultural development planning policy until 2020, in which priority is given to the expansion of areas specialized in rubber, fruit trees, and safe vegetables. Therefore, the area planted with annual crops continued to be replaced by perennial crops (People's Committee of Binh Duong Province 2010; People's Committee of Binh Duong Province 2018). Meanwhile, the conversion from perennial croplands into annual croplands is mainly due to the impact of the sharp drop in rubber prices since 2011 (Hurni and Fox 2018). Many farmers have cut down their rubber plantations to grow short-term crops, mainly grass for animal husbandry. In addition, the agroforestry practices, i.e. the intercropping practices of the young rubber with annual crops such as legumes, corn, sesame rice, cassava, papaya, and bananas, are a factor increasing annual croplands in the classified maps. These issues are also found in other rubber-growing areas (Stroesser et al. 2018; Kusakabe and Chanthoumphone 2021; Huang et al. 2022; Su et al. 2022). In summary, the transition between annual and perennial croplands is due to the effects of agricultural development policies and practices, natural conditions, and product price fluctuations in the market.

In the study area, unused land is a type of land-use with special characteristics. It is considered as the intermediate class in the transition between other classes. Therefore,

besides the variation in area, its spatial distribution is hardly fixed, especially in agricultural regions. The conversion to and from unused land comes from two main reasons: agricultural activities and planning of urban and industrial development. In the former, the agricultural land is often converted to unused land and then re-cultivated in a short time, which can easily be seen in between 1995 and 2001, as an example. In the latter, its behaviour is more complicated. The planned areas are levelled and converted into bare land (considered as temporarily unused land) to prepare for construction in the following stages. However, there is a fact in Vietnam that these areas may be quickly converted into built-up areas or remain as bare land for a long time depending on the development plan and investment progress. Therefore, the accumulation of this kind of 'unbuilt-up' bare land coupled with the emergence of new bare land areas for other reasons can cause the area of unused land to fluctuate over time.

4.5.2. Factors affecting urban expansion from 1995 to 2020

The orientation of urban development in Binh Duong found in this study clearly reflects the natural conditions and history of the province as well as the province's development policies and land-use planning.

First, Thu Dau Mot city, which is the urban centre of the province, is a long-established urban area on the banks of the Saigon River. Along with Thu Dau Mot, there are some towns developed in the south, such as Lai Thieu, Bung, and Di An. This area had favourable conditions for urbanization and industrialization such as its location near Ho Chi Minh City, convenient rail, and road traffic, a long history of urban development, commercial activities and handicrafts, and dense population (Le 2019). Therefore, after re-establishing the province in 1997, the provincial government focused on urbanization and industrialization in the southern region, including Thu Dau Mot, Thuan An, and Di An (Party Committee of Binh Duong Province 1997; Party Committee of Binh Duong Province 2001).

Second, since 2007, the government has, in turn, issued decisions on the master plan for socio-economic development and urban development of Binh Duong province until 2020, with a vision for 2030 (Prime Minister of Vietnam 2007; People's Committee of Binh Duong Province 2012; Prime Minister of Vietnam 2014), to upgrade Binh Duong into a centrally-controlled city (as known as a municipality, Vietnamese: thành phố trực thuộc trung ương) in the period 2020–2030. With the orientation to become an 'industrial metropolis' and develop evenly based on a regional strategy to reduce the imbalance in urban distribution, the provincial government has expanded the distribution of industrial zones to the north. Accordingly, the Binh Duong metropolis is divided into three subregions: (1) Southern urban region (Di An, Thuan An): compact urban model, high density. (2) Central urban region (Thu Dau Mot, Ben Cat, Tan Uyen): multifunctional and multicentre model, medium density. The administrative-political centre is moved to the new urban area of Hoa Phu–Phu Tan (as known as the Binh

Duong new city) in the north of Thu Dau Mot, located in the Binh Duong Industrial–Service–Urban Complex. (3) Northern urban region (the other districts): satellite urban model, low density. Under these policies, the urban and industrial areas in Binh Duong gradually expanded to the north in latter periods.

Third, the increasing urbanization rate of the province stems from industrialization. In addition to the master plans, a series of flexible policies have been implemented by the provincial government for industrial development. They comprise policies on land acquisition, development of transport systems and industrial infrastructure, reform of administrative procedures, attraction of investment capital, and attraction of human resources (Le 2019; Le et al. 2019; Nguyen et al. 2019; Do et al. 2020). These policies can be briefly summarized as follows: (1) accelerating the land clearance and conversion from agricultural land to urban and industrial land by policies on resettlement, vocational training, and job creation for people whose land has been acquired; (2) simplifying administrative procedures to attract domestic and foreign investment to industrial parks; (3) mobilizing nonbudget capital to develop the infrastructure of industrial zones and improve the transport system based on the build-operate-transfer model; (4) implementing policies on housing and social benefits for workers. These policies have created a favourable investment environment and confidence for investors. As a result, investment capital into the province continuously increased. The total foreign direct investment capitals in the periods of 2001–2005, 2006–2010, and 2010–2015 were USD 1.8 billion, 5.7 billion, and 10.2 billion, respectively, in which investment capital in industrial zones accounted for 90% (Binh Duong Statistical Office 2016). In consequence, many projects have been invested, new factories have been continuously formed and expanded. Since the first industrial park was established in 1995, as of 2019, Binh Duong has 29 industrial parks and 12 industrial clusters, with an average occupancy rate of over 70% in which many of them have been filled over 90%.

Last, as a result of the mentioned policies, the industrial development in the province created a job market with high and stable income. The monthly average income per capita at current prices of Binh Duong province and the Southeast region is higher than that of other regions of Vietnam (Table 4.7). People tend to migrate from low- to high income areas (Phan and Coxhead 2010). Career opportunities with high income combined with housing and social benefits have led to not only the movement of people from rural to urban areas within the province but also the migration of foreigners (Figure 4.9), thereby increasing the population. Industrial development and population growth increased the demand for housing, transportation, and other utilities, creating a force for the expansion of mixed residential areas and other facilities. Since 2002, with the explosion of urban migration, the provincial government first implemented the concept of industrial park development in association with mixed residential areas and utilities

for the planning of My Phuoc industrial park. This concept was then widely applied to the planning of urban and industrial zones in later stages.

Table 4.7. Monthly income per capita at current prices of Binh Duong province, economic regions, and the whole country of Vietnam (in thousand VND).

Year Region	1999	2010	2019
Binh Duong province	no data	2698	7433
Southeast	571	2304	6280
Whole country of Vietnam	295	1387	4295
Mekong River Delta	342	1247	3886
Central Highlands	345	1088	3095
North Central area and Central coastal area	229	1018	3331
Red River Delta	282	1580	5191
Northern midlands and mountain areas	199	905	2640

Source: General Statistics Office of Vietnam (2021).

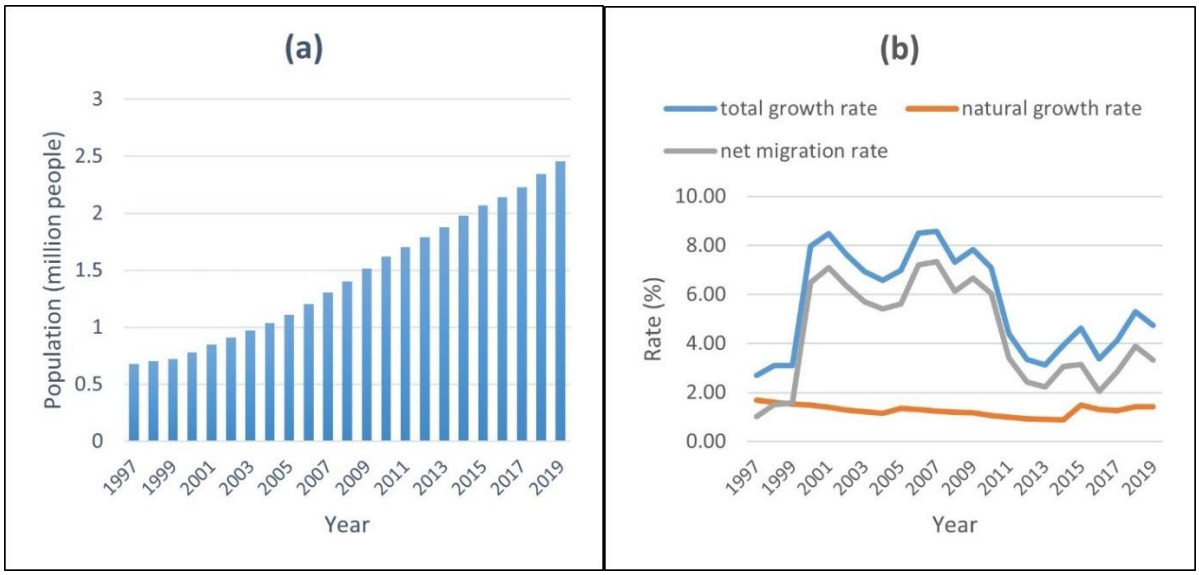


Figure 4.9. (a) Population growth and (b) growth rate in Binh Duong province (1997–2019).

Overall, the pattern of urban expansion in Binh Duong from 1995 to 2020 is the gradual transition from a compact urban form to a dispersed urban form, which is also found in many other cities in developing regions (Marengo 2015; Sumari et al. 2019; Xu, Dong, et al. 2019; Xu, Jiao, et al. 2019). This pattern is the result of the combined effects of natural conditions, history, economic development, demographics, land-use and urban planning, and policies. These findings are consistent with the results of many studies on urban sprawl in cities around the world (Reilly et al. 2009; Mahendra and Seto 2019; Shao, Sumari, et al. 2021; Mahtta et al. 2022). Of which, either population growth or economic growth was often considered the main driver for urban expansion, depending on the regional variations (Mahendra and Seto 2019; Mahtta et al. 2022). For Binh Duong, the importance of these two drivers changes over time. This is consistent with the findings of Mahtta et al. (2022) when they studied the role of population and economic growth in urban land expansion in more than 300 cities between 1970 and

2014. The authors revealed that although population growth was the primary driver, the effect of economic growth has increased significantly in importance since 2000, especially in low-income, low-middle-income, and upper-middle-income countries. They also stated that this increase occurred up to a point. Population growth, mainly due to migration, would again be an important driver since a country entered the highest income category. This is what happened in Binh Duong during the study period although only on a local scale.

4.5.3. Take-away for practice

Some practical experience, which may help in land planning and policymaking for other localities, can be learned from the pattern of land use change and urban sprawl in Binh Duong province as follow.

- Besides having good land-use and urban planning, focusing on infrastructure development, investment attraction, and human resources attraction are essential factors to accelerate the process of industrialization and urbanization.
- Despite rapid urbanization, inefficient land use (i.e. unused land) remains a dilemma. Measures to accelerate the progress of projects to increase the efficiency of urban land use are necessary.
- Urbanization often leads to an impact on the real estate market and causes housing prices to increase rapidly. This issue needs to be regulated.
- The rapid increase in housing prices due to urbanization along with fluctuations in agricultural market prices can have a negative impact on the land-use practices of farmers. It is necessary to regulate these issues to limit the conversion of land use purposes or the conversion of crops to follow short-term market trends.

4.6. Conclusion

Although the information on spatiotemporal land-use change and urbanization status is necessary for land-use planning and decisionmaking, it is still lacking in many urban regions, especially emerging urban areas such as Binh Duong province of Vietnam. By using land-use maps extracted from multitemporal Landsat images and spatial analysis techniques, the spatiotemporal pattern of land-use change and urban expansion in Binh Duong province from 1995 to 2020 were analyzed in this study.

The research results showed that there were different trends in the area variation of land-use types, and there was a large transition from agricultural and unused land to other types of land-use within 25 years. This study also revealed that the urban area of the province expanded 65 folds within the 25 years at an increasing rate. The expansion rates were uneven between subregions, and there was a gradual expansion and shift from south to north of the province and spreading to rural districts at an increasingly rapid rate during the study period. It led to a gradual transition from a compact urban form to a dispersed urban form. The factors affecting land-use change and urban expansion of Binh Duong province were also discussed. They comprise the natural conditions,

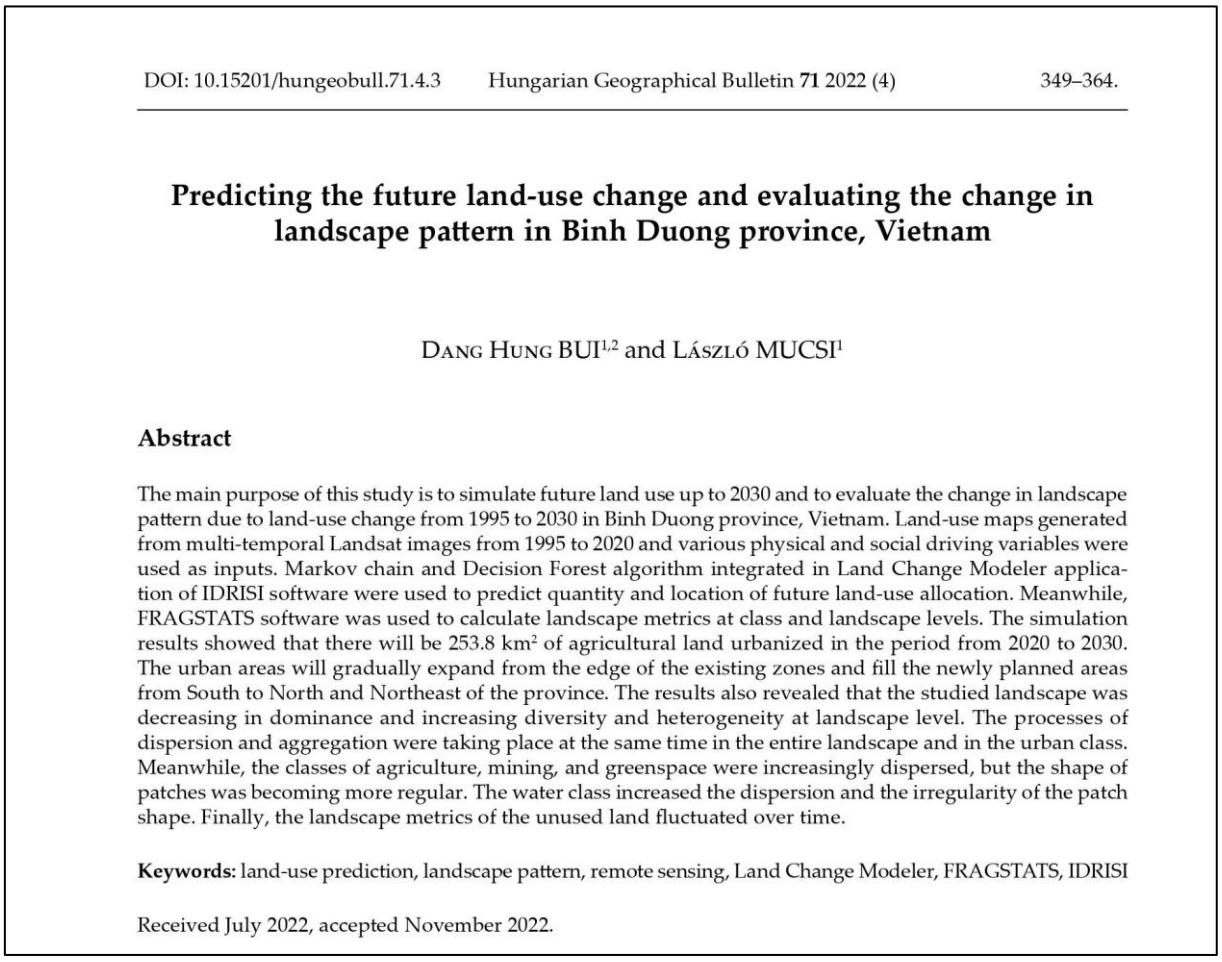
development history, policies and practices for urbanization, industrialization, and agricultural development, and fluctuation in the prices of products in the market.

The results of this study reveal a pattern of rapid urbanization in developing countries under the impact of land policies. Some practical lessons can be drawn from them. They can lay the groundwork for further studies on urban planning, land management, and policymaking in other localities not only in Vietnam but also in other countries. Further, land-use change can also cause adverse effects on the quality of the environment, landscape, and human health. Further studies on these issues are required.

5. Predicting the future land-use change and evaluating the change in landscape pattern in Binh Duong province, Vietnam

This article is published in *Hungarian Geographical Bulletin* as:

Bui DH, Mucsi L. 2022. Predicting the future land-use change and evaluating the change in landscape pattern in Binh Duong province, Vietnam. Hungarian Geographical Bulletin. 71(4):349-364. https://doi.org/10.15201/hungeobull.71.4.3



DOI: 10.15201/hungeobull.71.4.3 Hungarian Geographical Bulletin **71** 2022 (4) 349–364.

Predicting the future land-use change and evaluating the change in landscape pattern in Binh Duong province, Vietnam

DANG HUNG BUI[1,2] and LÁSZLÓ MUCSI[1]

Abstract

The main purpose of this study is to simulate future land use up to 2030 and to evaluate the change in landscape pattern due to land-use change from 1995 to 2030 in Binh Duong province, Vietnam. Land-use maps generated from multi-temporal Landsat images from 1995 to 2020 and various physical and social driving variables were used as inputs. Markov chain and Decision Forest algorithm integrated in Land Change Modeler application of IDRISI software were used to predict quantity and location of future land-use allocation. Meanwhile, FRAGSTATS software was used to calculate landscape metrics at class and landscape levels. The simulation results showed that there will be 253.8 km^2 of agricultural land urbanized in the period from 2020 to 2030. The urban areas will gradually expand from the edge of the existing zones and fill the newly planned areas from South to North and Northeast of the province. The results also revealed that the studied landscape was decreasing in dominance and increasing diversity and heterogeneity at landscape level. The processes of dispersion and aggregation were taking place at the same time in the entire landscape and in the urban class. Meanwhile, the classes of agriculture, mining, and greenspace were increasingly dispersed, but the shape of patches was becoming more regular. The water class increased the dispersion and the irregularity of the patch shape. Finally, the landscape metrics of the unused land fluctuated over time.

