Land cover change analysis of buffer areas in New Capital City of Nusantara, Indonesia: A cellular automata approach on satellite imageries data

M S C Claire[1], **S R Putri**[1,*], **A W Wijayanto**[1]

[1] Department of Statistical Computing, Politeknik Statistika STIS, Jakarta 13330, Indonesia

*Corresponding author's e-mail: salwa@stis.ac.id

Abstract. The proposed plan to move Indonesia's capital city to the New Capital City of Nusantara in East Kalimantan Province undoubtedly requires careful efforts to ensure food supply for the population. Population migration to the new capital may pose a food security challenge. To address this fundamental issue, one of the most crucial approaches is to establish buffer areas that can support the food needs of the new capital. The currently existing official Area Sampling Frame survey conducted by the government to assess food vulnerability faced several limitations, including weather conditions, field terrain variations, and high cost. In this study, we propose the utilization of remote sensing satellite imagery data in buffer areas to analyze changes and predict future land cover, which can provide valuable data for assessing food availability. We investigate the integration of a Cellular Automata method with the two most popular analytical methods of classical Logistic Regression and data-driven Artificial Neural Networks, known as CA-LR and CA-ANN, to identify and map land cover changes in the new capital buffer zones. Our findings reveal that both combined methods, CA-LR and CA-ANN, yield fairly promising results, with correctness and kappa statistic values exceeding 80%. Prediction results indicate that buffer areas are predominantly covered by trees, while built-up areas are still limited. The flooded vegetation cover, including rice fields, is predicted to decrease by 2024. This should be a matter of concern for stakeholders, considering the construction of the new capital city is still ongoing and the number of migrants is expected to keep rising.

1. Introduction

Recent studies highlight the fact that numerous countries throughout the world have successfully relocated their national capitals [1]. In Indonesia, the idea of moving the national capital has existed since the reign of the first president, Ir. Soekarno, but this was only realized during the presidency of Ir. Joko Widodo through Law No.3 of 2022. The Indonesian national capital will be relocated from DKI Jakarta to Penajam Paser Utara Regency and Kutai Kartanegara Regency in East Kalimantan Province [2]. One of the main reasons for relocating to the capital city is economic equality to attain people's well-being [1]. In achieving this goal, issues of ensuring food supplies become a challenging task in the idea of relocating the national capital. Unfortunately, attaining food security remains a struggle for Indonesia to this day. While Indonesia is commonly acknowledged as a nation with a strong emphasis on agriculture, the truth is that it continues to face persistent challenges related to various aspects of the agricultural sector, including but not limited to land-use alterations, human resource management, and provision of necessary inputs [3].

One crucial aspect that should be taken into consideration to maintain food security is population migration since it can affect population growth. The migration of the population to the new nation's capital could potentially lead to an increased demand for food, which may pose a challenge to food security if appropriate measures are not taken in advance. Related to this matter, the role of buffer zones in the plan to relocate the national capital has a significant impact as they enhance the development and ensure the continued existence of the new capital in the most efficient manner possible [4]. The buffer zone is defined as the area surrounding the core urban area [5]. In the context of the new capital city of Indonesia, there are four buffer areas: Samarinda City, Balikpapan City, Kutai Kartanegara Regency, and Paser Regency. Therefore, a comprehensive analysis of the land resource potential in these areas needs to be carried out to monitor food security.

The topic of food insecurity has not only become a challenge for Indonesia but it is a concern that continues to be a struggle for the world society. A person is considered to be food insecure when they lack safe and adequate nutritional food that is essential to their overall well-being and the ability to lead a healthy life [6]. According to estimates for 2022, the global population witnessed a distressing scenario as approximately 735 million individuals experienced food insecurity [7]. This circumstance is the reverberation of the global upheaval caused by the Covid-19 pandemic [8]. Food security is also highlighted in the sustainable development program arranged by the United Nations, namely the Sustainable Development Goals (SDGs) in its second goal, zero hunger. This program is a continuation of the Millennium Development Goals (MDGs) which ended in 2015. According to the SDGs' second goal monitoring results, global hunger, as assessed by the prevalence of undernourishment (PoU), remained relatively stable from 2021 to 2022, but it is still significantly higher than it was before the COVID-19 pandemic.

In Indonesia, the results of development monitoring show that Indonesia has succeeded in achieving 49 of the 67 MDG indicators. However, several indicators still need to be continued in the SDGs and one of them is reducing minimum consumption below 1,400 kcal/capita/day [9]. Based on data from Statistics Indonesia, the proportion of the population with minimum calorie intake below 1400 kcal/capita/day in Indonesia continued to increase from 2017 to 2019. Furthermore, the PoU in Indonesia has risen by 1.87 percent from 8.34 percent in 2020 to 10.21 percent in 2022 according to Statistics Indonesia data. This data also shows that the PoU value for East Kalimantan Province exceeds Indonesia's value in 2022. This indicates that there is still a chance for improvement in Indonesia generally and particularly in East Kalimantan's response, where the country's new capital city will be established, to food insecurity issues. The level of hunger and food insecurity in Indonesia shown by these two indicators is significantly influenced by the price of rice which is greatly influenced by rice production [10]. Consequently, optimum governance of the potential of the national capital's buffer zones, particularly for rice production in terms of food security, is crucial to strengthening the role of buffer zones in supporting the national capital's activities and achieving food security [4].

One of the initiatives aimed at addressing food vulnerability in Indonesia involves creating a Food Security and Vulnerability Atlas (FSVA) that provides information on the state of food vulnerability and security. The FSVA covers three key aspects: availability, affordability, and utilization of food. The aspect of food availability is measured by the ratio of normative carbohydrate consumption to food availability using data from the Area Sampling Frame (ASF) survey of Statistics Indonesia and the Ministry of Agriculture data [11]. ASF Survey is an area-based survey that estimates harvested area by direct observation of sample segments to improve the quality of harvested area data, especially rice [12]. The ASF survey produces high-quality estimates, but it comes at a considerable expense [13]. Furthermore, the conditions required for taking images directly in the field for the ASF survey, which requires them to be taken when there is adequate sunlight and no shadows or other items allowed in the area being shot except rice plant objects, make this survey weather sensitive. The potential existence of difficult-to-reach regions for survey agents is also an objection to the ASF survey's operation. Hence, exploring alternative methods to address the limitations of the ASF survey remains an appealing possibility.

Land cover analysis is one of the remote sensing approaches that could potentially be utilized for estimating land area and assessing food availability in a particular region. This approach may be used to swiftly estimate the extent of land and changes over a vast region without costing a lot of funds, overcoming the constraints of the ASF survey. Extensive research has been conducted on the analysis and prediction of land cover change. Various methods can be utilized for this purpose. One commonly used method is the Cellular Automata-Markov Chain approach [14–16], while others have integrated Artificial Neural Network and Cellular Automata [17–20]. Additionally, machine learning methods like Decision Trees and Support Vector Machines can also be employed to evaluate land cover change [21]. However, it is important to note that there is limited research specifically focused on analyzing changes and predicting land cover for the new capital city of Indonesia. Therefore, this study aims to fill this gap by examining land cover changes and predictions in the buffer zones surrounding Indonesia's new capital city. The findings of this research will be valuable for stakeholders in supporting the planning and development of the new capital city of Indonesia.

2. Data and Study Area

2.1. Study area

Following Indonesian Minister of Agriculture Decree No. 472/Kpts/RC.040/6/2018 on the Location of the National Agricultural, In East Kalimantan Province, there are three rice commodity development areas: Kutai Kartanegara, Paser, and North Penajam Paser. These locations are expected to meet or lessen East Kalimantan's dependency on imported rice commodities. These three regions also serve as buffer zones for Indonesia's new capital city so that it is expected to be able to support the food needs of the new capital city. Therefore, these three regions were chosen as the focus of this study.

