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

Summary of 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

1. 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 of Vietnam 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 geographic information system (GIS). However, 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. 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.

2. Research objective and hypotheses

The main objective of this study is to use and to develop GIS and remote sensing 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.

3. Data, methods, and workflow

The overall workflow of this dissertation is illustrated in Figure 1.

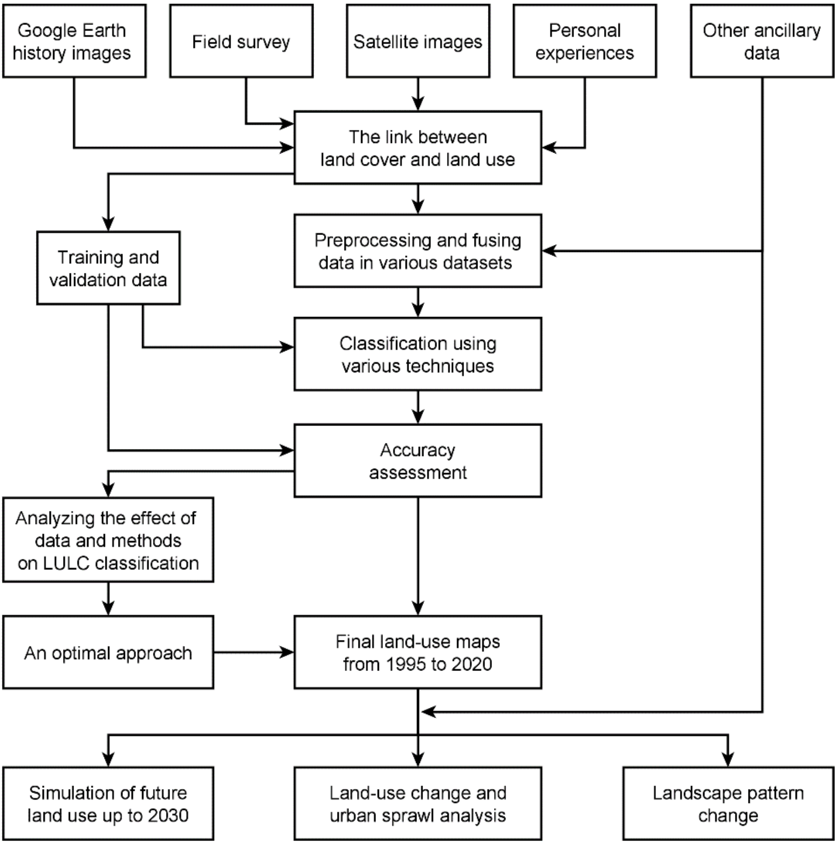

Figure 1. Overall workflow of the dissertation.

For satellite imagery, the optical and synthetic-aperture radar (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 (https://earthexplorer.usgs.gov/). Sentinel-1 Level-1 Ground Range Detected and Sentinel-2 Multispectral Instrument Level-2A images were downloaded from the Copernicus Scientific Data Hub (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 (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 (https://www.gso.gov.vn/). In addition, the Shuttle Radar Topography Mission Digital Elevation Model was downloaded from the USGS website. Population density raster data were downloaded from the WorldPop website (https://www.worldpop.org/). The road network map, land use status maps, and planning maps were collected from the provincial government. Other data were extracted from the OpenStreetMap project (https://www.openstreetmap.org/) and downloaded from the GEOFABRIK website (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.

In terms of methods, in order to solve the research hypotheses and achieve the research objective, I used and developed a series of remote sensing 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.

4. Key findings

By addressing the research hypotheses to achieve the research objective, this 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.

- **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%.

- **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.
- **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.

- **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.

- **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.

5. 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 contributes to the existing knowledge on land use land cover study using remote sensing 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 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 Binh Duong provice and other localities not only in Vietnam but also in other countries.

6. 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 Normalized Difference Vegetation Index and Normalized Difference Water Index 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 remote sensing data types such as hyperspectral and Light Detection and Ranging (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 also be pursued in further studies.

7. List of publications used in the dissertation

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.

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.

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.

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.