Keywords: land-use prediction, landscape pattern, remote sensing, Land Change Modeler, FRAGSTATS, IDRISI

Received July 2022, accepted November 2022.

Abstract

The main purpose of this study is to simulate future land use up to 2030 and to evaluate the change in landscape pattern due to land-use change from 1995 to 2030 in Binh Duong province, Vietnam. Land-use maps generated from multi-temporal Landsat images from 1995 to 2020 and various physical and social driving variables were used as inputs. Markov chain and Decision Forest algorithm integrated in Land Change Modeler application of IDRISI software were used to predict quantity and location of future land-use allocation. Meanwhile, FRAGSTATS software was used to calculate landscape metrics at class and landscape levels. The simulation results showed that there will be 253.8 km^2 of agricultural land urbanized in the period from 2020 to 2030. The urban areas will gradually expand from the edge of the existing zones and fill the newly planned areas from South to North and Northeast of the province. The results also revealed that the studied landscape was decreasing in dominance and increasing diversity and heterogeneity at landscape level. The processes of dispersion and aggregation were taking place at the same time in the entire landscape and in the urban class. Meanwhile, the classes of agriculture, mining, and greenspace were increasingly dispersed, but the shape of patches was becoming more regular. The water class increased the dispersion and the irregularity of the patch shape. Finally, the landscape metrics of the unused land fluctuated over time.

Keywords: land-use prediction; landscape pattern; remote sensing; Land Change Modeler; FRAGSTATS; IDRISI

5.1. Introduction

Socio-economic development can impact on land-use change process in many ways (Lambin and Meyfroidt 2010). In developing countries, the process of urbanization and the shift of socio-economic development policies, such as from agriculture-based to industry-oriented economy, lead to high land-use demand (Nourqolipour et al. 2016). As a result, the land-use transition is intense. Much of the transition in this context has been from natural and semi-natural to artificial landscapes. In recent years, due to population growth and urbanization, land-use change has taken place strongly in the vicinity of existing urban areas and in the key economic development zones in Vietnam (Truong et al. 2018; Nguyen and Kim 2020; Ha et al. 2020).

For example, in Binh Duong province, which is in the neighbourhood of the largest metropolis of Vietnam, and in the southern key economic zone, urbanization and industrialization have taken place very strongly in the last 25 years (Le 2019; Le et al. 2019). As a result, a large amount of agricultural land has been converted into industrial zones and urban areas. This type of conversion is still ongoing.

Land-use change affects landscape patterns and, as a result, ecosystem functions (Lin et al. 2013; Estoque and Murayama 2016; Tolessa et al. 2017; Kertész and Křeček 2019; Tang et al. 2020). Therefore, quantification of changes in landscape patterns,

including shape, size, and spatial distribution, is essential, especially where land-use change is dramatic, such as in emerging urban areas. The quantification facilitates comparison and assessment of landscape change during past and future land-use change. At the same time, it can also partly reveal the impact trend of land-use changes on the structure and function of diverse types of landscapes and ecosystems. This information may be useful for decision-making and land-use planning toward efficient use of resources and sustainable development (Vaz et al. 2014; Abdolalizadeh et al. 2019). The landscape pattern change is often assessed by landscape metrics at the three levels including patch, class, and landscape (Turner and Gardner 2015; Gergel and Turner 2017; Gudmann et al. 2020).

To calculate landscape metrics, land-use maps are often used as input. The maps in the past can be generated by historical geodetic measurement and administrative land-change records over the years. Another fast and effective method that is widely applied is to interpret from remote sensing images (Rahman 2016; B. Zhang et al. 2017; Singh et al. 2018). Although the use of remote sensing images to create land-use maps has some limitations such as resolution, classification algorithms, the ability to distinguish land use, etc., this is still a useful approach due to its promptness and proactivity. Meanwhile, future land-use maps can be collected from land-use planning maps or from simulation based on past variability trends and future demand in terms of quantity and spatio-temporal distribution (Zheng et al. 2015; Saxena and Jat 2019; Yin et al. 2021).

There are many models developed for land-change simulation, such as CLUE-S, CLUMondo, Land Change Modeler (LCM), LucSim, DinamicaEGO, SLEUTH, etc. Each model has its own pros and cons, and the choice of model to use depends on the goals and the available data of the study (Camacho Olmedo et al. 2018). LCM is one of the popular applications used to assess and simulate land-use change. The advantage of this application is that it is simple to use, easy to set up input parameters, has clear instructions, and many simulation algorithms are integrated. Many studies have used this application for land-use change prediction for various purposes (Megahed et al. 2015; Nor et al. 2017; Islam et al. 2018; Mishra et al. 2018).

With the mentioned issues in mind, this study was carried out for two main purposes including (1) Using LCM to simulate the future land use in Binh Duong province in 2025 and 2030, and (2) Quantification and evaluation of landscape change due to land-use change from 1995 to 2020 and forecast to 2030.

5.2. Materials and methods

5.2.1. Study area

This study was conducted in Binh Duong province which located in the southeast region of Vietnam (Figure 5.1). The land-use change in the province took place dramatically from 1997 when the province was re-established. Agricultural land and unused land were converted to other uses, most of which were devoted to expanding residential and

industrial areas. These changes were mainly due to socio-economic factors including urbanization, industrialization, and structural changes in agricultural production, and related policies (Le 2019; Le et al. 2019; Bui and Mucsi 2022).

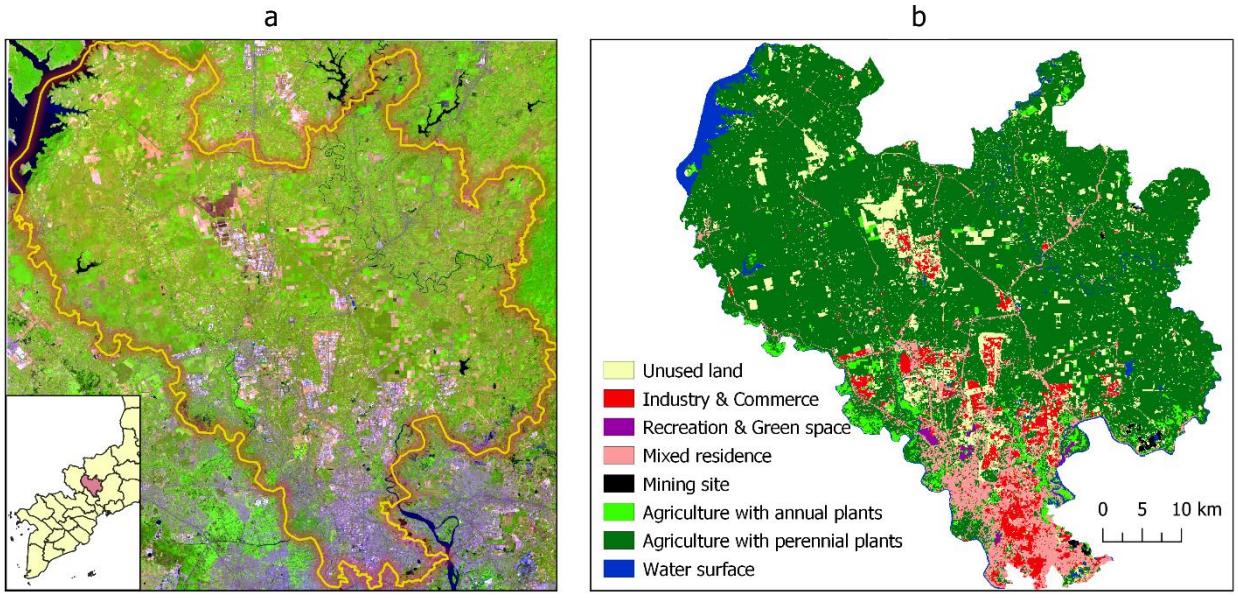


Figure 5.1. Study area in two maps. a = Composite from Landsat-8 OLI image (RGB: 6-5-2) acquired on 06/01/2020 and downloaded from the USGS website (https://earthexplorer.usgs.gov/); b = Land-use map in 2020 derived from the study of Bui, D.H. and Mucsi, L. (2022).

5.2.2. Data

This study used the land-use maps in 1995, 2001, 2005, 2010, 2015, and 2020 which were generated from multi-temporal Landsat images from the study of Bui, D.H. and Mucsi, L. (2022). The map was in the WGS-84 UTM 48N projection and a spatial resolution of 30 m and consisted of 8 land-use types (Table 5.1). The overall accuracy of these maps was reported to be 89.2, 88.9, 89.6, 90.8, 93.0, and 90.1 percent, respectively. The producer's accuracy ranged from 70.8 to 100 percent, while the user's accuracy ranged from 70.9 to 100 percent. Therefore, it is appropriate and dependable to use them for land-change prediction and landscape analysis.

Table 5.1. Land-use categories.

ID	Original land-use class	New class for Land Change Modeler	New ID
1	Unused land	Agricultural land	1
2	Industry and Commerce	Industry and Commerce	2
3	Recreation and Greenspace	Others	3
4	Mixed residence	Mixed residence	4
5	Mining site	Others	3
6	Agriculture with annual plants	Agricultural land	1
7	Agriculture with perennial plants	Agricultural land	1
8	Water surface	Others	3

To explore the drivers for the land-use change, which was a key step for the simulation model of land-use change, several kinds of data were collected and pre-

processed. SRTM 1 Arc-Second 30m digital elevation model (DEM) was downloaded from the website https://earthexplorer.usgs.gov/. Slope and aspect were extracted from the DEM. Population density raster data were downloaded from the website https://www.worldpop.org/ with a spatial resolution of 100 m. A raster of the mean population density in the period of 2010–2020 was calculated and resampled to a 30-m spatial resolution using bilinear method. All these raster data were pre-projected to WGS-84 UTM 48N. In addition, open water surfaces were extracted from the Open Street Map project (https://www.openstreetmap.org/) and downloaded from the website https://download.geofabrik.de/. Forest protection areas and planned industrial parks for 2020 and 2030 were extracted from the planning map of the provincial government. The 3-level main road network was extracted from the administrative map in 2014 and modified based on the Google satellite images.

The location points of the administrative and economic centre of the province and districts (hereinafter referred to as province centre and district centres, respectively), airport, train stations, and river ports were manually digitalized based on the Google satellite images. All these data were collected in vector format. Therefore, they were rasterized to a spatial resolution of 30 m and a projection of WGS-84 UTM 48N. After that, the maps of Euclidean distance to the open water surfaces, planned industrial parks, main roads, province centre, district centres, and transportation ports were extracted in turns. Furthermore, based on the land-use maps, the map of Euclidean distance to current residential and industrial areas in 2001 and 2020 was also produced, respectively.

5.2.3. Land-use change prediction

This study tended to simulate the land use of the study area in 2025 and 2030 based on the land-use maps of previous periods and land-use change drivers. We only focused on simulating the transition from agricultural land to urban land, which was the major transition taking place in recent years. The land-used maps were re-classed from eight to four categories as shown in Table 5.1. It should be noted that the two urban classes were not grouped together because their expansion was driven by varied factors. As a result, the transition from agricultural land to urban land would be included two sub-models. One was the transition from agricultural land to industrial and commercial regions (agri_to_indus), and the another was the transition from agricultural land to mixed residential areas (agri_to_resid). Other conversion types were ignored.

The LCM application integrated in the Terrset IDRISI 2020 software was used. The simulation process consists of calibration, validation, and prediction. The overall process is illustrated in the Figure 5.2. The LCM includes six algorithm options for simulation, including Multi-layer Perceptron neural network, Decision Forest (DF), Logistic regression, Support Vector Machine, Weighted Normalized Likelihoods, and SimWeight. After some trials, the DF algorithm was chosen. The number of trees was

set at 100, and the number of variables at split was the square root of a number of input variables.

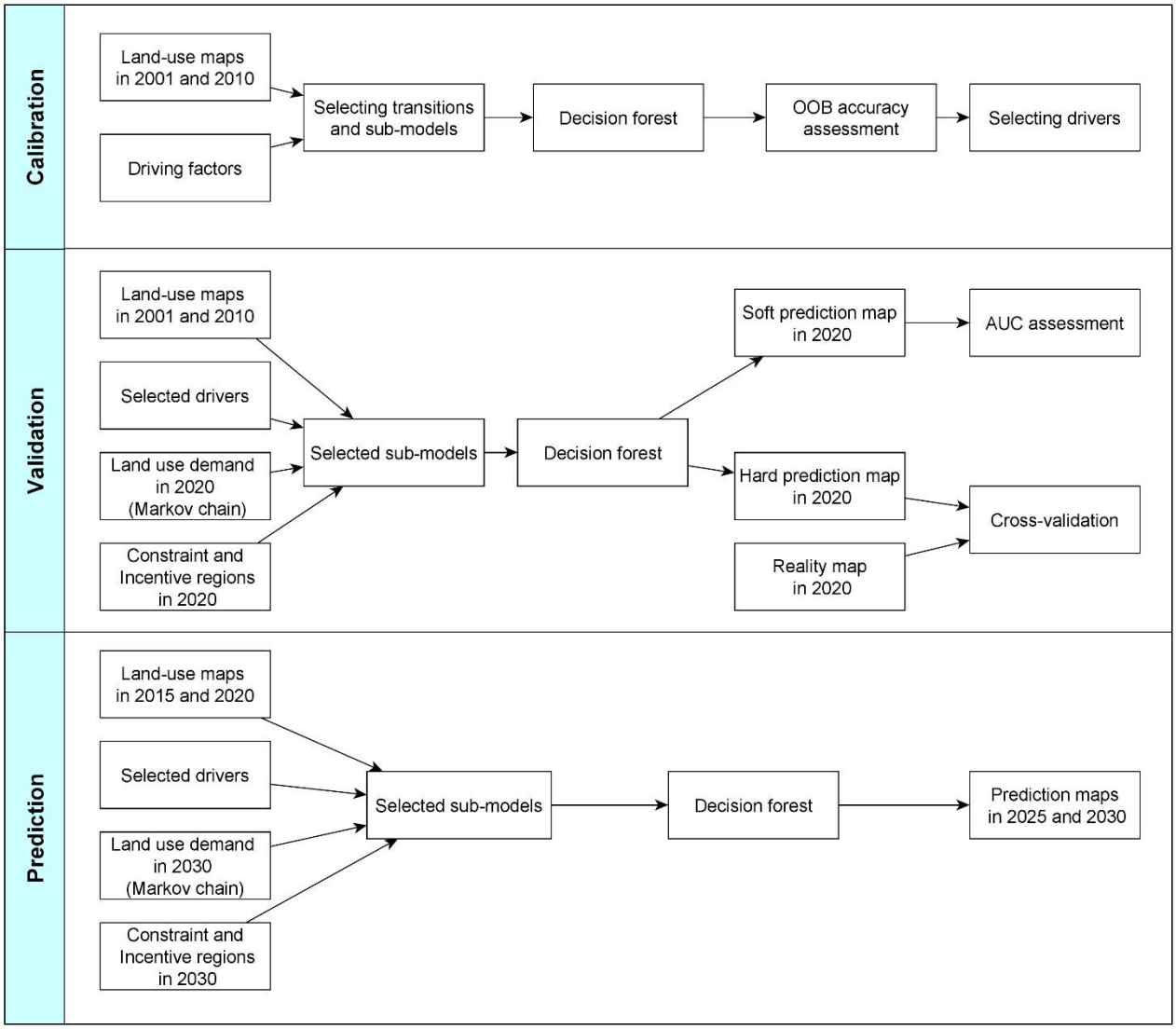


Figure 5.2. Simulation process.

At the calibration phase, the land-use maps of 2001 and 2010 were used as the earlier and later maps, respectively, combined with a set of variables to build the model. The purpose of these phases was to select appropriate variables as drivers for the land-use change transitions. The variable selection was based on the Out of bag (OOB) accuracy in the output report of the transition sub-models. If the OOB accuracy when holding a given variable constant was greater than OOB accuracy with all variables, it means that the given variable might not be significant in the model (Eastman 2020a), and it was excluded. At the validation phase, the predicted map in 2020 was simulated and compared with the reality map in 2020 to validate the model. The performance of the model was evaluated by the Kappa coefficients (Pontius 2000; Hagen 2002; Hagen 2003; Hagen-Zanker et al. 2005) and Figure of merit (FoM) (Pontius et al. 2008) for the hard-classification and the area under the curve (AUC) (Mas 2018) for the soft-classification outputs.