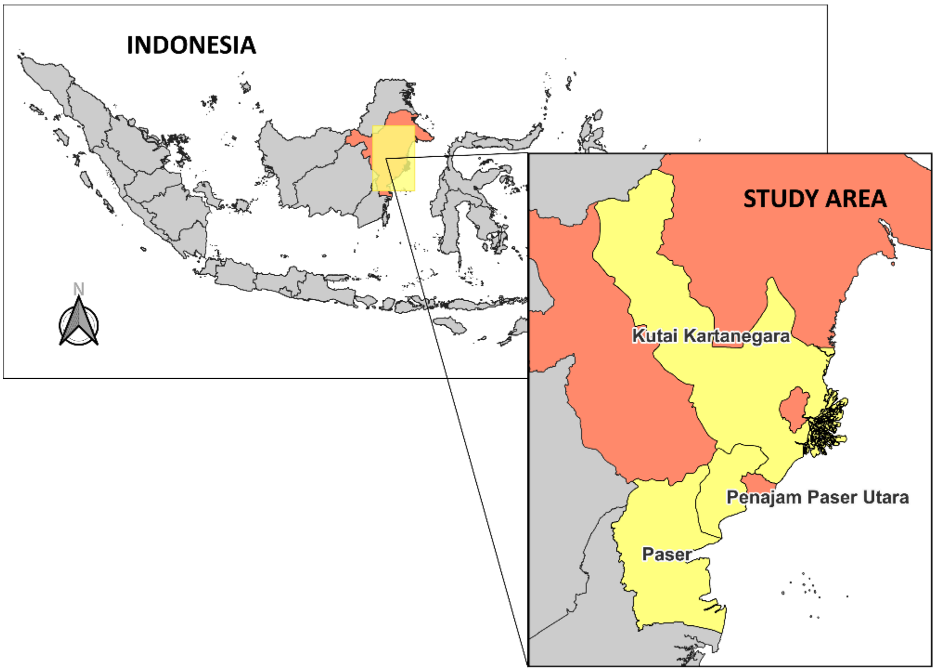

Figure 1. Research study area.

2.2. Data

Sentinel-2 satellite imagery was used to obtain the data for this study. Land use land cover (LULC) data is collected by utilizing Google Earth Engine (GEE), a cloud-based platform that offers the analysis of global environmental data. LULC data describes the earth's surface cover and its use regarding the socio-economic or functional aspects of the land [22]. This data consists of 9 classes with details that can be seen in Table 2. The spatial land cover prediction model will be developed using these land cover data. To assist in the development of a land cover prediction model, driving factors are used, namely slope and digital elevation model (DEM) obtained from the National Aeronautics and Space Administration of the Shuttle Radar Topography Mission (NASA-SRTM) Digital Elevation Model 30m. The area coverage of land cover, slope, and DEM data used is adjusted to the research study area using administrative boundary data.

Table 1. Indicator used in this research.

No	Indicator	Source	Product	Temporal Resolution	Spatial Resolution	References
1	Land Use Land Cover (LULC)	Sentinel-2	ESRI 10m Annual Land Use Land Cover (2017-2022)	10 Days	10 m (Raster Data)	[23]
2	Digital Elevation Model (DEM)	NASA SRTM	NASA SRTM Digital Elevation 30m	-	30 m (Raster Data)	[24]
3	Slope	NASA SRTM	NASA SRTM Digital Elevation 30m	-	30 m (Raster Data)	[24]
4	Administrative boundaries	Statistics Indonesia	Indonesia administrative boundaries	-	Vector Data	-

Table 2. Land cover classes.

Class	Land Cover	Definition
1	Water	Areas with a high concentration of water year-round
2	Trees	Areas with dense stands of tall plants that are at least 15 feet tall and typically have a closed or dense canopy.
4	Flooded Vegetation	Extensively watered and inundated agriculture, surrounded by clear water mixtures throughout the greater part of the year, such as flooded mangroves and rice paddies.
5	Crops	Crops that humans have planted or plotted which not surpass tree height, like corn/maize, soy, and wheat.
7	Built Area	Constructions built by humans such as housing, dense villages/cities, and paved roads.
8	Bare Ground	A rock or soil structure with little or no vegetation year-round.
9	Snow/Ice	A huge and homogeneous area of persistent ice or snow.
10	Clouds	No information was found on land cover caused by constant cloud cover.
11	Rangeland	Landscaped areas with similar grass cover and minimal to no higher plants.

3. Methods

The advancement of remote sensing gives the possibility of more advantages, economical, and instantaneous possible treatments to modern problems [25,26]. In numerous real-world scenarios, including estimating economic activity from space [26–28], forecasting rainfall simulation [29], agricultural statistics and monitoring [30,31], and oil palm cultivation detection [32–35], etc. remote sensing techniques have become increasingly popular in recent years. There is high temporal frequency and spatial resolution data accessible from remote sensing [28,32]. Furthermore, as this data might span a large geographic area, it may be useful to track geographical and temporal information and create a catalog of each specific object on Earth's surface [33,35].

Using satellite imaging data, this research was conducted to predict the land cover in the buffer zones around Indonesia's planned capital city in 2024. The procedure involves data collecting and pre-processing, model development, model validation, simulation of land cover prediction, and analysis of land cover changes from 2018 to 2024.

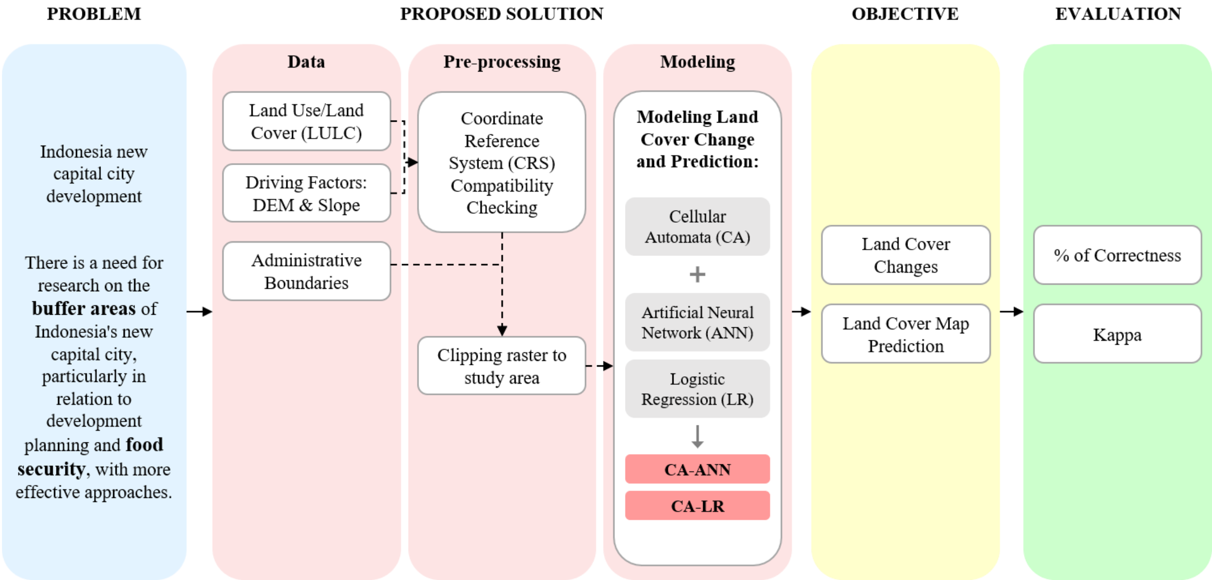

Figure 2. Research framework.

3.1. Data collection and preprocessing

The land cover data obtained from the Sentinel-2 satellite imagery through the Google Earth Engine (GEE) encompasses three years with a two-year gap, specifically 2018, 2020, and 2022. These particular years were chosen based on the availability of data and the requirements of the research to make predictions about the land cover in 2024. The research study area, which is located in the tropical region, only has two seasons, namely the rainy season and the dry season, hence there is no snow in this area. As a result, the downloaded land cover satellite imagery does not contain class 9, namely snow/ice. In addition to the land cover information, supplementary data such as slope and DEM (Digital Elevation Model) from NASA SRTM (Shuttle Radar Topography Mission) are also acquired using the GEE platform. Before further analysis, the data undergoes a pre-processing stage where the coordinate reference system (CRS) is aligned to WGS84 to ensure compatibility. The data obtained from GEE is then clipped to match the research-specific study area. This process also helps to reduce the computational workload of image processing.

3.2. Modeling land cover change

3.2.1. Defining inputs. To perform simulation and prediction of land cover changes, the 2018 (initial) and 2020 (final) land cover maps are loaded as input data along with two spatial variables, namely DEM and slope.

3.2.2. Correlation analysis of spatial variables. In this study, the spatial variables that have been chosen for analysis are the Digital Elevation Model (DEM) and slope. These particular variables were chosen based on prior research. Pearson's correlation was chosen to examine the relationship between the two variables considering the variables are not categorical. Pearson correlation coefficients vary between -1 and 1. A greater association between variables is indicated by the closer the Pearson's coefficient to ±1 regardless of the direction.

3.2.3. Changes analysis. This stage is carried out by calculating the area and its changes from land cover initial year and final year. The percentage change of each land cover class can be identified from the processing results. Moreover, a change map will be generated as one of the outputs of this process. A transition matrix is used to see changes in pixels from one class to another class.

3.2.4. Land cover change simulation. At this stage, a land cover change simulation for 2022 is carried out using two methods specifically Artificial Neural Network Multi-Layer Perceptron (ANN) and Logistic Regression (LR) methods integrated with Cellular Automata which is performed using QGIS plugin MOLUSCE. A cellular automaton (CA) is employed to simulate the changes in land use and land cover (LULC) by utilizing the transition probabilities derived from the ANN and LR learning procedure. CA models are commonly used to forecast how previous changes may affect the future via local interactions between the lands. Besides the input factor, CA is used to determine a collection of automatons that are impacted by the neighborhood [16]. The driving factor and slope from DEMs are utilized for the study because they provide repeatable data on both physical and human-related effects upon LULC [36].

A neural network is a collection of associated input and output components, each with its weight. The advantages of neural networks include their capability to handle noisy data effectively, which means they can still produce accurate results even when the input data is not entirely clean or error-free. A neural network is capable to learn and discover complex relationships and patterns within the data, even when our knowledge about the data is limited. Moreover, neural networks possess the ability to recognize and classify patterns that they have not specifically been trained on, making them highly adaptable and versatile in various scenarios. Unlike many decision tree algorithms that are primarily designed for categorical data, neural networks can handle a wide range of numeric inputs and produce continuous outputs, making them particularly useful in domains where a continuous value prediction is required. Furthermore, neural networks can be employed even in situations where we have limited knowledge or understanding of the relationship between attributes and classes. These characteristics make neural networks a powerful tool in various fields and applications [37].

The ANN model structure used to predict the potential for land use transition can be seen in Figure 2. Input is passed simultaneously through the input layer, given a weight, and then processed in the hidden layer, producing an output layer consisting of land cover reclassified classes. The parameters used in the ANN architecture include a 3 px neighborhood, 0.10 learning rate, 1000 maximum iterations, and 10 hidden layers.

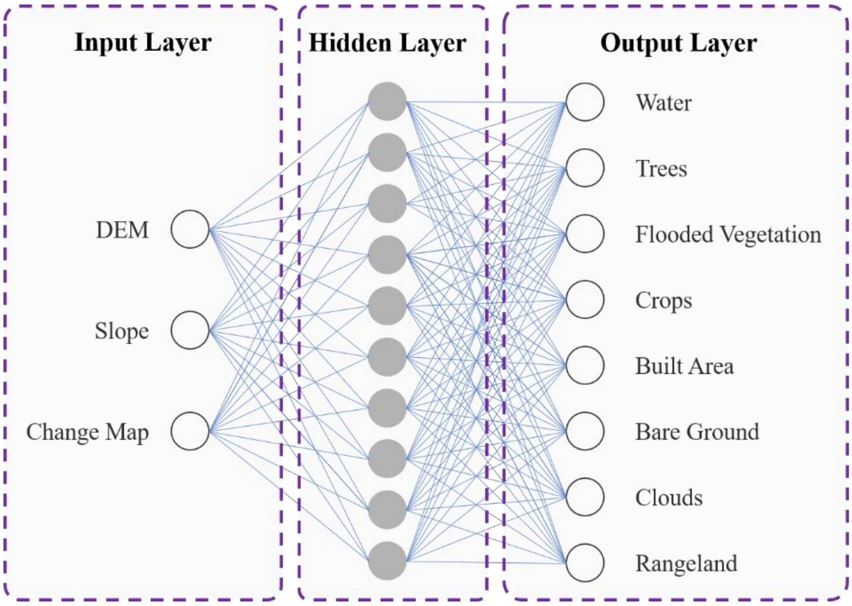

Figure 3. Structure of ANN model.

In addition to ANN, the Logistic Regression approach is also carried out to model the transitional potential. In general, logistic regression is a statistical method that is particularly effective in analysing and evaluating the connections between a categorical result variable and one or more categorical or continuous predictor variables. Logistics regression has the ability to overcome the difficulties of linear regression to describe a non-linear shape. More specifically, the logistic model forecasts the logarithm of the odds of the outcome variable based on the predictor variable(s). As previously mentioned, the logit represents the natural logarithm (ln) of the odds, and odds refer to the ratios of the probabilities of the outcome variable occurring. A simple logistic model is defined by the following formula [38].

$$logit(Y) = natural\,log(odds) = ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta X \quad (1)$$

$$\pi = \frac{e^{\alpha+\beta X}}{1+e^{\alpha+\beta X}} \quad (2)$$

Where α is the intercept, β is the regression coefficient, and e = 2.71828. The LR model was built using 100 maximum iterations and 3 px neighborhoods. Meanwhile, the samples used to build transitional potential in both methods were 1000 samples taken randomly from the input raster pixels.

3.3. Model validation
Model validation involves comparing the predicted or simulated land cover with the reference 2022 land cover to assess its accuracy. This validation process utilizes two main measures: the percentage of correctness and the kappa coefficient. A model is considered to have excellent agreement if the kappa value exceeds 80% [39].

3.4. Land cover prediction
Once the model successfully fulfills the predetermined criteria during the validation stage, we proceed to generate a prediction of the alteration in land cover for the year 2024.

4. Results

4.1. Inputs data and driving factors
The input data used to build the land cover prediction model is LULC 2018 and LULC 2020 with DEM and Slope support factors. This data will be used to predict land cover in 2024, which is preceded by predicting land cover in 2022 to validate the model that has been made.

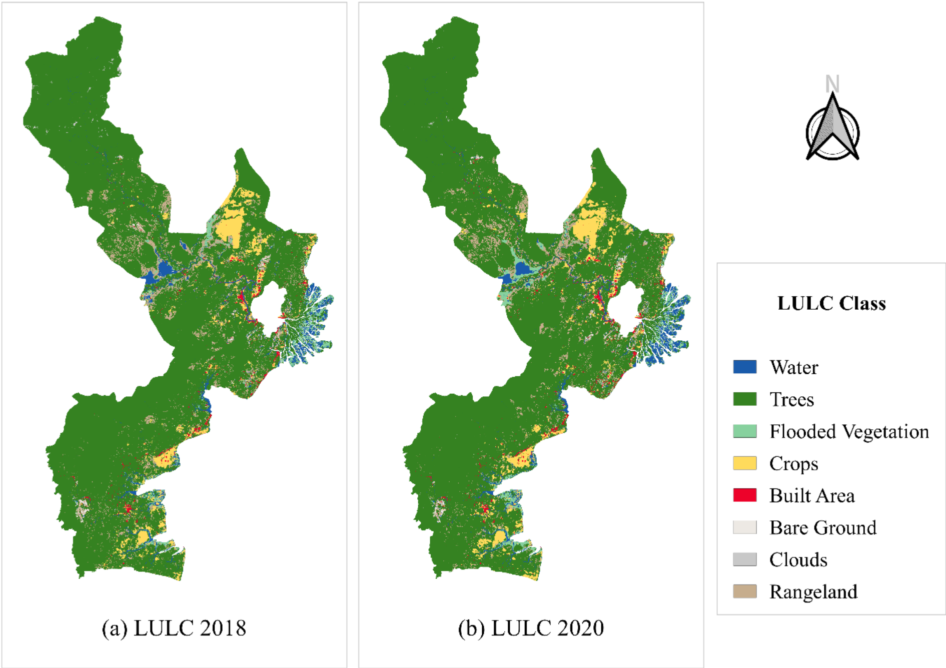

Figure 4. Land Uses and Land Covers (LULC) Input. (a) LULC 2018; (b) LULC 2020.