After the performance of the model was confirmed and satisfied, the prediction phase was performed. In this phase, the land-use maps in 2015 and 2020 were used as input to predict the maps in 2025 and 2030 with the same set of drivers selected at the calibration and validation phases. The reason to use the maps 2015 and 2020 was that the urban area in Binh Duong province has expanded at an increasing rate from 1995 to 2020 (Bui and Mucsi 2022), therefore, the two latest maps used may capture the most recent trend of urban expansion. This may more accurately reflect future land-use demand for the simulation. The land-use demand was calculated based on the Markov chain with the assumption that future conversion would be at a similar rate to the current period (Zheng et al. 2015). This calculation was built-in into LCM. The LCM also allows setting constraints and incentives for a particular type of conversion. The weights for these regions can be set between in a range of 0 to 1, where 0 is strictly forbidden and 1 is strongly encouraged. In this study, the protection forest was considered the prohibited area for both types of urbanization (weight of 0). For the agri_to_indus sub-model, it was encouraged to develop inside the planned industrial parks with a weight of 1, and the rest was set to a weight of 0.1. For the agri_to_resid sub-model, the weights were set to 1 and 0 for areas outside and inside the planned industrial parks, respectively.

After the prediction phases, the 4-class predicted maps in 2025 and 2030 were overlaid with the land-use map in 2020 (8 classes) to generate the 8-class land-use maps in 2025 and 2030, which would be used for calculating landscape metrics.

5.2.4. Landscape metrics

To measure the change in landscape patterns over time, this study used landscape metrics (McGarigal et al. 2012; Turner and Gardner 2015; Gergel and Turner 2017) at landscape and class levels. Because the mixed residential, industrial, and commercial areas formed the urban landscape, they were re-classed into a common class named urban. From the land-use maps, landscape metrics were calculated in FRAGSTATS 4.2 software based on the eight-cell neighbour rule (McGarigal et al. 2012). The metrics were chosen so that they were representative of the features of the landscape, were not redundant, and have been widely and effectively used in previous studies (Su et al. 2014; Dadashpoor et al. 2019). The features measured included dominance, diversity, and fragmentation. The selected metrics is shown in Table 5.2, and a detailed definition and description of the metrics can be found in the FRAGSTATS Manual document (McGarigal et al. 2012).

5.3. Results

5.3.1. Simulation of land-use change in future

5.3.1.1. Driving factors

Based on the results of the analysis of OBB accuracy in the DF outputs, the drivers included in the two sub-models are presented in Table 5.3. The drivers included in these sub-models are reasonable. A common point of both sub-models is that natural factors

related to topography (DEM, slope, aspect) do not affect urbanization. Possibly because the terrain of the whole area is relatively flat, except for a few low-mountain areas within the protected area, the weighting of these factors is likely to be the same in most places. The impact of other drivers of each sub-model was explained in detail below.

Table 5.2. Landscape metrics used.

Metric	Name	Level used	
		Landscape	Class
AREA_MN	Mean Patch Size	x	x
CONTAG	Contagion Index	x	–
IJI	Interspersion and Juxtaposition Index	x	x
LPI	Largest Patch Index	x	x
LSI	Landscape Shape Index	x	x
NP	Number of Patches	x	x
PD	Patch Density	–	x
PLAND	Percentage of Landscape	–	x
SHDI	Shannon's Diversity Index	x	–
SHEI	Shannon's Evenness Index	x	–

Table 5.3. Drivers for sub-models.

No.	Input drivers	Selected drivers by Decision Forest algorithm	
		Agri_to_resid	Agri_to_indus
1	DEM	–	–
2	Slope	–	–
3	Aspect	–	–
4	Distance to water sources	–	x
5	Distance to province centre	x	–
6	Distance to district centres	x	x
7	Distance to existing residential areas	x	–
8	Distance to existing industrial areas	–	x
9	Distance to planned industrial zones	–	x
10	Distance to main road	x	–
11	Distance to ports	–	x
12	Mean population density in 2010-2020	x	–

For the agri_to_resid sub-model, the included drivers can be explained by the following reasons. First, the new settlements are often formed from the edge of existing neighbourhoods. Second, the more populous the places, the higher the demand for housing and utilities. Third, the choice of housing also depends on the accessibility to utility services, which are often concentrated in the central areas of the province and districts. Last, to access these facilities as well as workplaces, accessibility to the transportation network is clearly an influencing factor. Meanwhile, the excluded factors may be due to several reasons. According to the general development orientation of the province, residential areas are formed close to industrial zones, which make up industrial – urban – service complexes, thus, making the distance to the existing industrial park redundant. Except for Tan Son Nhat Airport, the rest of ports (train stations and river harbours) are cargo stations, not passenger stations, so it has no impact. The distance to

the water source is not included probably because residential areas mainly use water from boreholes or water supply systems, which are relatively well distributed in urban areas.

Similarly, for the sub-model of agri_to_indus, the impact of included drivers can be explained as follow. First, new factories tend to form near previously developed places where infrastructure already exists. Second, the selection of sites within or near planned industrial zones is also to take advantage of the planned infrastructure and preferential policies from the provincial government. Third, reducing the distance to district centres and ports can increase market access and reduces transportation costs. Last, the ability to access water is probably to serve the needs of exploiting water resources for production activities. Meanwhile, the excluded drivers can be explained by some reasons. Similar to the case of the agri_to_resid sub-model, the distance to the existing residential areas is redundant. Distance to the province centre is also redundant compared to the distance to district centres. Besides, population density does not affect industrial development, maybe because of convenient transportation, people can go to work farther, so it is not necessary to form factories near densely populated areas to utilize human resources. Interestingly, the distance travelled does not affect the model either. Maybe because the current transport system has developed relatively widely, and the planning of new industrial zones also leads to the expansion of the transport network to access these zones. Therefore, this variable has no effect.

5.3.1.2. The performance of selected model

Four different maps of the study area (reality map, hard prediction map, soft prediction map and cross-validation map in 2020) are illustrated in Figure 5.3.

For hard prediction, the Kappa coefficients and FoM were used to evaluate the accuracy of the predicted map in 2020 and thereby validate the performance of the selected model. The results showed that Kappa, Kappa location, and Kappa histogram coefficients reached 0.71, 0.72, and 0.99, respectively. The simulated map contained the percentages of hits, null successes, misses, and false alarms of 3.77, 87.75, 4.54, and 3.94 percent, respectively. As a result, the FoM achieved 30.77 percent, producer's accuracy achieved 41.71 percent, and user's accuracy achieved 48.88 percent.

It can be seen that these values were relatively low. An important source of error was that the hard classification result was only one outcome in many equally plausible scenarios (Eastman 2020b). Therefore, it was difficult to predict exactly the location in terms of pixel-level where the change would take place. As can be seen visually, the hits, false alarms, and misses tended to occur in the same location in close proximity. This revealed that predicting the location of the change was relatively accurate. For 2-dimensional assessment, when validating by fuzzy Kappa using the exponential decay function (radius of neighbourhood = 4, halving distance = 2), the fuzzy Kappa value reached 0.77 and the average similarity achieved 0.94, which is much better than the

traditional kappa coefficients. In addition, for 3-dimensional assessment, according to Pontius, R.G. *et al.* (2008), the FoM is proportional to net changes in the study area. In this study, the actual rate of change from agricultural land to urban in the period 2010–2020 accounted for 4.32 percent of the entire area and 8.52 percent of the total agricultural area in 2010. The calculated FoM value was relatively high compared to these rates. Furthermore, Peng, K. *et al.* (2020) mentioned that "the spatial allocation algorithm cannot well simulate the isolated patches that newly emerged". Last but not least, the FoM value in this study was higher than that in other studies, where the FoM was less than 20 percent (Megahed et al. 2015; Peng et al. 2020).

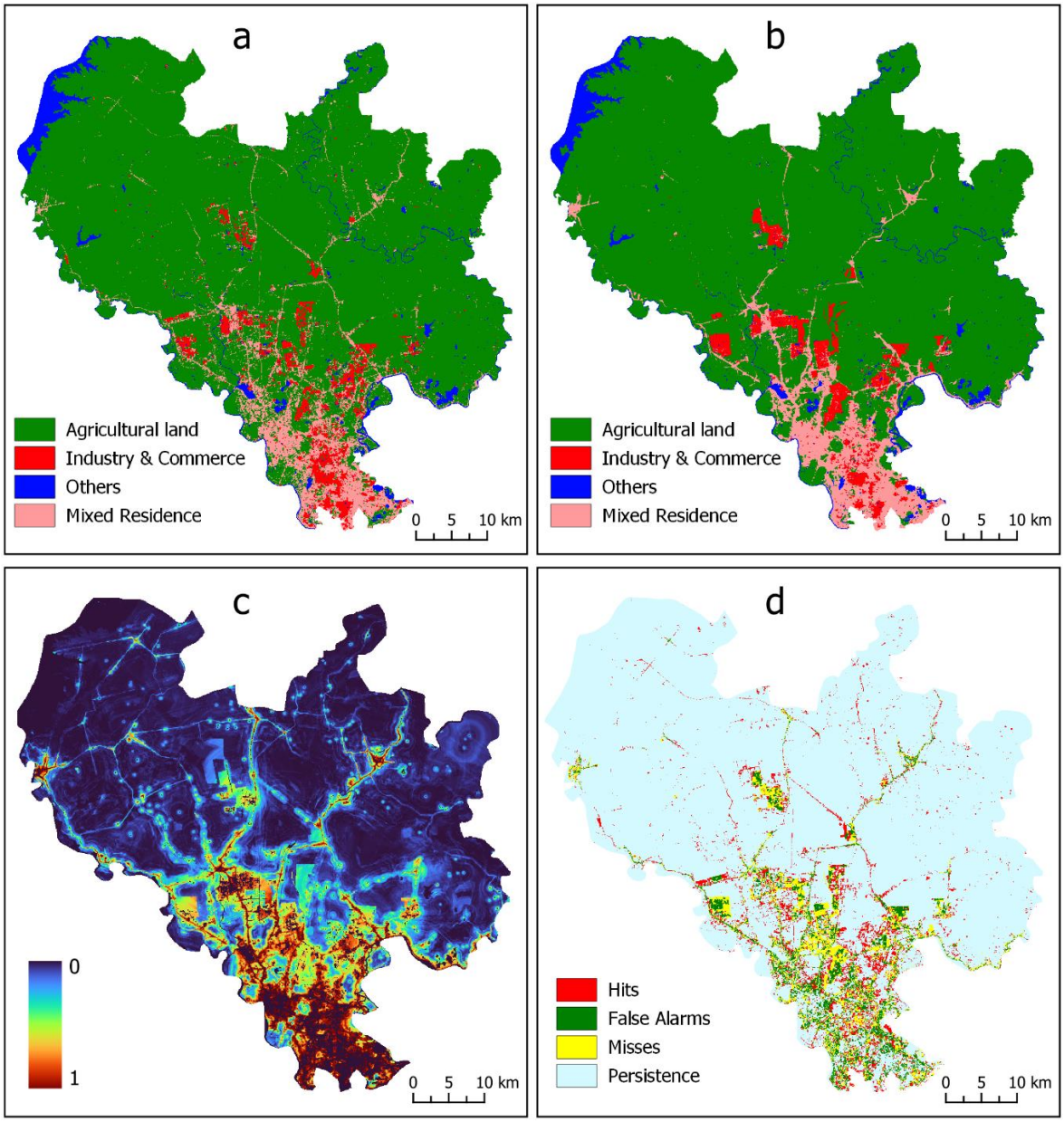


Figure 5.3. Reality map (a), hard-prediction map (b), soft-prediction map (c), and cross-validation map (d) for the study area in 2020.

The soft prediction result was validated by the AUC. The AUC is an index used to evaluate "how well a continuous surface predicts the locations given the distribution of a Boolean variable" (Eastman 2020b), and it was calculated from the receiver

operator characteristic (ROC). The AUC value ranges from 0.5 (bad model) to 1 (perfect model) (Estoque and Murayama 2016; Peng et al. 2020). The AUC in our model reached 0.96, which validated that the model could simulate potential areas for urban expansion from agricultural land with high accuracy.

5.3.1.3. Predicted maps and land-use change in 2025 and 2030

The simulation results gave that a total of 126.9 km^2 and 253.8 km^2 of agricultural land are expected to urbanize by 2025 and 2030, respectively. Specifically, residential areas may expand to 309.3 km^2 in 2025 and 395.9 km^2 in 2030, corresponding to an increase of 86.5 km^2 (138.8%) and 173.1 km^2 (177.7%), respectively, compared to 2020. The residential development is still concentrated in the South of the province and around the centre of the districts, where the infrastructure for development is an advantage. Meanwhile, the area of industrial and commercial zones may reach 150.4 km^2 in 2025 and 190.8 km^2 in 2030, corresponding to an increase of 40.4 km^2 (136.7%) and 80.8 km^2 (173.4%), respectively, compared to 2020. The new factories are going to fill the existing industrial parks and expand to the new planned industrial zones in the North and Northeast.

Corresponding to this urban expansion, from 2020 to 2025, perennial cropland, unused land, and annual cropland may be decreased by 77.8 km^2, 40.7 km^2, and 8.4 km^2, corresponding to a decline of 4.0, 16.4, and 9.4 percent, respectively, compared to 2020. Meanwhile, by 2030, these land-use types may be decreased by a total of 168.8 km^2, 67 km^2, and 18 km^2, corresponding to a decline of 8.8, 27.0, and 20.3 percent, respectively, compared to 2020. The predicted land use in 2025 and 2030 are illustrated in Figure 5.4.

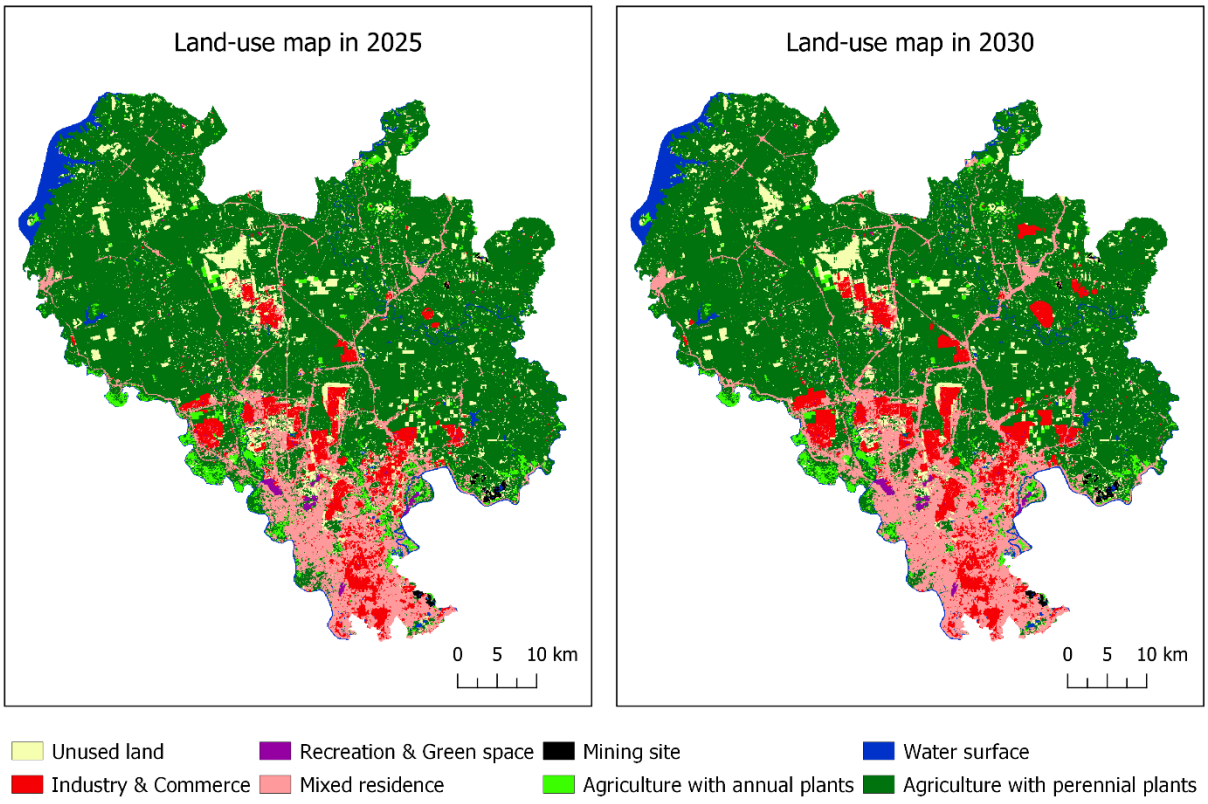


Figure 5.4. Predicted land use in 2025 (left) and in 2030 (right).

5.3.2. Landscape pattern change

5.3.2.1. Landscape level

The trends of the landscape indices at the landscape level are shown in Figure 5.5. Landscape change was analysed according to dominance, diversity, and fragmentation.