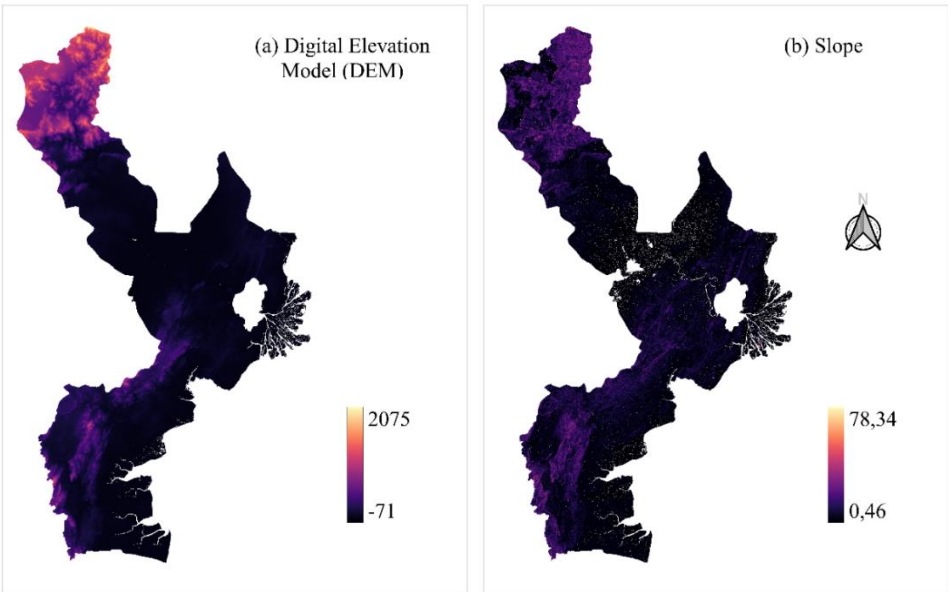

Figure 5. Driving factors: (a) DEM, (b) Slope.

According to the findings of the correlation analysis, it has been determined that the Pearson correlation coefficient between the Digital Elevation Model (DEM) and the slope is 0.6104. This indicates a moderate level of strength in the association between these two spatial variables.

4.2. Land cover change assessment

Table 3 displays the annual variations in land cover area. The buffer zone surrounding Indonesia's new capital city is primarily composed of trees, constituting over 80% of the total land cover. The presence of flooded vegetation and cultivated crops has witnessed an upward trend from 2018 to 2022, although it is anticipated to diminish in 2024 according to the forecasts generated by the Artificial Neural Network (ANN). The logistic regression (LR) result also predicts a reduction in the percentage of flooded vegetation land cover in 2024. This is not a good sign because the reduction of this area can affect food production, including rice. Thus, this can affect food security in the buffer zone.

Table 3. Annual area change.

Class	Area (%)				
	2018	2020	2022	2024	
				ANN	LR
Water	3.06	3.21	3.37	3.56	2.44
Trees	83.36	83.18	82.19	83.72	82.35
Flooded Vegetation	1.76	2.10	3.31	1.64	2.50
Crops	4.39	4.58	5.16	4.37	5.89
Built Area	0.88	1.01	0.96	0.94	0.65
Bare Ground	0.38	0.47	0.47	0.43	0.30
Clouds	0.26	0.08	0.15	0.08	0.06
Rangeland	5.91	5.36	4.38	5.25	5.81
Total	100.00	100.00	100.00	100.00	100.00

Table 4. Temporal changes.

Class	Temporal Changes (%)			
	2018-2020	2020-2022	2022-2024	
			ANN	LR
Water	0.147	0.165	0.190	-0.931
Trees	-0.181	-0.991	1.532	0.163
Flooded Vegetation	0.348	1.203	-1.662	-0.809
Crops	0.193	0.579	-0.793	0.731
Built Area	0.133	-0.050	-0.020	-0.309
Bare Ground	0.093	-0.001	-0.039	-0.168
Clouds	-0.182	0.071	-0.073	-0.099
Rangeland	-0.550	-0.976	0.867	1.423

Table 4 shows the temporal changes in land cover from 2018-2020, 2020-2022, and 2022-2024. In general, changes in rangeland and flooded vegetation are prominent compared to other classes. Notably, the most substantial transformation occurred in flooded vegetation between the years 2020 and 2022. Interestingly, land cover classes associated with food production, such as flooded vegetation and crops, are projected to exhibit a decline from 2022 to 2024. Conversely, there is an anticipated increase in the coverage of trees during the same time frame.

Table 5. Transition matrix from 2018 to 2024 Artificial Neural Networks (ANN) prediction.

Year	2024								
	LULC Class	Water	Trees	Flooded Vegetation	Crops	Built Area	Bare Ground	Clouds	Rangeland
2018	Water	0.849	0.038	0.087	0.010	0.002	0.006	0.000	0.008
	Trees	0.001	0.952	0.003	0.013	0.001	0.002	0.000	0.027
	Flooded Vegetation	0.354	0.116	0.413	0.018	0.001	0.002	0.001	0.096
	Crops	0.031	0.255	0.023	0.651	0.009	0.003	0.001	0.029
	Built Area	0.003	0.132	0.000	0.021	0.809	0.017	0.007	0.010
	Bare Ground	0.055	0.188	0.026	0.030	0.019	0.344	0.005	0.333
	Clouds	0.016	0.619	0.009	0.056	0.113	0.049	0.055	0.084
	Rangeland	0.011	0.437	0.046	0.055	0.005	0.016	0.002	0.429

Using the CA-ANN method, transition potential from 2018-2024 is shown in Table 5. Flooded vegetation is considered unstable land cover compared to trees which have a value of 0.952. Among all categories, water land cover has the most contribution to flooded vegetation. Meanwhile, the value of crops is 0.651. This shows that crops are more stable than flooded vegetation, but this number is also not satisfactory. This is quite unfortunate considering that flooded vegetation and crops are land covers that are closely related to food availability.

Table 6. Transition matrix from 2018 to 2024 Logistic Regression (LR) prediction.

Year	2024								
	LULC Class	Water	Trees	Flooded Vegetation	Crops	Built Area	Bare Ground	Clouds	Rangeland
2018	Water	0.640	0.079	0.214	0.021	0.003	0.010	0.000	0.033
	Trees	0.001	0.922	0.004	0.026	0.001	0.001	0.000	0.044
	Flooded Vegetation	0.166	0.171	0.548	0.024	0.002	0.002	0.000	0.087
	Crops	0.019	0.232	0.038	0.665	0.016	0.002	0.001	0.027
	Built Area	0.005	0.338	0.040	0.092	0.446	0.021	0.004	0.053
	Bare Ground	0.033	0.503	0.033	0.066	0.013	0.114	0.003	0.235
	Clouds	0.012	0.695	0.011	0.071	0.066	0.027	0.043	0.074
	Rangeland	0.003	0.550	0.062	0.091	0.005	0.012	0.001	0.275

The CA-LR method result shows a similar result with the CA-ANN method where the tree class is the most stable category. Based on the CA-LR method result, flooded vegetation has better stability than the results of the CA-ANN method. In addition, the class that makes the largest contribution to flooded vegetation is the same as the results shown by CA-ANN, namely water.

Below, we display graphs of annual area change by regency. From these graphs can be seen the composition of land cover and its annual changes at the regency level.