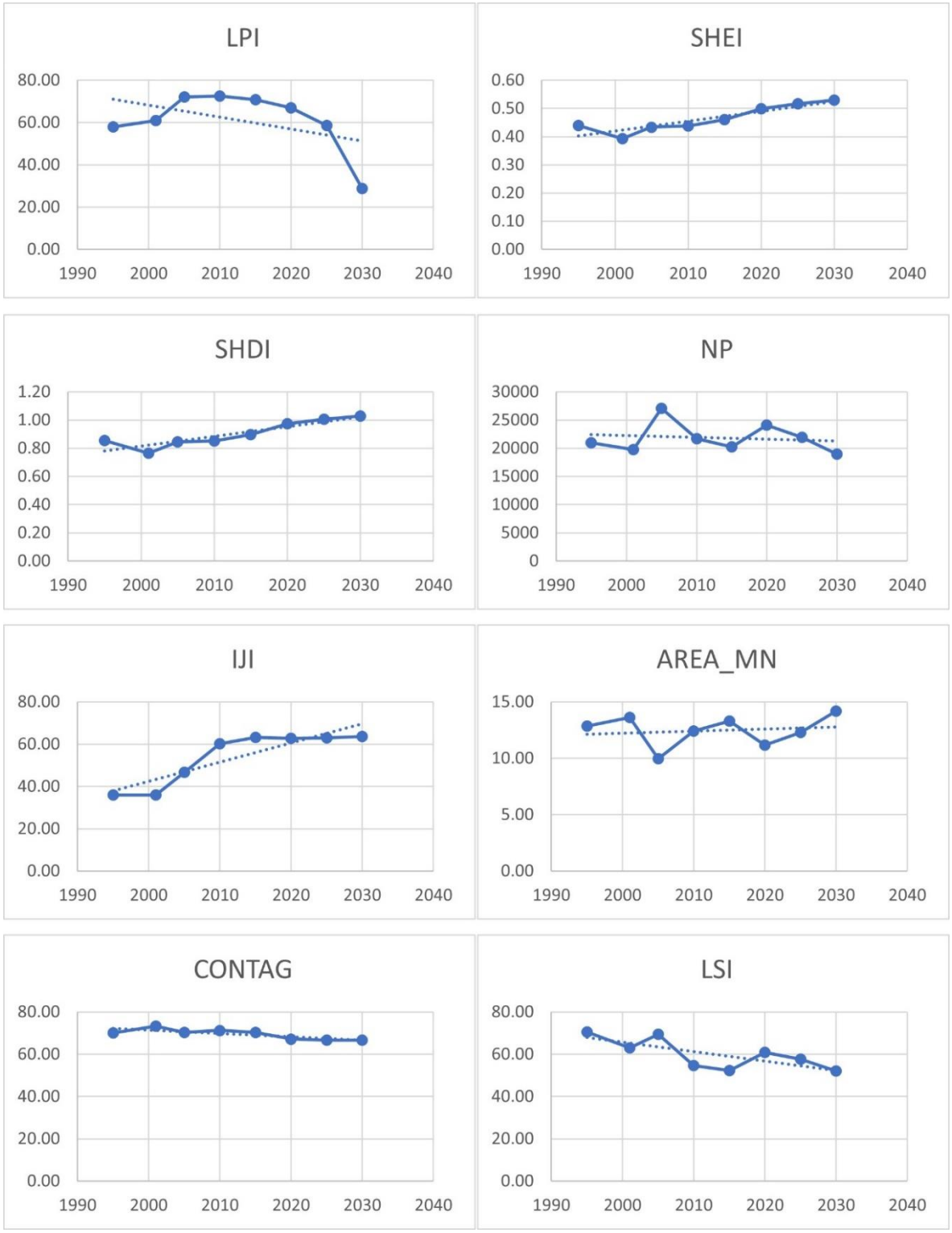


Figure 5.5. Landscape metrics calculated at landscape level. LPI = Largest Patch Index; SHEI = Shannon's Evenness Index; SHDI = Shannon's Diversity Index; NP = Number of Patches; IJI = Interspersion and Juxtaposition Index; AREA_MN = Mean Patch Size; CONTAG = Contagion Index; LSI = Landscape Shape Index.

Dominance: The dominance in the studied landscape was revealed by the LPI and SHEI. LPI increased in the period 1995–2010, then decreased in the period 2010–2020. It was also predicted to continuously decrease sharply in the period 2020–2030. Meanwhile, the SHEI decreased during the period 1995–2001 but increased continuously from 2001 to 2020 and was expected to continue to increase until 2030. The overall trend for LPI was to decrease while SHEI was to increase over the entire study period. This showed that although there was still a high dominance of a class in the landscape (in this case, the woodland), the area proportion of the classes was tending towards a more uniform distribution. In other words, there is a trend of transitioning from a landscape with only one dominant land-use type to a mixed landscape with many different land uses (Weng 2007).

Diversity: Landscape diversity was reflected by the SHDI, which tended to increase over the study period. Of which, the SHDI decreased in the period 1995–2001, increased continuously in the period 2001–2020, and was forecasted to continue to increase until 2030. This indicated an increase in diversity, which also means heterogeneity, in the landscape.

Fragmentation: The results showed an increasing trend of AREA_MN and IJI and a decreasing trend of NP, CONTAG, and LSI. NP and AREA_MN were the two indices that had an opposing trend and represented the characteristics of the land-use transformation in the study area, which had both dispersion and aggregation processes. When the landscape was fragmented, new fragments were formed (NP increased), and the average area of fragments decreased (AREA_MN decreased). But as these individual patches were gradually expanded, and clumped together into a larger patch, NP would be decreased and AREA_MN would be increased. The increasing trend of AREA_MN and decreasing trend of NP in the whole study period revealed that the aggregation process may be probably stronger, especially from 2020 to 2030. An increase in the IJI indicated that the landscape was more dispersion. However, this trend only took place strongly in the period 2001–2015, which most influenced the overall trend, while in other periods the increase was insignificant. A decrease in the CONTAG indicated a slight decrease in the degree of aggregation and infectivity between regions of the same class, i.e., an increase in the degree of interlacing, while a decrease in LSI revealed that structure fragments become less irregular and less complex.

In general, the results showed that the indices have a fluctuation over time, and the fragmentation of the landscape still existed in parallel with the aggregation, but the aggregation was somewhat stronger. This can be largely attributed to the strong transition from crops to woody land from 1995 to 2005, and then urban expansion in later stages, when urban areas formed separately at first were gradually expanded and became more interconnected, forming more compact regions with more regular shapes. In addition, part of this may also be because the prediction was only interested in the transition from agricultural land to urban.

5.3.2.2. Class level

The calculation results of the class-level metrics are presented in Table 5.4.

Table 5.4. Landscape metrics calculated at class level.

Land-use type	Year	PLAND	NP	PD	LPI	LSI	AREA_MN	IJI
Agriculture with perennial plants	1995	70.52	3472	1.29	57.91	71.20	54.69	37.08
	2001	75.64	2834	1.05	60.91	60.79	71.86	32.49
	2005	75.51	3383	1.26	72.07	64.28	60.10	47.87
	2010	77.13	2832	1.05	72.55	44.99	73.32	66.77
	2015	75.29	2728	1.01	70.91	43.41	74.31	71.27
	2020	71.52	3310	1.23	66.94	54.20	58.17	73.36
	2025	68.62	3520	1.31	58.55	52.97	52.49	72.48
	2030	65.24	3122	1.16	28.82	49.26	56.27	72.36
Agriculture with annual plants	1995	20.20	10238	3.80	2.16	113.30	5.31	31.78
	2001	17.98	11060	4.11	2.12	119.89	4.38	29.78
	2005	13.09	14550	5.40	0.78	129.10	2.42	39.67
	2010	5.49	8519	3.16	0.47	93.41	1.74	45.32
	2015	3.69	5853	2.17	0.21	73.99	1.70	52.95
	2020	3.29	5482	2.04	0.14	69.11	1.62	50.53
	2025	2.98	5024	1.87	0.14	66.74	1.60	53.45
	2030	2.62	4458	1.66	0.14	63.01	1.58	54.19
Urban	1995	0.19	381	0.14	0.03	21.54	1.34	63.72
	2001	0.95	787	0.29	0.14	33.34	3.27	63.33
	2005	2.50	1128	0.42	1.90	38.66	5.97	66.03
	2010	4.77	2633	0.98	2.84	55.96	4.88	58.14
	2015	7.39	2949	1.10	4.63	56.56	6.75	63.28
	2020	12.36	4514	1.68	8.24	70.90	7.37	58.48
	2025	17.07	3779	1.40	12.34	62.17	12.16	59.66
	2030	21.79	2874	1.07	17.22	49.90	20.41	61.28
Mining site	1995	0.04	28	0.01	0.01	7.74	4.17	51.16
	2001	0.05	39	0.01	0.01	8.55	3.63	63.53
	2005	0.11	17	0.01	0.06	6.83	17.87	74.70
	2010	0.16	34	0.01	0.04	7.07	12.51	79.77
	2015	0.22	26	0.01	0.10	7.12	22.73	83.54
	2020	0.25	27	0.01	0.09	8.21	25.16	86.22
	2025	0.25	27	0.01	0.09	8.21	25.16	86.36
	2030	0.25	27	0.01	0.09	8.21	25.16	86.10
Recreation and Green space	1995	0.02	25	0.01	0.01	4.07	1.81	64.62
	2001	0.04	67	0.02	0.03	6.41	1.55	72.27
	2005	0.03	106	0.04	0.02	6.84	0.79	77.29
	2010	0.19	347	0.13	0.09	11.42	1.45	77.48
	2015	0.30	804	0.30	0.09	20.79	1.02	72.71
	2020	0.45	1291	0.48	0.09	28.50	0.94	68.84
	2025	0.45	1291	0.48	0.09	28.50	0.94	61.69
	2030	0.45	1291	0.48	0.09	28.50	0.94	56.19

Land-use type	Year	PLAND	NP	PD	LPI	LSI	AREA_MN	IJI
Water surface	1995	2.38	458	0.17	1.55	25.50	14.02	33.12
	2001	2.61	593	0.22	1.64	28.38	11.83	37.75
	2005	2.30	554	0.21	1.33	29.73	11.20	52.99
	2010	2.96	708	0.26	1.64	33.65	11.26	54.51
	2015	2.97	729	0.27	1.55	34.57	10.97	56.80
	2020	2.91	887	0.33	1.47	36.80	8.83	63.31
	2025	2.91	887	0.33	1.47	36.80	8.83	65.51
	2030	2.91	887	0.33	1.47	36.80	8.83	66.93
Unused land	1995	6.64	6346	2.36	0.24	87.63	2.82	38.99
	2001	2.73	4360	1.62	0.10	70.75	1.68	55.99
	2005	6.45	7324	2.72	0.29	90.14	2.37	50.55
	2010	9.30	6595	2.45	1.15	86.88	3.80	58.98
	2015	10.13	7152	2.66	1.63	85.72	3.81	54.59
	2020	9.22	8616	3.20	0.79	90.33	2.88	50.84
	2025	7.71	7397	2.75	0.79	83.81	2.81	49.47
	2030	6.73	6340	2.35	0.51	76.58	2.86	47.92

Agriculture with perennial plants (AP): The PLAND and LPI of AP increased between 1995 and 2010, decreased between 2010 and 2020, and were expected to continue to decline until 2030, while the NP and AREA_ MN fluctuated. The PLAND always accounted for the largest proportion in the landscape (over 65%), and the LPI and AREA_MN were also much higher than the rest classes, while its NP is smaller than that of agriculture with annual plants (AA) and unused land (UL). It showed that AP was the dominant class in terms of the area and size of the patches. Since 2010, there has been a trend of gradually decreasing dominance and increasing dispersion (PLAND, LPI, and AREA_MN decreased, and NP and IJI increased), but the degree of dominance and aggregation was still high, and the shape of the patch was gradually less complex (LSI decreased).

Agriculture with annual plants: The PLAND of AA was steadily decreasing from about 20.0 percent in 1995 to 3.29 percent in 2020 and to 2.62 percent in 2030. Its LPI, AREA_MN, and NP showed a strong downward trend. The NP was reduced but still higher than the rest classes except for the UL. Meanwhile, the IJI increased, and LSI decreased. This showed that AA was increasingly decreasing in area, and at the same time, the degree of fragmentation was high. The shape of patches of AA was the most irregular compared to other classes, but it tended to become more regular over time.

Urban: The PLAND of urban grew rapidly from 0.19 percent in 1995 to 12.36 percent in 2020 and is forecasted to be 21.79 percent in 2030. The LPI, NP, and AREA_MN increased. This revealed two parallel processes in this class including (1) A gradual expansion from the edge of existing cities and interconnection between urban regions, which increased clumping and aggregation (LPI and AREA_MN increased and

IJI decreased) and (2) The formation of new discrete urban areas (NP increased). Thus, the dispersion here was due to the second process, not division from existing urban patches. In addition, these two processes also caused the shape of patches to fluctuate (LSI fluctuated).

Mining site (MS) and Recreation and Green space (RG): These were two rare classes in the landscape accounting for a small proportion (< 0.5%). However, they also showed an increasing trend over the years in terms of PLAND and LPI. For the RG class, NP increased, AREA_MN decreased, IJI changed slightly, and LSI increased. They revealed that RG areas were formed more, and they were more discrete and less connected. Furthermore, the shape of its patches more complicated. Similar to the urban class, the fragmentation here was mainly due to new formations, not division from existing patches. Meanwhile, for the MS class, NP fluctuated, AREA_MN increased, IJI increased, and LSI decreases slightly. This showed that the area of quarries was gradually expanding and was more dispersed with a more regular shape.

Water surface: The PLAND slightly increased, NP increased, AREA_MN decreased, IJI increased, LSI increased, and LPI was relatively stable over the years. This revealed that the new water surface areas were formed separately and more irregularly.

Unused land: This class had special characteristics. It was an intermediary for conversion between other classes, so the indices of this class often fluctuated strongly over the years.

In general, from 1995 to 2020, the study area experienced an intense change in the direction of increasing the fragmentation and dispersion of natural and semi-natural landscapes. These changes might be largely influenced by two parallel processes of urban landscape including aggregation and dispersion. These changing trends are forecast to continue. Clearly, changes in land use, and consequent changes in landscape pattern, are often aimed at serving the needs of socio-economic development. However, the fragmentation and dispersion of natural and semi-natural landscapes can have negative impacts on the ecological environment, ecosystem services, and benefits humans derive from them (Estoque and Murayama 2016; Tolessa et al. 2017), and, thus, may influence the sustainable development goals. Some of the major environmental conflicts that will arise in the next decade in the study area may include (1) a decline in provisioning services (food, raw material) due to the decline in agricultural land, (2) a decline of the regulating services (climate, water/water flow, erosion and fertility, purification and detoxification of water, air, and soil) due to an increase in impervious surfaces, and (3) a decline in supporting services (ecosystem process maintenance) due to fragmentation of natural and semi-natural habitats. Due to the limitation of the objective of this study, we did not quantify these aspects. For a more definitive assessment, further studies are needed.

5.4. Conclusions

Based on land-use maps of previous periods, this study used the LCM application of IDRISI software to forecast land use in Binh Duong province, Vietnam to 2030, mainly the transition from agricultural-land types to urban-land types. The Markov chain and the Decision Forest algorithm were used to predict future land-use allocation in terms of quantity and location, respectively. Various drivers were assessed. The research results revealed that the drivers of distances to province centre, district centres, existing residential areas, and main road and mean population density had an impact on the conversion from agricultural land to residential land, while the transition from agricultural land to industrial and commercial areas was driven by the factors of distances to water sources, district centers, existing industrial areas, planned industrial zones, and transportation ports. The selected model has been validated with the accuracy of the hard prediction being Kappa = 0.71, Kappa location = 0.72, Kappa histogram = 0.99, fuzzy Kappa = 0.77, and FoM = 30.77 percent and the accuracy of the soft prediction being AUC = 0.96. This result indicated that the model was suitable to predict the future land use in the study area. The simulation results showed that, in the period from 2020 to 2030, there will be 253.8 km^2 of agricultural land urbanized. The residential areas and the industrial-commercial zones are expected to expand to 395.9 km^2 and 190.8 km^2 , respectively. These areas will expand in the direction of gradually expanding from the edge of the existing zones and filling the newly planned areas from south to north and northeast.

This study also measured landscape pattern change caused by land-use change using landscape metrics calculated on FRAGSTATS software. At the landscape level, the results revealed that the studied landscape was increasingly decreasing in dominance and increasing diversity and heterogeneity. In addition, the processes of dispersion and aggregation are taking place at the same time. At the class level, the classes of agriculture, mining, and greenspace were increasingly dispersed, but the shape of patches was becoming more regular. Meanwhile, the urban class had similar characteristics to the entire landscape in terms of two parallel processes including dispersion and aggregation. The water class increased the dispersion and the irregularity of the patch shape. Finally, the landscape metrics of the unused land fluctuated over time.

This study provides insight into the causes and consequences of land-use change, especially in emerging urban areas in developing countries where sustainable development often has to trade-off with economic development goals. Changes in land use and landscape can affect the ecological environment, ecosystem services, and the benefits humans derive from them. Further studies on these issues are needed.

6. Conclusions

6.1 Summary of key findings

By addressing the research hypotheses to achieve the research objective, my study has achieved the following key findings:

- **Thesis 1.** I proved that land cover and land use in Binh Duong province were not only connected by spatial distribution and spectral properties but also by temporal characteristics. On the one hand, each land use type has its own spatial pattern and structure characterized by the properties of the land cover classes within it, such as composition, spatial distribution, spectral signature, and dominant class as well as the shape and size of objects. On the other hand, the change or non-change of land cover at a given site over different times of the year may also demonstrate the manner in which humans interact with the land, thereby showing the type of land use. This connection can easily be measured and analyzed based on RS and GIS techniques. Once the connection is clearly defined and suitable classification schemes are established, it is possible to convert a land cover map to a land use map based on their relationship. This thesis point comes from the first publication.
- **Thesis 2.** I supported that data sources, data structure, image processing, and fusion technique have diverse effects on land use land cover classification efficiency. Within the scope of this study, I demonstrated that:
    a. Using multi-temporal images in a pixel-based classification improved the accuracy of the generated land cover map (OA = 93.86%) compared to those using single-date images (OA = 89.59–90.78%).
    b. By capturing both spectral and spatial information, the segmentation technique and object-based classification could create boundaries between regions with different land use types and then relatively precisely formed land use function regions, which paved the way for producing the final land use map (OA = 93.45%).
    c. The fusion of SAR and optical data based on D-S theory at the decision level yielded better mapping results compared to using single-time single-sensor images or stacked optical-SAR images. The datasets fused at the decision level increased the OA by a range of 0.75% to 2.07% compared to the optical datasets. The fusion of SAR and optical data with their derived textures and indices at the decision level using D-S theory brought the best results.
    d. The integration of SAR and optical products using the layer-stacking technique at the pixel level did not give more power to the classification process. It reduced the accuracy of the mapping result by 4.88% to 6.58% compared to that of the optical datasets.
    e. The inclusion of GLCM textures and spectral indices in the datasets helped improve the mapping results in this study. However, while the effectiveness

of the textures is clear, the contribution of the spectral indices is still controversial.