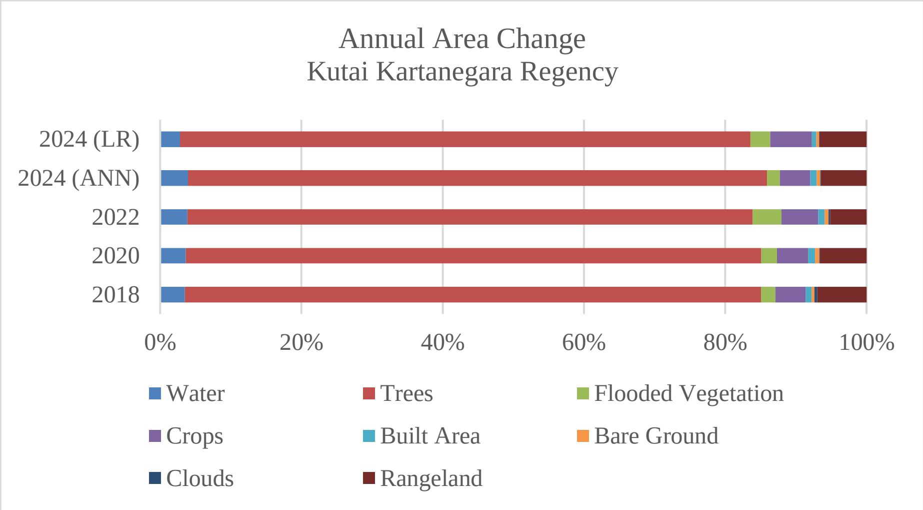

Figure 6. Annual area change of Kutai Kartanegara Regency.

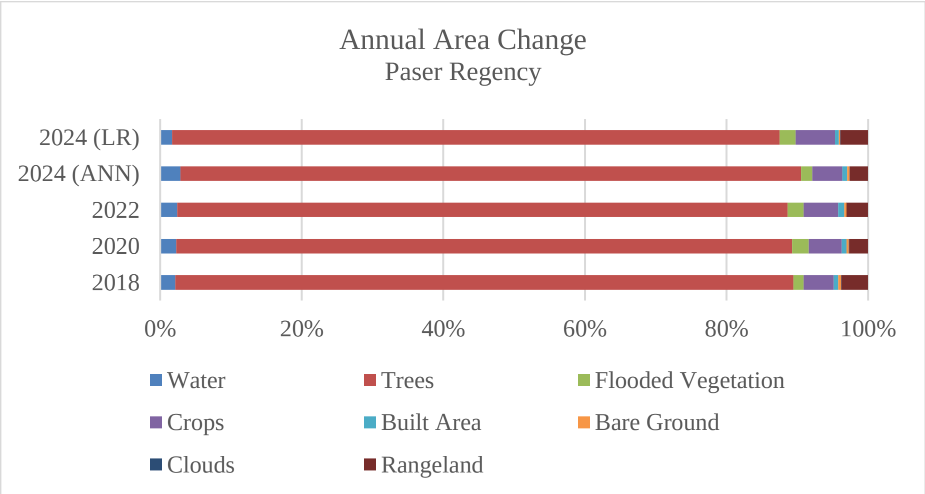

Figure 7. Annual area change of Paser Regency.

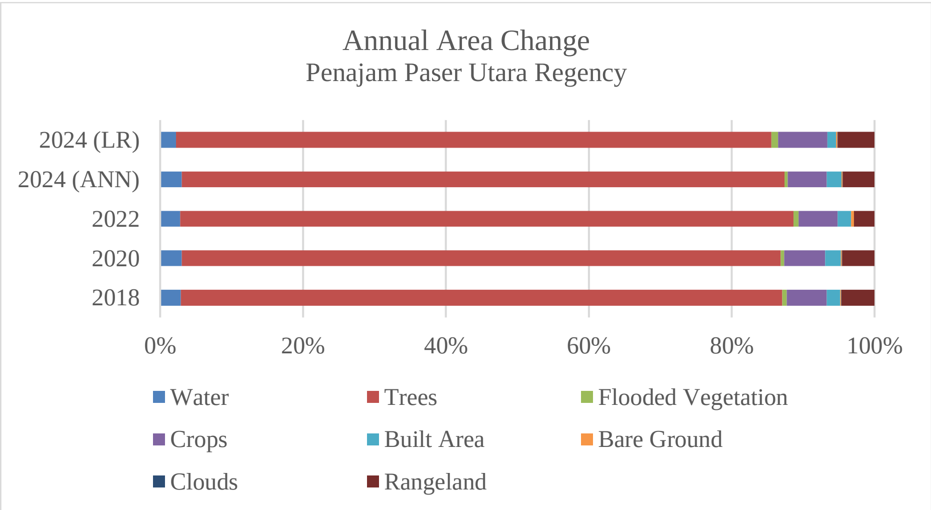

Figure 86. Annual area change of Penajam Paser Utara Regency.

The three buffer areas covered in this research have relatively similar land cover composition from 2018 to 2024 with the dominance of the Trees class. It can be seen that the built area in these regions is relatively small compared to the total land cover. Among these three regions, Kutai Kartanegara Regency has the highest proportion of flooded vegetation area. Meanwhile, the proportion of crop land cover is more prominent in Penajam Paser Utara Regency.

By examining the composition of the land cover in the buffer areas, there is potential land cover that can be utilized to maintain food security. The government can consider converting tree land into agricultural land, especially rice fields. Thus, it is expected that it can increase rice production.

4.3. Transition potential modeling and model validation

CA-ANN and CA-LR approaches were utilized to build the prediction model, which was assisted by the DEM and Slope spatial data. We projected land cover in 2022 using LULC data from 2018 and 2020. Table 7 shows a comparison of the actual and projected land cover outcomes in 2022.

Table 7. Comparison of actual and predicted Land Uses Land Covers 2022.

Class	CA-ANN Method			CA-LR Method		
	2022 Actual Area (%)	2022 Predicted Area (%)	Difference (%)	2022 Actual Area (%)	2022 Predicted Area (%)	Difference (%)
Water	3.37	3.44	0.07	3.37	2.74	-0.64
Trees	82.19	83.63	1.44	82.19	83.08	0.89
Flooded Vegetation	3.31	1.79	-1.51	3.31	2.35	-0.96
Crops	5.16	4.41	-0.75	5.16	5.45	0.29
Built Area	0.96	0.95	-0.01	0.96	0.73	-0.23
Bare Ground	0.47	0.43	-0.04	0.47	0.30	-0.16
Clouds	0.15	0.08	-0.07	0.15	0.06	-0.09
Rangeland	4.38	5.26	0.88	4.38	5.30	0.91
Total	100.00	100.00		100.00	100.00	

When comparing the actual land cover to the predicted results obtained from both the CA-ANN and CA-LR methods, notable differences can be observed. The CA-ANN method exhibits the greatest disparities in the flooded vegetation and tree categories. However, the built area is accurately predicted, indicated by the minimal variance between the actual and predicted land cover. The comparison of results from the CA-LR method also reveals discrepancies in the flooded vegetation classes. Both methods demonstrate an underestimation of five classes and an overestimation of three classes, as outlined in Table 7. Despite these variations, there is no significant disparity between the actual area and the predicted results for the year 2022.

Table 8. Validation result.

Method	% of Correctness	Kappa (overall)
CA-ANN	95.6383	0.9144
CA-LR	94.6773	0.8958

The validation results of both the Artificial Neural Network (ANN) and Logistic Regression (LR) methods indicate a slight difference in the percentage of correctness. The ANN method yields a higher correctness value of 95.6383 compared to LR which has a value of 94.6773. Additionally, the overall Kappa value of the ANN method is slightly higher than that of LR. However, it is noteworthy that both methods exhibit a validation measure value of over 80%, indicating a strong level of agreement for the model.

4.4. Land cover future prediction

The findings from the land cover predictions conducted in the buffer region of Indonesia's upcoming capital city reveal that the predominant land cover class in this area is still Trees. Furthermore, the built-up area's land cover remains relatively minimal. The area of flooded vegetation is also relatively small. This observation suggests that the development within the buffer zone has not progressed significantly. Generally, it can be seen that the land cover in this buffer area has not undergone substantial alterations between 2018 and the projected outcomes for 2024.

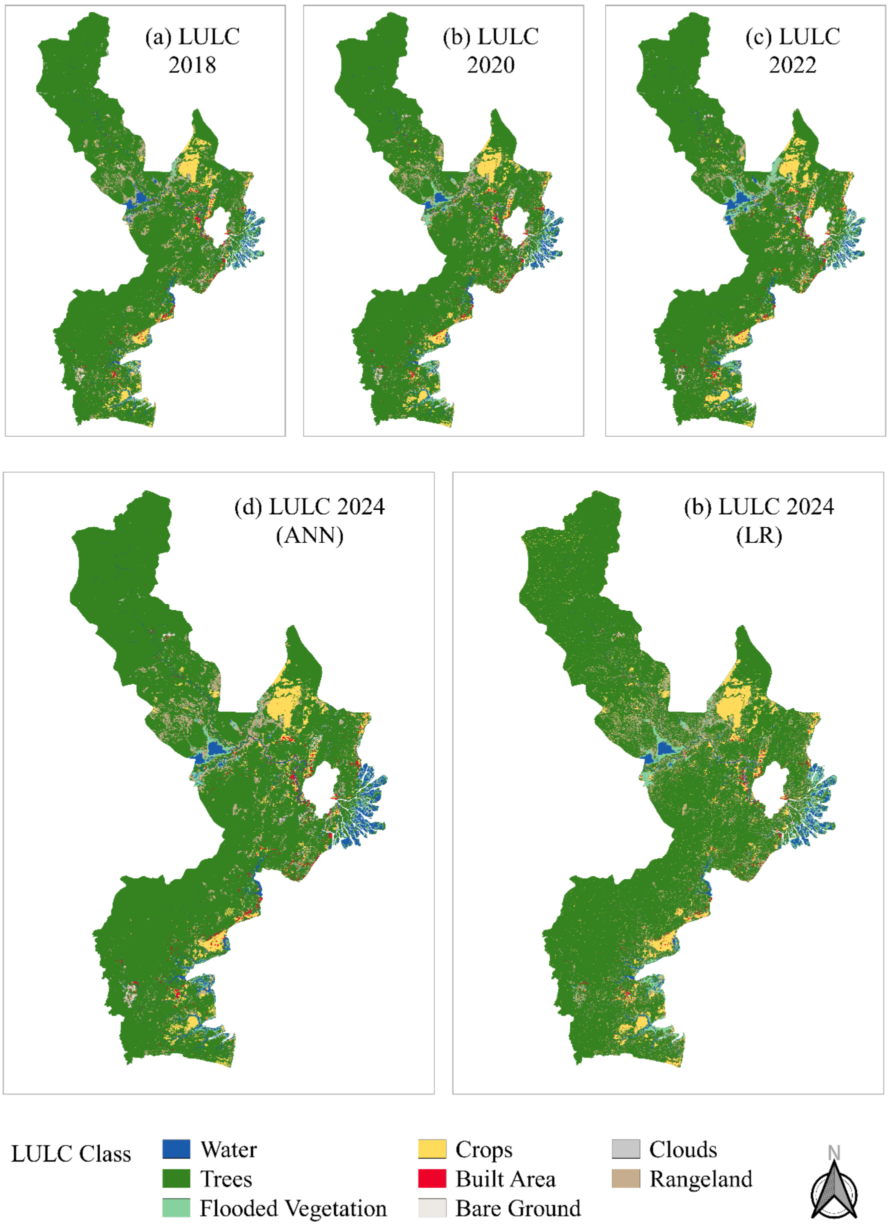

Figure 9. LULC prediction result. (a) LULC 2018; (b) LULC 2020; (c) LULC 2022; (d) prediction of LULC 2024 using CA-ANN; (e) prediction of LULC 2024 using CA-Logistic Regression.

The study area of this research covers three district-level administrative areas which are part of the buffer zones for the new capital city of Indonesia. The results of the prediction of land cover in 2024 using the CA-ANN and CA-Logistic Regression methods for these three regions can be seen in the images below.

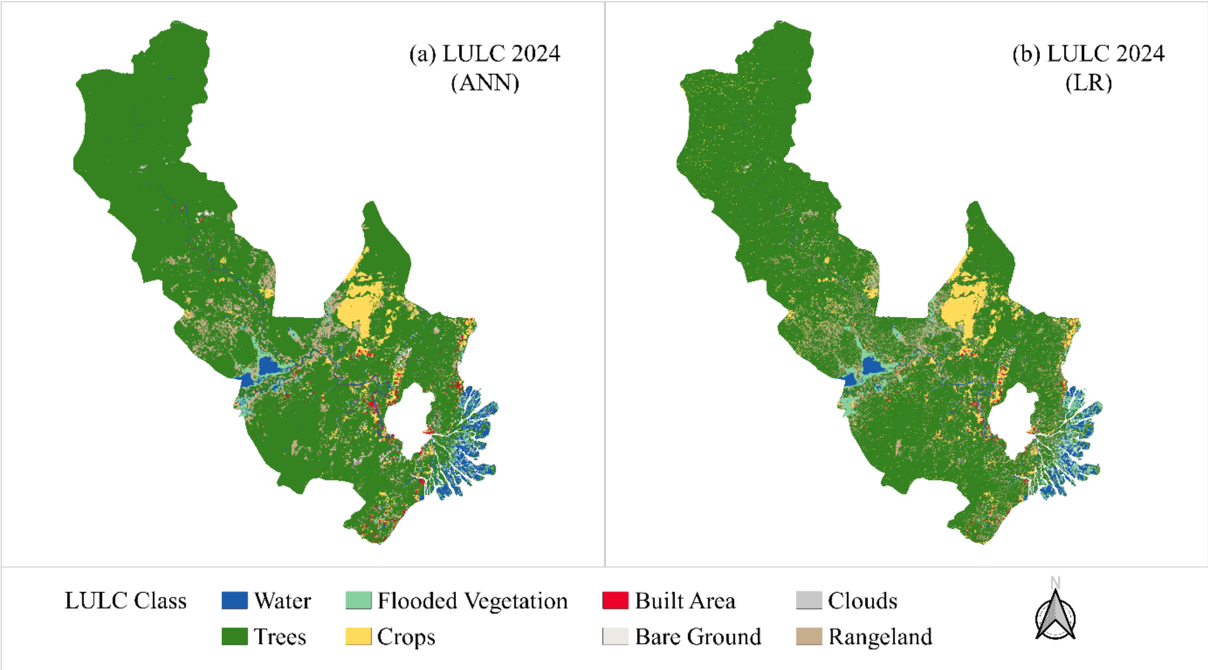

Figure 10. LULC 2024 prediction result for Kutai Kartanegara Regency.

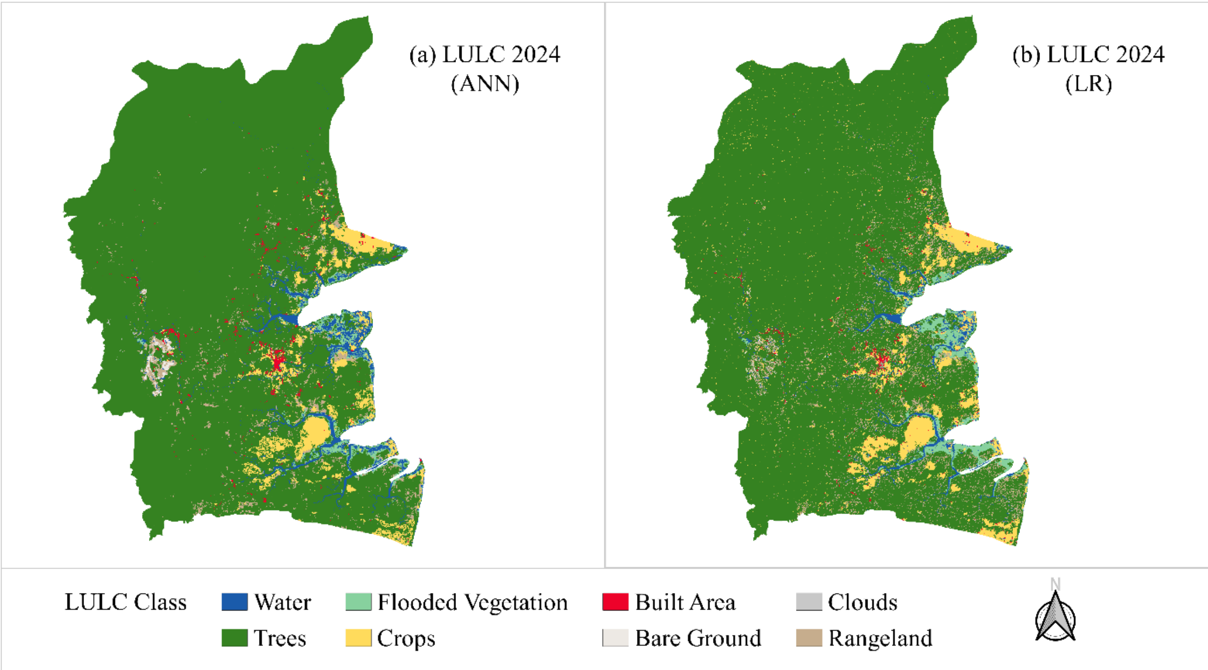

Figure 11. LULC 2024 prediction result for Paser Regency.