I also developed a novel approach that is a combination of pixel-based and object-based classifications using a random forest classifier, GIS techniques, and decision rules on multi-temporal RS data. This is the optimal mapping approach given the data availability in the study area and the objective of the study. It provides the ability to effectively extract and translate a land cover map into a land use map. Also, with the long-term availability of Landsat data, it is suitable for the generation of time-series maps, which can be used for land change analysis. In this study, the land use maps generated based on this approach have high accuracy. The extracted maps for the years 1995, 2001, 2005, 2010, 2015, and 2020 gained OA ranging from 88.9% to 93.0%, where the PA and UA of the classes ranged from 70.8% to 100%. This thesis point comes from the first, second, and the third publications.

- **Thesis 3.** I analyzed and confirmed that there were different trends in the area variation of land use types, and there was a large transition from agricultural and unused land to other types of land use in the study area from 1995 to 2020. Mixed residential areas, industrial and commercial zones, recreational regions and green spaces, and mining sites have seen a continuous increase in area, corresponding to 217.9, 109.7, 11.7, and 5.6 km^2, respectively. By contrast, the area of annual croplands continuously decreased, with the total decline being about 455.3 km^2. Meanwhile, the agricultural land for perennial plants increased about 177.7 km^2 in 1995–2010 and decreased about 151.0 km^2 in 2010–2020. The unused land had a strong fluctuation in their area and spatial distribution, whereas the water surface area fluctuated slightly but still increased overall. This thesis point comes from the third publication.
- **Thesis 4.** I measured that the urban area of the province expanded rapidly in the 25 years at an increasing rate. The developed area increased 65 folds, from 5.1 km^2 in 1995 to 332.8 km^2 in 2020. I also proved that the expansion rates were uneven between subregions, and there was a gradual expansion and shift from south to north of the province and spreading to rural districts at an increasingly rapid rate during the study period. It led to a gradual transition from a compact urban form to a dispersed urban form. The factors affecting land use change and urban expansion in Binh Duong province comprised the natural conditions, development history, policies and practices for urbanization, industrialization, and agricultural development, and fluctuation in the prices of products in the market. This thesis point comes from the third publication.
- **Thesis 5.** I confirmed that it is possible to identify variables driving land use change and simulate future land use in terms of quantity and location based on the Markov chain and decision forest algorithm with acceptable reliability. I

discovered that the change was driven by many variables. In which, the drivers of distances to the province centre, district centres, existing residential areas, and main road and mean population density has an impact on the conversion from agricultural land to residential land. Meanwhile, the transition from agricultural land to industrial and commercial areas is driven by the variables of distances to water sources, district centres, existing industrial areas, planned industrial zones, and transportation ports. The selected model was validated with an acceptable accuracy of the hard and soft predictions. The former gained a Kappa of 0.71, a Kappa location of 0.72, a Kappa histogram of 0.99, a fuzzy Kappa of 0.77, and an FoM of 30.77 percent. Meanwhile, the latter achieved an AUC of 0.96. According to the prediction model, in the period from 2020 to 2030, there will be 253.8 km^2 of agricultural land urbanized. The residential areas and the industrial-commercial zones are expected to expand to 395.9 km^2 and 190.8 km^2, respectively. The residential development will be still concentrated in the South of the province and around the centre of the districts. Meanwhile, the new factories are going to fill the existing industrial parks and expand to the new industrial zones in the North and Northeast. This thesis point comes from the fourth publication.

- **Thesis 6.** I measured that by the impacts of land use change and urban expansion, the studied landscape was increasingly decreasing in dominance and increasing diversity and heterogeneity at the landscape level. In addition, the processes of dispersion and aggregation are taking place at the same time. For the class level, the classes of agriculture, mining, and greenspace were increasingly dispersed, but the shape of their patches was becoming more regular. Meanwhile, the urban class has aggregation and dispersion processes occurring parallel similar to those of the entire landscape. The water class had an increase in the dispersion and the irregularity of their patch shape. Finally, the landscape metrics of the unused land fluctuated over time. In general, from 1995 to 2020, the study area experienced an intense change in the direction of increasing the fragmentation and dispersion of natural and semi-natural landscapes. These changes might be largely influenced by the two parallel processes of the urban landscape. In addition, these changing trends are forecast to continue in the next decade. This thesis point comes from the fourth publication.

6.2 Implications

In scientific terms, my detection of the connection between land cover and land use of the study area as well as my evaluations and findings of the mapping performance of the different methods in this study contribute to the existing knowledge on land use land cover study using RS data and GIS techniques. It gives a valuable reference for further studies in the selection of data sources, data structures, image processing techniques, and classification and fusion methods to improve mapping performance. In particular, my novel approach developed in this study, which helps to generate and translate a land

cover map into a land use map from satellite images and GIS techniques, offers many advantages. It promotes the reproducibility and proactivity of the research as well as cost-efficiency and time savings. The output of this approach, i.e., the land cover map and land use map, can be used for different purposes.

In practical terms, by analyzing and discovering land use change and urban expansion, their driving factors, and their effects on the landscape pattern in Binh Duong province of Vietnam, this study reveals a pattern of rapid urbanization in developing countries under the impact of land policies. Some practical lessons can be drawn from them. They can lay the groundwork for further studies on urban planning, land management, and policymaking in Binh Duong provice and other localities not only in Vietnam but also in other countries.

6.3 Limitations, recommendations, and future research

This study still has some limitations due to its research time and resources. They are listed below along with some main recommendations and future research.

- This study mainly used a random forest algorithm for classification. An experiment and comparison of its performance with other classification algorithms are needed.
- The spectral indices extracted from optical bands should be used with caution in future studies due to their controversial effectiveness. Furthermore, only the NDVI and NDWI extracted from optical data were taken into consideration in this study. Therefore, the impact of the other extracted indices should be studied.
- The workflow of the developed approach is transferable; however, the criteria used in this process were formed based on personal experience, visual observations, and experiments (trial and error). Although these values are applicable to neighboring areas, such as localities in Southern Vietnam with characteristics similar to those of the study area, it is suggested that these values need to be re-assessed and revised when applying them to other regions where the land cover and land use characteristics differ from those of this study area. Furthermore, the integration of landscape metrics into classification stages also should be investigated to help the forming of land use function regions.
- The research in this dissertation has not included evaluating the effectiveness of using multi-temporal SAR images, which are not affected by clouds, especially in the context of tropical regions where cloud cover is a challenge. This study will be carried out in the future.
- Other RS data types such as hyperspectral and LiDAR data as well as other data sources should also be considered in further studies. In addition, the classification at the sub-pixel level should also be taken into account.
- Changes in land use and landscape pattern can cause adverse effects on the environmental quality, ecosystem services, and the benefits humans derive from them. These issues will be pursued in further studies.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Prof. Dr. habil. Mucsi László, for his immeasurable guidance, support, patience, enthusiasm, and encouragement during my PhD study. During the past 4 years, I have always felt that working and studying under his supervision was a real blessing for me. Without his comprehensive help, my PhD study would not have been possible.

I would like to thank the governments of Vietnam and Hungary for funding my study through the Stipendium Hungaricum scholarship programme and giving me the opportunity to have wonderful experiences of living and studying in Hungary. I would like to thank the Industrial University of Ho Chi Minh City as well as the Institute for Environmental Science, Engineering, and Management for nominating me to participate in this programme. Thank my colleagues for their kind words and encouragement during my study period.

I would like to thank the Department of Geoinformatics, Physical and Environmental Geography of the University of Szeged for providing me with a great research environment. I would also like to thank all the teachers and staff as well as international and Hungarian PhD students in the department for their friendliness, kindness, and support.

I would like to thank the local authorities and agencies in Binh Duong province for their help in collecting data for the study.

I would like to thank the individuals, organizations, and funding sources, which have been fully acknowledged in my articles, for their support in the publication process.

I would like to thank the Vietnamese community in Szeged for always making me feel warm and familiar like at home. I would also like to thank the Hungarians I have known for their kindness, friendliness, enthusiasm, and honesty. These make me have a special affection for Szeged in particular and for Hungary in general.

Finally and most importantly, I would like to thank my family and besties for their encouragement, understanding, and support, especially during the challenging time of the COVID-19 pandemic. Without them, I would not have been able to overcome all the difficulties and get where I am today. I never said these words because I am not used to saying them, but “I always love you all, and I mean it from the bottom of my heart!”.

References

Abdi AM. 2020. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GIScience Remote Sens [Internet]. 57(1):1–20. https://doi.org/10.1080/15481603.2019.1650447

Abdolalizadeh Z, Ebrahimi A, Mostafazadeh R. 2019. Landscape pattern change in Marakan protected area, Iran. Reg Environ Chang [Internet]. 19(6):1683–1699. https://doi.org/10.1007/s10113-019-01504-9

Acheampong M, Yu Q, Enomah LD, Anchang J, Eduful M. 2018. Land use/cover change in Ghana's oil city: Assessing the impact of neoliberal economic policies and implications for sustainable fevelopment goal number one – A remote sensing and GIS approach. Land use policy [Internet]. 73(November 2016):373–384. https://doi.org/10.1016/j.landusepol.2018.02.019

Adepoju KA, Adelabu SA. 2020. Improving accuracy evaluation of Landsat-8 OLI using image composite and multisource data with Google Earth Engine. Remote Sens Lett [Internet]. 11(2):107–116. https://doi.org/10.1080/2150704X.2019.1690792

ADPC. 2020. The Regional Land Cover Monitoring System [Internet]. [accessed 2020 Aug 1]. https://www.landcovermapping.org/en/landcover/

Aguirre-Gutiérrez J, Seijmonsbergen AC, Duivenvoorden JF. 2012. Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. Appl Geogr [Internet]. 34:29–37. https://doi.org/https://doi.org/10.1016/j.apgeog.2011.10.010

Amponsah O, Blija DK, Ayambire RA, Takyi SA, Mensah H, Braimah I. 2022. Global urban sprawl containment strategies and their implications for rapidly urbanising cities in Ghana. Land use policy [Internet]. 114:105979. https://doi.org/https://doi.org/10.1016/j.landusepol.2022.105979

Andrade-Núñez MJ, Aide TM. 2018. Built-up expansion between 2001 and 2011 in South America continues well beyond the cities. Environ Res Lett [Internet]. 13(8):84006. https://doi.org/10.1088/1748-9326/aad2e3

Arowolo AO, Deng X, Olatunji OA, Obayelu AE. 2018. Assessing changes in the value of ecosystem services in response to land-use/land-cover dynamics in Nigeria. Sci Total Environ [Internet]. 636:597–609. https://doi.org/10.1016/j.scitotenv.2018.04.277

Ban Y, Hu H, Rangel IM. 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land-cover mapping: Object-based and knowledge-based approach. Int J Remote Sens. 31(6):1391–1410. https://doi.org/10.1080/01431160903475415

Bauer T, Steinnocher K. 2001. Per parcel land use classification in urban areas applying a rule-based technique. GeoBIT/GIS [Internet]. 6:24–27. http://giscenter.isu.edu/training/pdf/Geotech_seminar/BauerSteinnocher.pdf

Belgiu M, Drăguţ L. 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J Photogramm Remote Sens [Internet]. 114:24–31. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2016.01.011

Binh Duong Statistical Office. 2016. Binh Duong - 20 years of construction and development. Binh Duong, Vietnam.

Binh Duong Statistical Office. 2019. Statistical yearbook of Binh Duong 2018. Binh Duong, Vietnam.

Binh Duong Statistical Office. 2020. Statistical yearbook of Binh Duong 2019. Binh Duong, Vietnam.

Blaschke T, Merschdorf H. 2014. Geographic information science as a multidisciplinary and

multiparadigmatic field. Cartogr Geogr Inf Sci [Internet]. 41(3):196–213. https://doi.org/10.1080/15230406.2014.905755

Boivin C, Stat.ASSQ. 2020. dst: Using the Theory of Belief Functions [Internet]. https://cran.r-project.org/package=dst

Breiman L. 2001. Random forests. Mach Learn. 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Brown DG, Duh J Der. 2004. Spatial simulation for translating from land use to land cover. Int J Geogr Inf Sci. 18(1):35–60. https://doi.org/10.1080/13658810310001620906

Bui DH, Mucsi L. 2021. From land cover map to land use map: A combined pixel-based and object-based approach using multi-temporal Landsat data, a random forest classifier, and decision rules. Remote Sens [Internet]. 13(9):1700. https://doi.org/10.3390/rs13091700

Bui DH, Mucsi L. 2022. Land-use change and urban expansion in Binh Duong province, Vietnam, from 1995 to 2020. Geocarto Int [Internet]. 37(27):17096–17118. https://doi.org/10.1080/10106049.2022.2123564

Bui LT, Nguyen PH, Nguyen DCM. 2020. Model for assessing health damage from air pollution in quarrying area – Case study at Tan Uyen quarry, Ho Chi Minh megapolis, Vietnam. Heliyon [Internet]. 6(9):e05045. https://doi.org/10.1016/j.heliyon.2020.e05045

Cai G, Ren H, Yang L, Zhang N, Du M, Wu C. 2019. Detailed urban land use land cover classification at the metropolitan scale using a three-layer classification scheme. Sensors (Switzerland). 19(14). https://doi.org/10.3390/s19143120

Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F. 2018. Geomatic Approaches for Modeling Land Change Scenarios. [place unknown]: Springer Cham. https://doi.org/https://doi.org/10.1007/978-3-319-60801-3

Cao H, Liu J, Chen J, Gao J, Wang G, Zhang W. 2019. Spatiotemporal patterns of urban land use change in typical cities in the Greater Mekong Subregion (GMS). Remote Sens. 11(7). https://doi.org/10.3390/rs11070801

Carranza-García M, García-Gutiérrez J, Riquelme JC. 2019. A framework for evaluating land use and land cover classification using convolutional neural networks. Remote Sens. 11(3). https://doi.org/10.3390/rs11030274

Ceccarelli T, Smiraglia D, Bajocco S, Rinaldo S, Angelis A De, Salvati L, Perini L. 2013. Land cover data from Landsat single-date imagery: an approach integrating pixel-based and object-based classifiers. Eur J Remote Sens [Internet]. 46(1):699–717. https://doi.org/10.5721/EuJRS20134641

Chen Q, Whitbrook A, Aickelin U, Roadknight C. 2014. Data classification using the Dempster-Shafer method. J Exp Theor Artif Intell. 26(4):493–517. https://doi.org/10.1080/0952813X.2014.886301

Chen Yuehong, Zhou Y, Ge Y, An R, Chen Yu. 2018. Enhancing land cover mapping through integration of pixel-based and object-based classifications from remotely sensed imagery. Remote Sens. 10(1). https://doi.org/10.3390/rs10010077

Chen YY, Huang W, Wang WH, Juang JY, Hong JS, Kato T, Luyssaert S. 2019. Reconstructing Taiwan's land cover changes between 1904 and 2015 from historical maps and satellite images. Sci Rep. 9(1):1–12. https://doi.org/10.1038/s41598-019-40063-1

Chuvieco E, Huete A. 2016. Fundamentals of satellite remote sensing: An environmental approach. Boca Raton, FL: CRC Press.

Cihlar J, Jansen L. 2001. From Land Cover to Land Use: A Methodology for Efficient Land

Use Mapping over Large Areas. Prof Geogr. 53(2):275–289. https://doi.org/10.1080/00330124.2001.9628460

Comber AJ, Wadsworth RA, Fisher PF. 2008. Using semantics to clarify the conceptual confusion between land cover and land use: the example of 'forest.' J Land Use Sci [Internet]. 3(2–3):185–198. https://doi.org/10.1080/17474230802434187

Congalton RG, Green K. 2019. Assessing the Accuracy of Remotely Sensed Data Principles and Practices, Third Edition. Boca Raton: CRC Press.

Dadashpoor H, Azizi P, Moghadasi M. 2019. Land use change, urbanization, and change in landscape pattern in a metropolitan area. Sci Total Environ [Internet]. [accessed 2020 Sep 24] 655:707–719. https://doi.org/10.1016/j.scitotenv.2018.11.267

van Delden H, van Vliet J, Rutledge DT, Kirkby MJ. 2011. Comparison of scale and scaling issues in integrated land-use models for policy support. Agric Ecosyst Environ [Internet]. 142(1):18–28. https://doi.org/https://doi.org/10.1016/j.agee.2011.03.005

Dempster AP. 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping. Ann Math Stat. 38(2):325–339. https://doi.org/10.1214/AOMS/1177698950

Department of Natural Resources and Environment of Binh Duong Province. 2020. Report on environmental status of Binh Duong province in the period 2016 - 2020 (Vietnamese). Binh Duong, Vietnam.

Disperati L, Virdis SGP. 2015. Assessment of land-use and land-cover changes from 1965 to 2014 in Tam Giang-Cau Hai Lagoon, central Vietnam. Appl Geogr [Internet]. 58:48–64. https://doi.org/https://doi.org/10.1016/j.apgeog.2014.12.012

Do QC, Nguyen H Le, Hoang XD. 2020. From Industrial Policy to Economic and Social Upgrading in Vietnam. Singapore.