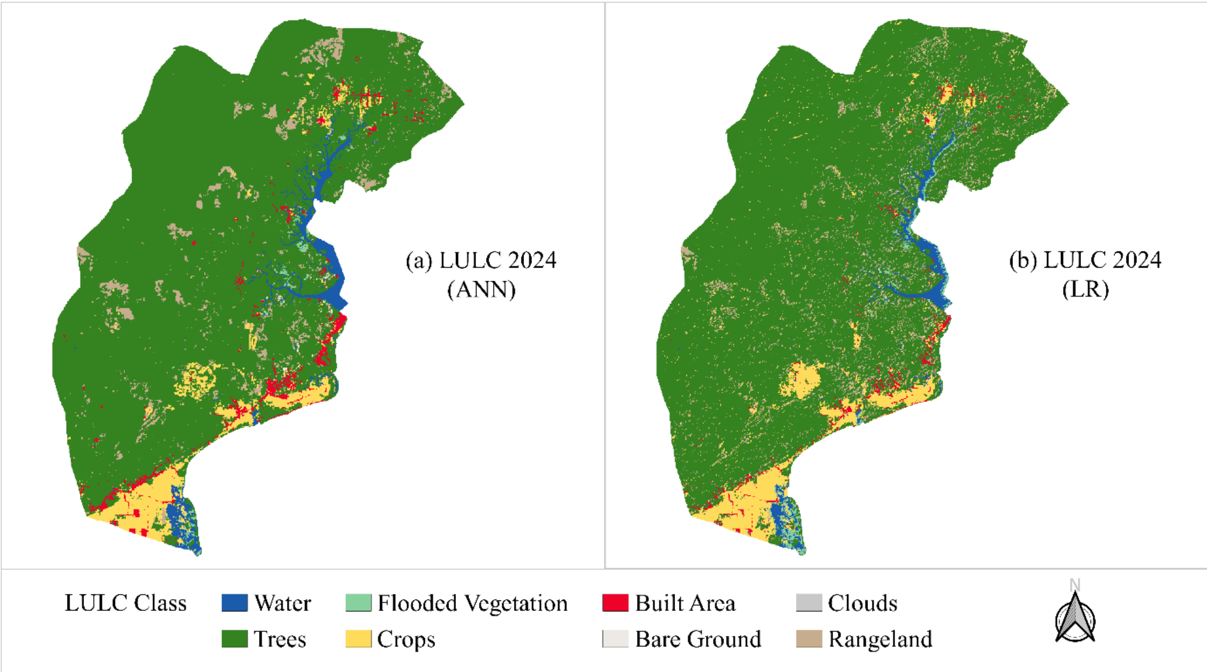

Figure 12. LULC 2024 prediction result for Penajam Paser Utara Regency.

The visual representation of the land cover prediction result by integrating Cellular Automata (CA) with ANN is quite in line with the results obtained using a combination of CA-Logistic Regression. The visual analysis reveals that, on the whole, the predictions made using ANN tend to exhibit a higher prevalence of built-up land cover compared to the LR predictions, particularly in the Penajam Paser Utara Regency. Furthermore, when examining the Paser Regency area, it becomes evident that the LR predictions indicate a stronger presence of flooded vegetation class compared to the ANN predictions.

5. Discussion

The land cover predictions carried out in this study show pretty good results. This can be seen from the validation results explained in the previous section. Therefore, the results of this study can be used for further analysis to enhance our understanding of food security. After predicting land cover, the analysis can be continued to estimate food availability in a particular region. By selecting the relevant land cover classes and implementing mathematical operations, such as multiplying with productivity data for rice or other food commodities, the obtained land cover prediction data can be further processed to generate more detailed information such as rice production. Additionally, zonal statistics can be employed to analyze production patterns in different administrative areas of interest.

Previous studies have already utilized remote sensing data to predict food security [16]. The availability and completeness of data are factors that must be taken into account when calculating food adequacy. To achieve more accurate rice production estimates, it is preferable to have land cover data that includes specific classes for rice fields rather than more general classifications like flooded vegetation or crops. Furthermore, additional supporting factors, beyond just DEM and slope, can be incorporated to enrich the analysis. The current research solely focuses on utilizing land cover data obtained from remote sensing satellite imagery. For a more comprehensive food security analysis using land cover predictions, it is recommended to combine satellite image data with administrative records and field data to obtain more accurate outcomes.

6. Conclusion

The CA-ANN and CA-LR methods produced comparable outcomes when predicting land cover. The regions of Kutai Kartanegara, Paser, and North Penajam Paser in East Kalimantan Province, which have been designated as buffer zones for Indonesia's new capital city, share a similar land cover composition. This composition remains relatively stable from 2018 to 2024 with trees being the dominant cover. These three regions will be focused on the development of rice commodity agriculture. Consequently, it is hoped that these regions will be able to provide ample food supply, particularly for the new capital region.

Among the three areas, Kutai Kartanegara Regency has the largest proportion of flooded vegetation areas, while North Penajam Paser Regency stands out for its significant amount of cropland cover. Unfortunately, the land cover predictions show that the amount of flooded vegetation which includes paddy fields, is expected to decrease by 0.8 to 1.6 percent between 2022 and 2024. Although this decrease may not be significant, it is important to consider that the construction of the new capital city will continue in the coming years, leading to an influx of people migrating to the area. Apart from that, the large amount of tree land cover is a potential that can be developed to maintain food security. The government could consider converting a part of tree land cover into rice fields as an alternative solution to increase rice production. In this way, rice production in the buffer area is expected to be able to meet food needs in the new capital city of Indonesia.

Besides comparable outcomes, the integration of Cellular Automata with ANN and Logistic Regression has proven to yield satisfactory results in predicting land cover. This can be shown by the high percentage of correctness and kappa values, which exceed 80%. Although the difference in the validation measures between the two methods is not significant, it is apparent that ANN outperforms Logistic Regression in terms of validation results. Nevertheless, both methods demonstrate a high level of accuracy in predicting land cover.