Eastman JR. 2020a. TerrSet 2020 Geospatial Monitoring and Modeling System Manual. [place unknown]: ClarkLabs.

Eastman JR. 2020b. TerrSet 2020 Geospatial Monitoring and Modeling System Tutorial. [place unknown]: ClarkLabs.

ESRI. 2022. What is GIS [Internet]. [accessed 2022 Oct 4]. https://www.esri.com/en-us/what-is-gis/overview

Estoque RC, Murayama Y. 2016. Quantifying landscape pattern and ecosystem service value changes in four rapidly urbanizing hill stations of Southeast Asia. Landsc Ecol. 31(7):1481–1507. https://doi.org/10.1007/s10980-016-0341-6

Feng Q, Yang J, Zhu D, Liu J, Guo H, Bayartungalag B, Li B. 2019. Integrating multitemporal Sentinel-1/2 data for coastal land cover classification using a multibranch convolutional neural network: A case of the Yellow River Delta. Remote Sens. 11(9). https://doi.org/10.3390/rs11091006

Fenta AA, Yasuda H, Haregeweyn N, Belay AS, Hadush Z, Gebremedhin MA, Mekonnen G. 2017. The dynamics of urban expansion and land use/land cover changes using remote sensing and spatial metrics: the case of Mekelle City of northern Ethiopia. Int J Remote Sens [Internet]. 38(14):4107–4129. https://doi.org/10.1080/01431161.2017.1317936

Filipponi F. 2019. Sentinel-1 GRD Preprocessing Workflow. Proceedings. 18(1):11. https://doi.org/10.3390/ecrs-3-06201

Fonteh ML, Theophile F, Cornelius ML, Main R, Ramoelo A, Cho MA. 2016. Assessing the Utility of Sentinel-1 C Band Synthetic Aperture Radar Imagery for Land Use Land Cover Classification in a Tropical Coastal Systems When Compared with Landsat 8. J Geogr Inf Syst. 08(04):495–505. https://doi.org/10.4236/jgis.2016.84041

Fox J, Castella JC. 2013. Expansion of rubber (Hevea brasiliensis) in Mainland Southeast Asia: What are the prospects for smallholders? J Peasant Stud. 40(1):155–170. https://doi.org/10.1080/03066150.2012.750605

de Furtado LFA, Silva TSF, Fernandes PJF, de Novo EML de M. 2015. Land cover classification of Lago Grande de Curuai floodplain (Amazon, Brazil) using multi-sensor and image fusion techniques. Acta Amaz. 45(2):195–202. https://doi.org/10.1590/1809-4392201401439

GADM. 2020. GADM maps and data [Internet]. [accessed 2020 Aug 1]. https://gadm.org/

General Statistics Office of Vietnam. 2020. Population [Internet]. [accessed 2020 Dec 8]. https://www.gso.gov.vn/en/population/

General Statistics Office of Vietnam. 2021. Health, Culture, Sport, Living standards, Social order, Safety and Environment [Internet]. [accessed 2021 Oct 25]. https://www.gso.gov.vn/en/health-culture-sport-living-standards-social-order-safety-and-environment/

Gergel SE, Turner MG. 2017. Learning Landscape Ecology A Practical Guide to Concepts and Techniques. [place unknown]: Springer-Verlag New York.

Ghosh A, Sharma R, Joshi PK. 2014. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. Appl Geogr [Internet]. 48:31–41. https://doi.org/https://doi.org/10.1016/j.apgeog.2014.01.003

Giri C, Pengra B, Long J, Loveland TR. 2013. Next generation of global land cover characterization, mapping, and monitoring. Int J Appl Earth Obs Geoinf [Internet]. 25:30–37. https://doi.org/https://doi.org/10.1016/j.jag.2013.03.005

Giri CP. 2012. Remote Sensing of Land Use and Land Cover Principles and Applications (1st ed.). Giri CP, editor. Boca Raton: CRC press. https://doi.org/https://doi.org/10.1201/b11964

Di Gregorio A. 2005. Land cover classification system: classification concepts and user manual. [place unknown]: Food & Agriculture Org.

Grigoraș G, Urițescu B. 2019. Land Use/Land Cover changes dynamics and their effects on Surface Urban Heat Island in Bucharest, Romania. Int J Appl Earth Obs Geoinf. 80(February):115–126. https://doi.org/10.1016/j.jag.2019.03.009

De Gruijter J, Brus DJ, Bierkens MFP, Knotters M. 2006. Sampling for natural resource monitoring. Berlin, Heidelberg, Germany: Springer. https://doi.org/https://doi.org/10.1007/3-540-33161-1

Guan X, Liao S, Bai J, Wang F, Li Z, Wen Q, He J, Chen T. 2017. Urban land-use classification by combining high-resolution optical and long-wave infrared images. Geo-Spatial Inf Sci [Internet]. 20(4):299–308. https://doi.org/10.1080/10095020.2017.1403731

Gudmann A, Csikós N, Szilassi P, Mucsi L. 2020. Improvement in satellite image-based land cover classification with landscape metrics. Remote Sens. 12(21):1–21. https://doi.org/10.3390/rs12213580

Gulácsi A, Kovács F. 2020. Sentinel-1-imagery-based high-resolution water cover detection on wetlands, aided by Google Earth Engine. Remote Sens. 12(10):1–20. https://doi.org/10.3390/rs12101614

Ha T V., Tuohy M, Irwin M, Tuan P V. 2020. Monitoring and mapping rural urbanization and land use changes using Landsat data in the northeast subtropical region of Vietnam. Egypt J Remote Sens Sp Sci [Internet]. [accessed 2020 Aug 9] 23(1):11–19. https://doi.org/10.1016/j.ejrs.2018.07.001

Hagen-Zanker A, Straatman B, Uljee I. 2005. Further developments of a fuzzy set map comparison approach. Int J Geogr Inf Sci [Internet]. 19(7):769–785. https://doi.org/10.1080/13658810500072137

Hagen A. 2002. Multi-method assessment of map similarity. 5th Agil Conf Geogr Inf Sci.:1–8.

Hagen A. 2003. Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci [Internet]. 17(3):235–249. https://doi.org/10.1080/13658810210157822

Hall DL, Llinas J. 1997. An introduction to multisensor data fusion. Proc IEEE. 85(1):6–23. https://doi.org/10.1109/5.554205

Henits L, Jürgens C, Mucsi L. 2016. Seasonal multitemporal land-cover classification and change detection analysis of Bochum, Germany, using multitemporal Landsat TM data. Int J Remote Sens. 37(15):3439–3454. https://doi.org/10.1080/01431161.2015.1125558

Henits L, Mucsi L, Liska CM. 2017. Monitoring the changes in impervious surface ratio and urban heat island intensity between 1987 and 2011 in Szeged, Hungary. Environ Monit Assess. 189(2). https://doi.org/10.1007/s10661-017-5779-8

Hu J, Ghamisi P, Zhu X. 2018. Feature Extraction and Selection of Sentinel-1 Dual-Pol Data for Global-Scale Local Climate Zone Classification. ISPRS Int J Geo-Information [Internet]. [accessed 2020 May 16] 7(9):379. https://doi.org/10.3390/ijgi7090379

Huang IY, James K, Thamthanakoon N, Pinitjitsamut P, Rattanamanee N, Pinitjitsamut M, Yamklin S, Lowenberg-DeBoer J. 2022. Economic outcomes of rubber-based agroforestry systems: a systematic review and narrative synthesis. Agrofor Syst [Internet].:1–20. https://doi.org/10.1007/s10457-022-00734-x

Hunt GC, Rosentrater L, Frolking S, Iii BM, Hurtt GC, Rosentrater L, Frolking S, Moore III B. 2001. Linking remote-sensing estimates of land cover and census statistics on land use to produce maps of land use of the conterminous United States. Global Biogeochem Cycles [Internet]. 15(3):673–685. https://doi.org/10.1029/2000GB001299

Hurni K, Fox J. 2018. The expansion of tree-based boom crops in mainland Southeast Asia: 2001 to 2014. J Land Use Sci [Internet]. 13(1–2):198–219. https://doi.org/10.1080/1747423X.2018.1499830

Islam K, Rahman MF, Jashimuddin M. 2018. Modeling land use change using Cellular Automata and Artificial Neural Network: The case of Chunati Wildlife Sanctuary, Bangladesh. Ecol Indic. https://doi.org/10.1016/j.ecolind.2018.01.047

Jansen LJM, Di Gregorio A. 2003. Land-use data collection using the “land cover classification system”: Results from a case study in Kenya. Land use policy. 20(2):131–148. https://doi.org/10.1016/S0264-8377(02)00081-9

JAXA EORC. 2020. Homepage of High-Resolution Land Use and Land Cover Map Products [Internet]. [accessed 2020 Aug 1]. https://www.eorc.jaxa.jp/ALOS/en/lulc/lulc_index.htm

Ji L, Gong P, Geng X, Zhao Y. 2015. Improving the accuracy of the water surface cover type in the 30 m FROM-GLC product. Remote Sens. 7(10):13507–13527. https://doi.org/10.3390/rs71013507

Jia Y, Ge Y, Ling F, Guo X, Wang J, Wang L, Chen Y, Li X. 2018. Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. Remote Sens. 10(3). https://doi.org/10.3390/rs10030446

Jiao L. 2015. Urban land density function: A new method to characterize urban expansion. Landsc Urban Plan [Internet]. 139:26–39.

https://doi.org/https://doi.org/10.1016/j.landurbplan.2015.02.017

Johnson BA, Jozdani SE. 2018. Identifying generalizable image segmentation parameters for urban land cover mapping through meta-analysis and regression tree modeling. Remote Sens. 10(1). https://doi.org/10.3390/rs10010073

Kantakumar LN, Kumar S, Schneider K. 2016. Spatiotemporal urban expansion in Pune metropolis, India using remote sensing. Habitat Int [Internet]. 51:11–22. https://doi.org/10.1016/j.habitatint.2015.10.007

Kertész Á, Křeček J. 2019. Landscape degradation in the world and in Hungary. Hungarian Geogr Bull. 68(3):201–221. https://doi.org/10.15201/hungeobull.68.3.1

Kim JH. 2015. Crossing-over between land cover and land use: Exploring spatially varying relationships in two large US metropolitan areas. Appl Geogr [Internet]. 60:37–45. https://doi.org/10.1016/j.apgeog.2015.03.002

Kosztra B, Büttner G, Hazeu G, Arnold S. 2017. Updated CLC illustrated nomenclature guidelines. [place unknown].

Kovács Z, Farkas ZJ, Egedy T, Kondor AC, Szabó B, Lennert J, Baka D, Kohán B. 2019. Urban sprawl and land conversion in post-socialist cities: The case of metropolitan Budapest. Cities [Internet]. 92:71–81. https://doi.org/https://doi.org/10.1016/j.cities.2019.03.018

Kulkarni SC, Rege PP. 2020. Pixel level fusion techniques for SAR and optical images: A review. Inf Fusion [Internet]. 59(July 2019):13–29. https://doi.org/10.1016/j.inffus.2020.01.003

Kusakabe K, Chanthoumphone C. 2021. Transition From Subsistence Agriculture to Rubber Plantations in Northern Laos: Analysis of Household Livelihood Strategies by Ethnicity and Gender. SAGE Open [Internet]. 11(2):21582440211011464. https://doi.org/10.1177/21582440211011463

Lambin EF, Meyfroidt P. 2010. Land use transitions: Socio-ecological feedback versus socio-economic change. Land use policy [Internet]. 27(2):108–118. https://doi.org/https://doi.org/10.1016/j.landusepol.2009.09.003

Le VH. 2019. The process of urbanization in Binh Duong province 1986 - 2010 (Vietnamese). Ho Chi Minh City, Vietnam: Vietnam National University-HoChiMinh City - University of Social Sciences and Humanities.

Le VN, Truong HT, Ton Nu QT, Le VH, Nguyen NK. 2019. Binh Duong urbanization in the period 1997 - 2017 (Vietnamese). Binh Duong, Vietnam: Thu Dau Mot University.

Lennert J, Farkas JZ, Kovács AD, Molnár A, Módos R, Baka D, Kovács Z. 2020. Measuring and Predicting Long-Term Land Cover Changes in the Functional Urban Area of Budapest. Sustainability. 12(8). https://doi.org/10.3390/su12083331

Li H, Wang C, Zhong C, Su A, Xiong C, Wang J, Liu J. 2017. Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index. Remote Sens [Internet]. [accessed 2020 May 16] 9(3):249. https://doi.org/10.3390/rs9030249

Li X, Chen G, Liu X, Liang X, Wang S, Chen Y, Pei F, Xu X. 2017. A New Global Land-Use and Land-Cover Change Product at a 1-km Resolution for 2010 to 2100 Based on Human–Environment Interactions. Ann Am Assoc Geogr [Internet]. 107(5):1040–1059. https://doi.org/10.1080/24694452.2017.1303357

Liaw A, Wiener M. 2002. Classification and Regression by randomForest. R News [Internet]. 2(3):18–22. https://cran.r-project.org/doc/Rnews/

Lin T, Xue X, Shi L, Gao L. 2013. Urban spatial expansion and its impacts on island ecosystem services and landscape pattern: A case study of the island city of Xiamen, Southeast China. Ocean Coast Manag [Internet]. 81:90–96.

https://doi.org/10.1016/j.ocecoaman.2012.06.014

Linh NHK, Erasmi S, Kappas M. 2012. Quantifying Land Use/Cover Change And Landscape Fragmentation In Danang City, Vietnam: 1979-2009. ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci [Internet]. XXXIX-B8:501–506. https://doi.org/10.5194/isprsarchives-XXXIX-B8-501-2012

Liu D, Xia F. 2010. Assessing object-based classification: advantages and limitations. Remote Sens Lett [Internet]. 1(4):187–194. https://doi.org/10.1080/01431161003743173

Liu Y, Gong W, Hu X, Gong J. 2018. Forest type identification with random forest using Sentinel-1A, Sentinel-2A, multi-temporal Landsat-8 and DEM data. Remote Sens. 10(6):1–25. https://doi.org/10.3390/rs10060946

Lu D, Li G, Moran E, Dutra L, Batistella M. 2011. A Comparison of Multisensor Integration Methods for Land Cover Classification in the Brazilian Amazon. GIScience Remote Sens [Internet]. 48(3):345–370. https://doi.org/10.2747/1548-1603.48.3.345

Lu D, Li G, Moran E, Dutra L, Batistella M. 2014. The roles of textural images in improving land-cover classification in the Brazilian Amazon. Int J Remote Sens. 35(24):8188–8207. https://doi.org/10.1080/01431161.2014.980920

Lu D, Weng Q. 2007. A survey of image classification methods and techniques for improving classification performance. Int J Remote Sens. 28(5):823–870. https://doi.org/10.1080/01431160600746456

Mahendra A, Seto KC. 2019. Upward and outward growth: managing urban expansion for more equitable cities in the global south. Work Pap.

Mahtta R, Fragkias M, Güneralp B, Mahendra A, Reba M, Wentz EA, Seto KC. 2022. Urban land expansion: the role of population and economic growth for 300+ cities. npj Urban Sustain [Internet]. 2(1):5. https://doi.org/10.1038/s42949-022-00048-y

Malinverni ES, Tassetti AN, Mancini A, Zingaretti P, Frontoni E, Bernardini A. 2011. Hybrid object-based approach for land use/land cover mapping using high spatial resolution imagery. Int J Geogr Inf Sci [Internet]. 25(6):1025–1043. https://doi.org/10.1080/13658816.2011.566569

Marengo MC. 2015. Sprawl and Density, Towards a Dispersed Urban form The Case of Córdoba City - Argentina. J Eng Archit. 3(2):45–56. https://doi.org/10.15640/jea.v3n2a5

Mas JF. 2018. Receiver Operating Characteristic (ROC) Analysis. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F, editors. Geomat Approaches Model L Chang Scenar [Internet]. Cham: Springer International Publishing; p. 465–467. https://doi.org/10.1007/978-3-319-60801-3_30

McGarigal K, Cushman SA, Ene E. 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Comput Softw Progr Prod by Authors Univ Massachusetts, Amherst http//www.umass.edu/landeco/research/fragstats/fragstats.html.

Megahed Y, Cabral P, Silva J, Caetano M. 2015. Land cover mapping analysis and urban growth modelling using remote sensing techniques in greater Cairo region-Egypt. ISPRS Int J Geo-Information. 4(3):1750–1769. https://doi.org/10.3390/ijgi4031750

Merriam-Webster. 2022. Urban sprawl. Merriam-Webster.com Dict [Internet]. [accessed 2022 Jul 15]. https://www.merriam-webster.com/dictionary/urban sprawl

Mertes CM, Schneider A, Sulla-Menashe D, Tatem AJ, Tan B. 2015. Detecting change in urban areas at continental scales with MODIS data. Remote Sens Environ [Internet]. 158:331–347. https://doi.org/https://doi.org/10.1016/j.rse.2014.09.023

Mezaal MR, Pradhan B, Rizeei HM. 2018. Improving landslide detection from airborne laser

scanning data using optimized Dempster-Shafer. Remote Sens. 10(7). https://doi.org/10.3390/rs10071029

Mfuka C, Byamukama E, Zhang X. 2020. Spatiotemporal characteristics of white mold and impacts on yield in soybean fields in South Dakota. Geo-Spatial Inf Sci [Internet]. 23(2):182–193. https://doi.org/10.1080/10095020.2020.1712265

Minister of Natural Resources and Environment of Vietnam. 2018. Circular No. 27/2018/TT-BTNMT Regulation on land statistics, inventory and current land use mapping. Vietnam.

Mishra VN, Rai PK, Prasad R, Punia M, Nistor MM. 2018. Prediction of spatio-temporal land use/land cover dynamics in rapidly developing Varanasi district of Uttar Pradesh, India, using geospatial approach: a comparison of hybrid models. Appl Geomatics. 10(3):257–276. https://doi.org/10.1007/s12518-018-0223-5

Msofe NK, Sheng L, Lyimo J. 2019. Land Use Change Trends and Their Driving Forces in the Kilombero Valley Floodplain, Southeastern Tanzania. Sustain . 11(2). https://doi.org/10.3390/su11020505

Mucsi L, Liska CM, Henits L, Tobak Z, Csendes B, Nagy L. 2017. The evaluation and application of an urban land cover map with image data fusion and laboratory measurements. Hungarian Geogr Bull. 66(2):145–156. https://doi.org/10.15201/hungeobull.66.2.4

Nampak H, Pradhan B, Mojaddadi Rizeei H, Park HJ. 2018. Assessment of land cover and land use change impact on soil loss in a tropical catchment by using multitemporal SPOT-5 satellite images and Revised Universal Soil Loss Equation model. L Degrad Dev. 29(10):3440–3455. https://doi.org/10.1002/ldr.3112

Netzel P, Stepinski TF. 2015. Pattern-Based Assessment of Land Cover Change on Continental Scale With Application to NLCD 2001–2006. IEEE Trans Geosci Remote Sens. 53(4):1773–1781. https://doi.org/10.1109/TGRS.2014.2348715

Nguyen HTT, Doan TM, Tomppo E, McRoberts RE. 2020. Land use/land cover mapping using multitemporal sentinel-2 imagery and four classification methods-A case study from Dak Nong, Vietnam. Remote Sens [Internet]. [accessed 2020 Aug 9] 12(9):1–27. https://doi.org/10.3390/RS12091367

Nguyen LH, Joshi DR, Clay DE, Henebry GM. 2020. Characterizing land cover/land use from multiple years of Landsat and MODIS time series: A novel approach using land surface phenology modeling and random forest classifier. Remote Sens Environ [Internet]. 238(December 2018):111017. https://doi.org/10.1016/j.rse.2018.12.016

Nguyen Q, Kim D-C. 2020. Reconsidering rural land use and livelihood transition under the pressure of urbanization in Vietnam: A case study of Hanoi. Land use policy [Internet]. 99:104896. https://doi.org/https://doi.org/10.1016/j.landusepol.2020.104896

Nguyen VH, Ton Nu QT, Nguyen VS, Le VH, Truong TT, Nguyen TTH, Nguyen TXT, Nguyen NK. 2019. Context of urbanization in Binh Duong (Vietnamese). Binh Duong, Vietnam: Thu Dau Mot University.

Noi PT, Kappas M. 2017. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors [Internet]. 18(1). https://doi.org/10.3390/s18010018

Nor ANM, Corstanje R, Harris JA, Brewer T. 2017. Impact of rapid urban expansion on green space structure. Ecol Indic [Internet]. 81(September 2016):274–284. https://doi.org/10.1016/j.ecolind.2017.05.031

Nourqolipour R, Shariff ARBM, Balasundram SK, Ahmad NB, Sood AM, Buyong T. 2016. Predicting the Effects of Urban Development on Land Transition and Spatial Patterns

of Land Use in Western Peninsular Malaysia. Appl Spat Anal Policy [Internet]. 9(1):1–19. https://doi.org/10.1007/s12061-014-9128-9

Party Committee of Binh Duong Province. 1997. Documents of the 6th Congress of the Party Committee of Binh Duong Province (Vietnamese). Binh Duong, Vietnam.

Party Committee of Binh Duong Province. 2001. Documents of the 7th Congress of the Party Committee of Binh Duong Province (Vietnamese). Binh Duong, Vietnam.

Peng K, Jiang W, Deng Y, Liu Y, Wu Z, Chen Z. 2020. Simulating wetland changes under different scenarios based on integrating the random forest and CLUE-S models: A case study of Wuhan Urban Agglomeration. Ecol Indic [Internet]. 117(July):106671. https://doi.org/10.1016/j.ecolind.2020.106671

Peng W, Wang G, Zhou J, Zhao J, Yang C. 2015. Studies on the temporal and spatial variations of urban expansion in Chengdu, western China, from 1978 to 2010. Sustain Cities Soc [Internet]. 17:141–150. https://doi.org/10.1016/j.scs.2015.03.004

People's Committee of Binh Duong Province. 2010. Decision No. 4614/QD-UBND on approving the development planning of agriculture, forestry and fishery in Binh Duong province until 2020 (Vietnamese). Vietnam.

People's Committee of Binh Duong Province. 2012. Decision No. 1701/QD-UBND on approving the general planning project to develop Binh Duong metropolis until 2020, vision to 2030 (Vietnamese). Vietnam.

People's Committee of Binh Duong Province. 2018. Decision No. 157/QD-UBND on approving the adjustment of the planning for the development of agriculture, forestry, and fishery in Binh Duong province until 2020, supplementing the planning to 2025 (Vietnamese). Vietnam.

Pham HM, Yamaguchi Y. 2011. Urban growth and change analysis using remote sensing and spatial metrics from 1975 to 2003 for Hanoi, Vietnam. Int J Remote Sens [Internet]. 32(7):1901–1915. https://doi.org/10.1080/01431161003639652

Phan D, Coxhead I. 2010. Inter-provincial migration and inequality during Vietnam's transition. J Dev Econ [Internet]. 91(1):100–112. https://doi.org/10.1016/j.jdeveco.2009.06.008

Pohl C, van Genderen J. 2016. Remote sensing image fusion: A practical guide. London: CRC Press.

Pontius RG. 2000. Quantification error versus location error in comparison of categorical maps. Photogramm Eng Remote Sensing. 66(8):1011–1016.

Pontius RG, Boersma W, Castella J-C, Clarke K, de Nijs T, Dietzel C, Duan Z, Fotsing E, Goldstein N, Kok K, et al. 2008. Comparing the input, output, and validation maps for several models of land change. Ann Reg Sci [Internet]. 42(1):11–37. https://doi.org/10.1007/s00168-007-0138-2

Pontius RG, Millones M. 2011. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remote Sens. 32(15):4407–4429. https://doi.org/10.1080/01431161.2011.552923

Pratama AP, Yudhistira MH, Koomen E. 2022. Highway expansion and urban sprawl in the Jakarta Metropolitan Area. Land use policy [Internet]. 112:105856. https://doi.org/https://doi.org/10.1016/j.landusepol.2021.105856

Prime Minister of Vietnam. 2007. Decision No. 81/2007/QD-TTg on approving the master plan for socio-economic development of Binh Duong province until 2020 (Vietnamese). Vietnam.

Prime Minister of Vietnam. 2014. Decision No. 893/QD-TTg on approving the adjustment of the master plan for socio-economic development of Binh Duong province until 2020,

supplementing the planning to 2025 (Vietnamese). Vietnam.

Prins AJ, Van Niekerk A. 2020. Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms. Geo-Spatial Inf Sci [Internet]. 24(2):1–13. https://doi.org/10.1080/10095020.2020.1782776

Quan Y, Tong Y, Feng W, Dauphin G, Huang W, Xing M. 2020. A novel image fusion method of multi-spectral and sar images for land cover classification. Remote Sens. 12(22):1–25. https://doi.org/10.3390/rs12223801

Rahman MT. 2016. Detection of land use/land cover changes and urban sprawl in Al-Khobar, Saudi Arabia: An analysis of multi-temporal remote sensing data. ISPRS Int J Geo-Information. 5(2). https://doi.org/10.3390/ijgi5020015

Ran YH, Li X, Lu L, Li ZY. 2012. Large-scale land cover mapping with the integration of multi-source information based on the Dempster-Shafer theory. Int J Geogr Inf Sci. 26(1):169–191. https://doi.org/10.1080/13658816.2011.577745

Rawat JS, Kumar M. 2015. Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India. Egypt J Remote Sens Sp Sci [Internet]. 18(1):77–84. https://doi.org/10.1016/j.ejrs.2015.02.002

Reilly MK, O'Mara MP, Seto KC. 2009. From Bangalore to the Bay Area: Comparing transportation and activity accessibility as drivers of urban growth. Landsc Urban Plan [Internet]. 92(1):24–33. https://doi.org/https://doi.org/10.1016/j.landurbplan.2009.02.001

Rimal B, Zhang L, Keshtkar H, Wang N, Lin Y. 2017. Monitoring and modeling of spatiotemporal urban expansion and land-use/land-cover change using integrated Markov chain cellular automata model. ISPRS Int J Geo-Information. 6(9). https://doi.org/10.3390/ijgi6090288

Rizeei HM, Saharkhiz MA, Pradhan B, Ahmad N. 2016. Soil erosion prediction based on land cover dynamics at the Semenyih watershed in Malaysia using LTM and USLE models. Geocarto Int [Internet]. 31(10):1158–1177. https://doi.org/10.1080/10106049.2015.1120354

Ruiz Hernandez IE, Shi W. 2018. A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. Int J Remote Sens [Internet]. 39(4):1175–1198. https://doi.org/10.1080/01431161.2017.1395968

Sajikumar N, Remya RS. 2015. Impact of land cover and land use change on runoff characteristics. J Environ Manage [Internet]. 161:460–468. https://doi.org/10.1016/j.jenvman.2014.12.041

Sánchez-Cuervo AM, Aide TM, Clark ML, Etter A. 2012. Land Cover Change in Colombia: Surprising Forest Recovery Trends between 2001 and 2010. PLoS One [Internet]. 7(8):e43943. https://doi.org/10.1371/journal.pone.0043943

Saxena A, Jat MK. 2019. Capturing heterogeneous urban growth using SLEUTH model. Remote Sens Appl Soc Environ [Internet]. 13(December 2018):426–434. https://doi.org/10.1016/j.rsase.2018.12.012

Schiller G, Bimesmeier T, Pham ATV. 2020. Method for quantifying supply and demand of construction minerals in urban regions-A case study of hanoi and its Hinterland. Sustain. 12(11). https://doi.org/10.3390/su12114358

Schmitt M, Zhu XX. 2016. Data Fusion and Remote Sensing: An ever-growing relationship. IEEE Geosci Remote Sens Mag. 4(4):6–23. https://doi.org/10.1109/MGRS.2016.2561021

Schoeman F, Newby TS, Thomson MW, Van den Berg EC. 2013. South African National

Land-Cover Change Map. South African J Geomatics. 2(2):94–105.

Serra P, Pons X, Saurí D. 2008. Land-cover and land-use change in a Mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. Appl Geogr [Internet]. 28(3):189–209. https://doi.org/https://doi.org/10.1016/j.apgeog.2008.02.001

Shackelford AK, Davis CH. 2003. A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. IEEE Trans Geosci Remote Sens. 41(10):2354–2363. https://doi.org/10.1109/TGRS.2003.815972

Shafer G. 1976. A mathematical theory of evidence. Princeton: Princeton University Press.

Shalaby A, Tateishi R. 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. Appl Geogr [Internet]. 27(1):28–41. https://doi.org/https://doi.org/10.1016/j.apgeog.2006.09.004

Shao Y, Lunetta RS. 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J Photogramm Remote Sens [Internet]. 70:78–87. https://doi.org/10.1016/j.isprsjprs.2012.04.001

Shao Z, Cheng G, Li D, Huang X, Lu Z, Liu J. 2021. Spatio-temporal-spectral-angular observation model that integrates observations from UAV and mobile mapping vehicle for better urban mapping. Geo-Spatial Inf Sci [Internet]. 24(4):615–629. https://doi.org/10.1080/10095020.2021.1961567

Shao Z, Fu H, Fu P, Yin L. 2016. Mapping urban impervious surface by fusing optical and SAR data at the decision level. Remote Sens. 8(11):1–21. https://doi.org/10.3390/rs8110945

Shao Z, Sumari NS, Portnov A, Ujoh F, Musakwa W, Mandela PJ. 2021. Urban sprawl and its impact on sustainable urban development: a combination of remote sensing and social media data. Geo-spatial Inf Sci [Internet]. 24(2):241–255. https://doi.org/10.1080/10095020.2020.1787800

Shooshtari SJ, Gholamalifard M. 2015. Scenario-based land cover change modeling and its implications for landscape pattern analysis in the Neka Watershed, Iran. Remote Sens Appl Soc Environ [Internet]. 1:1–19. https://doi.org/https://doi.org/10.1016/j.rsase.2015.05.001

Singh SK, Laari PB, Mustak S, Srivastava PK, Szabó S. 2018. Modelling of land use land cover change using earth observation data-sets of Tons River Basin, Madhya Pradesh, India. Geocarto Int. 33(11):1202–1222. https://doi.org/10.1080/10106049.2017.1343390

Solberg AHS. 2006. Data fusion for remote sensing applications. Signal image Process Remote Sens.:249–271.

Steinhausen MJ, Wagner PD, Narasimhan B, Waske B. 2018. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. Int J Appl Earth Obs Geoinf [Internet]. 73(April):595–604. https://doi.org/10.1016/j.jag.2018.08.011

Stroesser L, Penot É, Michel I, Tongkaemkaew U, Chambon B. 2018. Income Diversification for Rubber Farmers Through Agroforestry Practices. How to Withstand Rubber Price Volatility in Phatthalung Province, Thailand. Rev Int des études du développement [Internet]. 235(3):117–145. https://doi.org/10.3917/ried.235.0117

Su S, Wang Y, Luo F, Mai G, Pu J. 2014. Peri-urban vegetated landscape pattern changes in relation to socioeconomic development. Ecol Indic [Internet]. 46:477–486. https://doi.org/https://doi.org/10.1016/j.ecolind.2014.06.044

Su Y, Sujakhu NM, Smith A. 2022. Gendered impacts of falling rubber prices: Changing livelihood strategies in China's rubber hearltand. ICRAF Work Pap.

Sumari NS, Cobbinah PB, Ujoh F, Xu G. 2020. On the absurdity of rapid urbanization: Spatio-temporal analysis of land-use changes in Morogoro, Tanzania. Cities [Internet]. 107:102876. https://doi.org/https://doi.org/10.1016/j.cities.2020.102876

Sumari NS, Xu G, Ujoh F, Korah PI, Ebohon OJ, Lyimo NN. 2019. A Geospatial Approach to Sustainable Urban Planning: Lessons for Morogoro Municipal Council, Tanzania. Sustain . 11(22). https://doi.org/10.3390/su11226508

Tabib Mahmoudi F, Arabsaeedi A, Alavipanah SK. 2019. Feature-Level Fusion of Landsat 8 Data and SAR Texture Images for Urban Land Cover Classification. J Indian Soc Remote Sens [Internet]. 47(3):479–485. https://doi.org/10.1007/s12524-018-0914-8

Tadese M, Kumar L, Koech R, Kogo BK. 2020. Mapping of land-use/land-cover changes and its dynamics in Awash River Basin using remote sensing and GIS. Remote Sens Appl Soc Environ [Internet]. 19(June):100352. https://doi.org/10.1016/j.rsase.2020.100352

Tang J, Li Y, Cui S, Xu L, Ding S, Nie W. 2020. Linking land-use change, landscape patterns, and ecosystem services in a coastal watershed of southeastern China. Glob Ecol Conserv [Internet]. 23:e01177. https://doi.org/10.1016/j.gecco.2020.e01177

Tapiador FJ, Casanova JL. 2003. Land use mapping methodology using remote sensing for the regional planning directives in Segovia, Spain. Landsc Urban Plan [Internet]. 62(2):103–115. https://doi.org/https://doi.org/10.1016/S0169-2046(02)00126-3

Tavares PA, Beltrão NES, Guimarães US, Teodoro AC. 2019. Integration of sentinel-1 and sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern Brazilian Amazon. Sensors (Switzerland). 19(5). https://doi.org/10.3390/s19051140

The European Space Agency. 2021. Sentinel Overview [Internet]. [accessed 2021 Apr 24]. https://sentinel.esa.int/web/sentinel/missions

Thomlinson JR, Bolstad P V, Cohen WB. 1999. Coordinating Methodologies for Scaling Landcover Classifications from Site-Specific to Global: Steps toward Validating Global Map Products. Remote Sens Environ [Internet]. 70(1):16–28. https://doi.org/https://doi.org/10.1016/S0034-4257(99)00055-3

Tian S, Zhang X, Tian J, Sun Q. 2016. Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China. Remote Sens. 8(11):1–14. https://doi.org/10.3390/rs8110954

Tolessa T, Senbeta F, Kidane M. 2017. The impact of land use/land cover change on ecosystem services in the central highlands of Ethiopia. Ecosyst Serv [Internet]. 23(June 2016):47–54. https://doi.org/10.1016/j.ecoser.2016.11.010

Toure SI, Stow DA, Shih H chien, Weeks J, Lopez-Carr D. 2018. Land cover and land use change analysis using multi-spatial resolution data and object-based image analysis. Remote Sens Environ [Internet]. 210(March):259–268. https://doi.org/10.1016/j.rse.2018.03.023

Trisurat Y, Shirakawa H, Johnston JM. 2019. Land-use/land-cover change from socio-economic drivers and their impact on biodiversity in Nan Province, Thailand. Sustain. 11(3). https://doi.org/10.3390/su11030649

Truong NCQ, Nguyen HQ, Kondoh A. 2018. Land Use and Land Cover Changes and Their Effect on the Flow Regime in the Upstream Dong Nai River Basin, Vietnam. Water [Internet]. 10(9). https://doi.org/10.3390/w10091206

Turner MG, Gardner RH. 2015. Landscape Ecology in Theory and Practice: Pattern and Process. [place unknown]: Springer-Verlag New York. https://doi.org/https://doi.org/10.1007/978-1-4939-2794-4

Ty PH, Phuc NQ, Westen G van. 2014. Vietnam in the debate on land grabbing: conversion of agricultural land for urban expansion and hydropower development. In: Kaag M, Zoomers A, editors. Glob L grab beyond hype [Internet]. 1st ed. Nova Scotia: Fernwood Publishing; p. 135–151. http://www.bloomsburycollections.com/book/the-global-land-grab-beyond-the-hype/ch8-vietnam-in-the-debate-on-land-grabbing-conversion-of-agricultural-land-for-urban-expansion-and-hydropower-development/

USGS. 2020. EarthExplorer [Internet]. [accessed 2020 Aug 1]. https://earthexplorer.usgs.gov/

USGS. 2022. What is remote sensing and what is it used for? [Internet]. [accessed 2022 Oct 4]. https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used

Vandana M, John SE, Maya K, Padmalal D. 2020. Environmental impact of quarrying of building stones and laterite blocks: a comparative study of two river basins in Southern Western Ghats, India. Environ Earth Sci [Internet]. 79(14):1–15. https://doi.org/10.1007/s12665-020-09104-1

Vaz E, De Noronha T, Nijkamp P. 2014. Exploratory Landscape Metrics for Agricultural Sustainability. Agroecol Sustain Food Syst [Internet]. 38(1):92–108. https://doi.org/10.1080/21683565.2013.825829

Vu Van N, Nguyen Hai A. 2015. Assessment of the Water Transfer Capacity from Be River Basin through Phuoc Hoa Hydraulic-Works. Clean - Soil, Air, Water. 43(5):645–651. https://doi.org/10.1002/clen.201300262

Wang L, Sousa WP, Gong P. 2004. Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. Int J Remote Sens [Internet]. 25(24):5655–5668. https://doi.org/10.1080/014311602331291215

Wei B, Xie Y, Wang X, Jiao J, He S, Bie Q, Jia X, Xue X, Duan H. 2020. Land cover mapping based on time-series MODIS-NDVI using a dynamic time warping approach: A casestudy of the agricultural pastoral ecotone of northern China. L Degrad Dev [Internet]. [accessed 2020 May 16] 31(8):1050–1068. https://doi.org/10.1002/ldr.3502

Weng YC. 2007. Spatiotemporal changes of landscape pattern in response to urbanization. Landsc Urban Plan. 81(4):341–353. https://doi.org/10.1016/j.landurbplan.2007.01.009

Wu Q, Li H qing, Wang R song, Paulussen J, He Y, Wang M, Wang B hui, Wang Z. 2006. Monitoring and predicting land use change in Beijing using remote sensing and GIS. Landsc Urban Plan. 78(4):322–333. https://doi.org/10.1016/j.landurbplan.2005.10.002

Xie Z, Chen Y, Lu D, Li G, Chen E. 2019. Classification of land cover, forest, and tree species classes with Ziyuan-3 multispectral and stereo data. Remote Sens. 11(2):1–27. https://doi.org/10.3390/rs11020164

Xu G, Dong T, Cobbinah PB, Jiao L, Sumari NS, Chai B, Liu Y. 2019. Urban expansion and form changes across African cities with a global outlook: Spatiotemporal analysis of urban land densities. J Clean Prod [Internet]. 224:802–810. https://doi.org/https://doi.org/10.1016/j.jclepro.2019.03.276

Xu G, Jiao L, Liu J, Shi Z, Zeng C, Liu Y. 2019. Understanding urban expansion combining macro patterns and micro dynamics in three Southeast Asian megacities. Sci Total Environ [Internet]. 660:375–383. https://doi.org/10.1016/j.scitotenv.2019.01.039

Xu L, Ma A. 2021. Coarse-to-fine waterlogging probability assessment based on remote sensing image and social media data. Geo-Spatial Inf Sci [Internet]. 24(2):279–301. https://doi.org/10.1080/10095020.2020.1812445

Xu X, Shrestha S, Gilani H, Gumma MK, Siddiqui BN, Jain AK. 2020. Dynamics and drivers of land use and land cover changes in Bangladesh. Reg Environ Chang [Internet]. 20(2):54. https://doi.org/10.1007/s10113-020-01650-5

Yang X, Chen LL, Li Y, Xi W, Chen LL. 2015. Rule-based land use/land cover classification

in coastal areas using seasonal remote sensing imagery: a case study from Lianyungang City, China. Environ Monit Assess [Internet]. 187(7):449. https://doi.org/10.1007/s10661-015-4667-3

Yin J, Yin Z, Zhong H, Xu S, Hu X, Wang J, Wu J. 2011. Monitoring urban expansion and land use/land cover changes of Shanghai metropolitan area during the transitional economy (1979-2009) in China. Environ Monit Assess. 177(1–4):609–621. https://doi.org/10.1007/s10661-010-1660-8

Yin L, Dai E, Xie G, Zhang B. 2021. Effects of land-use intensity and land management policies on evolution of regional land system: A case study in the hengduan mountain region. Land. 10(5). https://doi.org/10.3390/land10050528

Zakeri H, Yamazaki F, Liu W. 2017. Texture analysis and land cover classification of tehran using polarimetric synthetic aperture radar imagery. Appl Sci. 7(5). https://doi.org/10.3390/app7050452

Zhang B, Zhang Q, Feng C, Feng Q, Zhang S. 2017. Understanding land use and land cover dynamics from 1976 to 2014 in Yellow River Delta. Land. 6(1):1–20. https://doi.org/10.3390/land6010020

Zhang F, Tashpolat T, Kung H, Ding J. 2010. The change of land use/cover and characteristics of landscape pattern in arid areas oasis: An application in Jinghe, Xinjiang. Geo-spatial Inf Sci [Internet]. 13(3):174–185. https://doi.org/10.1007/s11806-010-0322-x

Zhang H, Xu R. 2018. Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta. Int J Appl Earth Obs Geoinf. 64(September 2017):87–95. https://doi.org/10.1016/j.jag.2017.08.013

Zhang H, Zhang Y, Lin H. 2012. A comparison study of impervious surfaces estimation using optical and SAR remote sensing images. Int J Appl Earth Obs Geoinf [Internet]. 18:148–156. https://doi.org/https://doi.org/10.1016/j.jag.2011.12.015

Zhang J. 2010. Multi-source remote sensing data fusion: Status and trends. Int J Image Data Fusion. 1(1):5–24. https://doi.org/10.1080/19479830903561035

Zhang Q, Wang J. 2003. A rule-based urban land use inferring method for fine-resolution multispectral imagery. Can J Remote Sens. 29(1):1–13. https://doi.org/10.5589/m02-075

Zhang Y, Li Q, Huang H, Wu W, Du X, Wang H. 2017. The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: A case study in Beijing, China. Remote Sens. 9(9). https://doi.org/10.3390/rs9090865

Zhang Y, Sun L. 2019. Spatial-temporal impacts of urban land use land cover on land surface temperature: Case studies of two Canadian urban areas. Int J Appl Earth Obs Geoinf [Internet]. 75:171–181. https://doi.org/https://doi.org/10.1016/j.jag.2018.10.005

Zheng HW, Shen GQ, Wang H, Hong J. 2015. Simulating land use change in urban renewal areas: A case study in Hong Kong. Habitat Int [Internet]. 46:23–34. https://doi.org/10.1016/j.habitatint.2014.10.008

Zheng X, Wang Y, Gan M, Zhang J, Teng L, Wang K, Shen Z, Zhang L. 2016. Discrimination of settlement and industrial area using landscape metrics in rural region. Remote Sens. 8(10):1–19. https://doi.org/10.3390/rs8100845

Zoungrana BJB, Conrad C, Amekudzi LK, Thiel M, Da ED, Forkuor G, Löw F. 2015. Multi-temporal landsat images and ancillary data for land use/cover change (LULCC) detection in the Southwest of Burkina Faso, West Africa. Remote Sens. 7(9):12076–12102. https://doi.org/10.3390/rs70912076

Summary

Owing to urbanization and industrialization as well as other human activities, land use in Binh Duong province of Vietnam has significantly changed since 1995. This study aimed to use and develop GIS and RS techniques for time-series land cover and land use monitoring and classification from 1995 to 2020 and prediction to 2030 for Binh Duong province. The hypotheses of this study were that: (1) There is a connection between the land cover and land use, and this connection can be measured and analyzed by geospatial information techniques in Binh Duong province; (2) There are diverse effects of data sources, data structure, image processing, and fusion technique on land use land cover classification efficiency, and it is possible to select an optimal mapping approach given the data availability in the study area and the objective of the study; (3) There is a significant change in land use patterns of the study area from 1995 to 2020; (4) The urban expansion process in the study area varies both spatially and temporally during the study period; (5) It is possible to predict future land use of the study area based on various natural and socioeconomic factors; and (6) Land use change and urban expansion cause significant changes in landscape patterns of the study area.

For data, various kinds of data were investigated and collected. They consisted of optical (Landsat-5, -7, -8, and Sentinel-2) and SAR (Sentinel-1) images acquired during the study period, administrative boundary data, training and validation data for the classification process, census data, SRTM DEM, population density raster data, road network, land use status map, planning maps, and other ancillary data. A field survey trip to the study area was also conducted between 18 January and 18 February 2020 to collect data and gain a deeper understanding of land cover and land use in the study area. In addition, Google Earth history images and my personal experiences were also used to interpret training and validation data.

In terms of methods, I used and developed a series of RS and GIS techniques to solve the research hypotheses and achieve the research objective. They included (1) image processing techniques for preprocessing optical and SAR data, extracting spectral indices and GLCM textures, and combining data at different levels, (2) land use land cover classification using pixel-based and object-based approaches, D-S theory, spatial analysis, decision rules, and random forest classifier, (3) accuracy assessment based on visual assessment and confusion matrix, (4) change detection based on spatial and temporal analysis and statistics such as transition matrices, urban growth rate calculation, and district-based, ring-based, and sector-based analysis, (5) simulation of future land use based on the Markov chain and decision forest algorithm, and (6) evaluation of landscape pattern change using landscape metrics. The ERDAS IMAGINE 2020, SNAP 8.0, QGIS 3, IDRISI TerrSet 2020, FRAGSTATS 4.2, and R 3.6 software, depending on the purpose, were used for these tasks.

For the results, I proved that land cover and land use in the study area were not only linked by spatial distribution and spectral properties but also by temporal characteristics. On the one hand, each land use type has its own spatial pattern and structure characterized by the properties of the land cover classes within it, such as composition, spatial distribution, spectral signature, and dominant class as well as the shape and size of objects. On the other hand, the change or non-change of land cover at a given site over different times of the year may also demonstrate the manner in which humans interact with the land, thereby showing the type of land use. This connection can easily be measured and analyzed based on RS and GIS techniques. Once the relationship between land cover and land use is clearly defined and suitable classification schemes are established, it is possible to convert a land cover map to a land use map based on their relationship. These results have confirmed the research hypothesis 1.

In addition, I supported that data sources, data structure, image processing, and fusion technique have diverse effects on land use land cover classification efficiency. First, using multi-temporal images in a pixel-based classification improved the accuracy of the generated land cover map compared to those using single-date images. Second, the segmentation technique and object-based classification could create boundaries between regions with different land use types and then relatively precisely formed land use function regions, which paved the way for producing the final land use map. Third, the fusion of SAR and optical data based on the D-S theory at the decision level yielded better land cover mapping results compared to using single-time single-sensor images or stacked optical-SAR images. Fourth, the integration of SAR and optical products using the layer-stacking technique at the pixel level did not give more power to the land cover classification process. Fifth, the inclusion of GLCM textures and spectral indices in the datasets helped improve the mapping results in this study. However, while the effectiveness of the textures is clear, the contribution of the spectral indices is still controversial. Last but not least, I developed a novel approach that is a combination of pixel-based and object-based classifications using a random forest classifier, GIS techniques, and decision rules on multi-temporal RS data. This is the optimal mapping approach given the data availability in the study area and the objective of the study. It provides the ability to effectively extract and translate a land cover map into a land use map. These results have confirmed the research hypothesis 2.

By applying the optimal mapping approach developed, land use maps of Binh Duong province in 1995, 2001, 2005, 2010, 2015, and 2020 were generated and analyzed. I analyzed and confirmed that there was a large transition from agricultural and unused lands to other uses. This resulted in an expansion of developed areas, recreational regions, mining sites, and water surfaces, a drastic decline of agricultural land for annual crops, and a fluctuation of perennial cropland and unused land. These results have confirmed the research hypothesis 3.

I also measured that the urban area has expanded 65 times within 25 years at an increasing rate. The AER and ECR were uneven between subregions, and there was a gradual expansion and shift from south to north of the province and spreading to rural districts at an increasingly rapid rate during the study period. It led to a gradual transition from a compact urban form to a dispersed urban form. The factors affecting the changes comprise natural conditions, development histories, policies and practices for urbanization, industrialization, and agricultural development, and product price fluctuations in the market. These results have confirmed the research hypothesis 4.

Furthermore, I confirmed that it is possible to identify variables driving land use change and simulate future land use in terms of quantity and location based on the Markov chain and decision forest algorithm with acceptable reliability. I discovered that the changes were driven by many variables. In which, the drivers of distances to the province centre, district centres, existing residential areas, and main road and mean population density has an impact on the conversion from agricultural land to residential land. Meanwhile, the transition from agricultural land to industrial and commercial areas is driven by the variables of distances to water sources, district centres, existing industrial areas, planned industrial zones, and transportation ports. Then, the simulation results showed that there will be 253.8 km^2 of agricultural land urbanized in the period from 2020 to 2030. The urban areas will gradually expand from the edge of the existing zones and fill the newly planned areas from South to North and Northeast of the province. These results have confirmed the research hypothesis 5.

Finally, by calculating and analyzing landscape metrics from 1995 to 2020, I measured that by the impacts of land use change and urban expansion, the studied landscape was decreasing in dominance and increasing diversity and heterogeneity at the landscape level. The processes of dispersion and aggregation were taking place at the same time in the entire landscape and in the urban class. Meanwhile, there was an intense change in the direction of increasing the fragmentation and dispersion of natural and semi-natural landscapes. These changing trends are forecast to continue in the next decade. These results have confirmed the research hypothesis 6.

In scientific terms, the evaluations and findings of this study contribute to the existing knowledge on land use land cover study using RS data and GIS techniques. In particular, my novel approach developed in this study, which helps to generate and translate a land cover map into a land use map from satellite images and GIS techniques, offers many advantages. It promotes the reproducibility and proactivity of the research as well as cost-efficiency and time savings. The output of this approach, i.e., the land cover map and land use map, can be used for different purposes. In practical terms, by analyzing and discovering land use change and urban expansion, their driving factors, and their effects on the landscape pattern in Binh Duong province of Vietnam, this study reveals a pattern of rapid urbanization in developing countries under the impact of land policies. Some practical lessons can be drawn from them. They can lay the groundwork

for further studies on urban planning, land management, and policymaking in Binh Duong province and other localities not only in Vietnam but also in other countries.

Declaration

I, Bui Dang Hung, declare that the dissertation, which I hereby submit for the doctoral degree at the Doctoral School of Geosciences, University of Szeged, is my own and original work and has not been previously submitted for an academic degree or title at this or any other institution. I certify that the help received in the preparation of this dissertation and all sources used have been fully acknowledged and cited.

Appendix A

Table A1. Summary of training and validation data for the pre-land cover map.

No.	Class	Training (Point/Polygon)	Validation (Point/Polygon)	Total (Point/Polygon)
1	Barren land	1423/56	143/25	1566/81
2	Barren land to grass/crops	266/11	35/5	301/16
3	Crops	190/13	13/5	203/18
4	Grass/crops to barren land	222/9	39/7	261/16
5	Grass	118/7	23/3	141/10
6	Impervious surface with high albedo	476/49	82/28	558/77
7	Impervious surface with low albedo	927/58	113/24	1040/82
8	Water	358/21	39/9	397/30
9	Mature woody trees to barren land	84/4	16/3	100/7
10	Mature woody trees	691/31	67/11	758/42
11	Young woody trees	154/8	16/3	170/11
	TOTAL	4909/267	586/123	5495/390

Table A2. Summary of validation data for land cover maps.

No.	Class	Final Land Cover Map (Point)	T1 Land Cover Map (Point)	T2 Land Cover Map (Point)
1	Barren land	159	178	198
2	Annual plants	87	52	36
3	Grass	23	23	35
4	Impervious surface	195	195	195
5	Water	39	39	39
6	Perennial plants	83	99	83
	TOTAL	586	586	586

Table A3. Summary of training data for land use function regions.

No.	Class	Segment
I. First round		
1	Recreation area	41
2	Mining site	81
3	Industrial area	115
4	Other	162
	TOTAL	399
II. Second round		
1	Industrial area	129
2	Other	220
	TOTAL	349

Table A4. Summary of validation data for the land use map.

No.	Class	Point
1	Unused land	159
2	Industry and commerce	82
3	Recreation and green space	23
4	Mixed residence	113
5	Mining site	25
6	Agriculture with annual plants	87
7	Agriculture with perennial plants	83
8	Water surface	39
	TOTAL	611