References

[1] Taufiq M. Pemindahan Ibu Kota dan Potensi Konektivitas Pemerataan Ekonomi. Jurnal Vokasi Indonesia 2020;8:27–39.
[2] Saraswati MK, Adi EAW. Pemindahan Ibu Kota Negara Ke Provinsi Kalimantan Timur Berdasarkan Analisis Swot. Jurnal Ilmu Sosial Dan Pendidikan (JISIP) 2022;6:4042–52. https://doi.org/10.36312/jisip.v6i1.3086/http.
[3] Rozaki Z. Food security challenges and opportunities in indonesia post COVID-19. Advances in Food Security and Sustainability, vol. 6, Elsevier Ltd; 2021, p. 119–68. https://doi.org/10.1016/bs.af2s.2021.07.002.
[4] Tim Penulis. Penguatan Daerah Penyangga Dalam Mendukung Ibu Kota Negara. Samarinda: Pusat Pelatihan dan Pengembangan dan Kajian Desentralisasi dan Otonomi Daerah; 2020.
[5] Qi Y, Henderson M, Xu M, Chen J, Shi P, He C, et al. Evolving core-periphery interactions in a rapidly expanding urban landscape: The case of Beijing. Landsc Ecol 2004;19:375–88.
[6] Food and Agriculture Organization of the United Nations (FAO). Hunger and food insecurity 2023. https://www.fao.org/hunger/en/ (accessed August 25, 2023).
[7] FAO, IFAD, UNICEF, WFP, WHO. In Brief to The State of Food Security and Nutrition in the World 2023. FAO; IFAD; UNICEF; WFP; WHO; 2023. https://doi.org/10.4060/cc6550en.
[8] Panghal A, Mor RS, Kamble SS, Khan SAR, Kumar D, Soni G. Global food security post COVID-19: Dearth or dwell in the developing world? Agron J 2022;114:878–84. https://doi.org/10.1002/agj2.20932.
[9] Kementerian Perencanaan Pembangunan Nasional Republik Indonesia/Badan Perencanaan Pembangunan Nasional. Sekilas SDGs 2023. https://sdgs.bappenas.go.id/sekilas-sdgs/ (accessed October 23, 2023).
[10] Mone DM V, Utami ED. Determinan Kelaparan di Indonesia Tahun 2015-2019. Seminar Nasional Official Statistics, 2021, p. 547–56.
[11] Badan Ketahanan Pangan. Peta Ketahanan dan Kerentanan Pangan 2020. 2020.
[12] BPS, BPPT. Upaya Perbaikan Data Padi dengan Metode Kerangka Sampel Area (KSA). 2018.
[13] Muchisha ND, Fitrianto A, Soleh AM. Klasifikasi Lahan Pertanian Jagung Menggunakan Machine Learning Berdasarkan Data Citra Landsat-8. 2021.

[14] Khawaldah HA, Farhan I, Alzboun NM. Simulation and prediction of land use and land cover change using GIS, remote sensing and CA-Markov model. Global Journal of Environmental Science and Management (GJESM) 2020;6:215–32. https://doi.org/10.22034/gjesm.2020.02.07.
[15] Sugandhi N, Supriatna, Kusratmoko E, Rakuasa H. Prediksi Perubahan Tutupan Lahan di Kecamatan Sirimau, Kota Ambon Menggunakan Celular Automata- Markov Chain. Jurnal Pendidikan Geografi (JPG) 2022;9. https://doi.org/10.20527/jpg.v9i2.13880.
[16] Virtriana R, Riqqi A, Anggraini TS, Fauzan KN, Ihsan KTN, Mustika FC, et al. Development of Spatial Model for Food Security Prediction Using Remote Sensing Data in West Java, Indonesia. ISPRS Int J Geoinf 2022;11. https://doi.org/10.3390/ijgi11050284.
[17] Lukas P, Melesse AM, Kenea TT. Prediction of Future Land Use/Land Cover Changes Using a Coupled CA-ANN Model in the Upper Omo–Gibe River Basin, Ethiopia. Remote Sens (Basel) 2023;15. https://doi.org/10.3390/rs15041148.
[18] Alshari EA, Gawali BW. Modeling Land Use Change in Sana'a City of Yemen with MOLUSCE. J Sens 2022. https://doi.org/10.1155/2022/7419031.
[19] Baig MF, Mustafa MRU, Baig I, Takaijudin HB, Zeshan MT. Assessment of Land Use Land Cover Changes and Future Predictions Using CA-ANN Simulation for Selangor, Malaysia. Water (Switzerland) 2022;14. https://doi.org/10.3390/w14030402.
[20] Rahmah AN, Subiyanto S, Amarrohman FJ. Pemodelan Perubahan Penggunaan Lahan dengan Artificial Neural Network (ANN) di Kota Semarang. Jurnal Geodesi Undip 2020;9.
[21] Otukei JR, Blaschke T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International Journal of Applied Earth Observation and Geoinformation 2010;12. https://doi.org/10.1016/j.jag.2009.11.002.
[22] Goverment of Canada. Land Cover & Land Use 2015. https://natural-resources.canada.ca/maps-tools-and-publications/satellite-imagery-and-air-photos/tutorial-fundamentals-remote-sensing/educational-resources-applications/land-cover-biomass-mapping/land-cover-land-use/9373#answer (accessed August 29, 2023).
[23] Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Brumby SP. Global land use / land cover with Sentinel 2 and deep learning. IEEE International Symposium on Geoscience and Remote Sensing (IGARSS), 2021.
[24] Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, et al. The shuttle radar topography mission. Reviews of Geophysics 2007;45. https://doi.org/10.1029/2005RG000183.
[25] Afira N, Wijayanto AW. Mono-temporal and multi-temporal approaches for burnt area detection using Sentinel-2 satellite imagery (a case study of Rokan Hilir Regency, Indonesia). Ecol Inform 2022;69.
[26] Putri SR, Wijayanto AW, Pramana S. Multi-source satellite imagery and point of interest data for poverty mapping in East Java, Indonesia: Machine learning and deep learning approaches. Remote Sensing Applications: Society and Environment. Remote Sens Appl 2023;29.
[27] Putri SR, Wijayanto AW, Sakti AD. Developing Relative Spatial Poverty Index Using Integrated Remote Sensing and Geospatial Big Data Approach: A Case Study of East Java, Indonesia. ISPRS Int J Geoinf 2022;11. https://doi.org/10.3390/ijgi11050275.
[28] Utami NA, Wijayanto AW, Pramana S, Astuti ET. Spatially granular poverty index (SGPI) for urban poverty mapping in Jakarta metropolitan area (JMA): a remote sensing satellite imageries and geospatial big data approach, Earth Science Informatics. Earth Sci Inform 2023.
[29] Putri SR, Wijayanto AW. Learning Bayesian Network for Rainfall Prediction Modeling in Urban Area using Remote Sensing Satellite Data (Case Study: Jakarta, Indonesia). Proceedings of The International Conference on Data Science and Official Statistics (ICDSOS), 2021, p. 77–90.
[30] Saadi TDT, Wijayanto AW. Machine Learning Applied to Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production Areas in

Nganjuk, East Java, Indonesia. International Journal of Remote Sensing and Earth Sciences 2021;18:19–32.

[31] Wijayanto AW, Triscowati DW, Marsuhandi AH. Maize Field Area Detection in East Java, Indonesia: An Integrated Multispectral Remote Sensing and Machine Learning Approach. International Conference on Information Technology and Electrical Engineering (ICITEE), IEEE; 2020, p. 168–73.

[32] Wijayanto AW, Afira N, Nurkarim W. Machine Learning Approaches using Satellite Data for Oil Palm Area Detection in Pekanbaru City, Riau. IEEE International Conference on Computational Intelligence and Cybernetics (CyberneticsCom), IEEE; 2022, p. 84–9.

[33] Nurmasari Y, Wijayanto AW. Oil Palm Plantation Detection in Indonesia Using Sentinel-2 and Landsat-8 Optical Satellite Imagery (Case Study : Rokan Hulu Regency, Riau Province). International Journal of Remote Sensing and Earth Sciences (IJReSES) 2021;18:1–18. https://doi.org/10.30536/j.ijreses.2021.v18.a3537.

[34] Putra YC, Wijayanto AW, Chulafak GA. Oil palm trees detection and counting on Microsoft Bing Maps Very High Resolution (VHR) satellite imagery and Unmanned Aerial Vehicles (UAV) data using image processing thresholding approach. Ecol Inform 2022;72.

[35] Putra YC, Wijayanto AW. Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning. Remote Sens Appl 2023;29.

[36] Muhammad R, Zhang W, Abbas Z, Guo F, Gwiazdzinski L. Spatiotemporal Change Analysis and Prediction of Future Land Use and Land Cover Changes Using QGIS MOLUSCE Plugin and Remote Sensing Big Data: A Case Study of Linyi, China. Land (Basel) 2022;11. https://doi.org/10.3390/land11030419.

[37] Han J, Kamber M, Pei J. Data Mining Concepts and Techniques. 3rd ed. Burlington: Morgan Kaufmann; 2011.

[38] Peng CYJ, Lee KL, Ingersoll GM. An introduction to Logistic Regression Analysis and Reporting. Journal of Educational Research 2002;96:3–14. https://doi.org/10.1080/00220670209598786.

[39] Hapsary MSA, Subiyanto S, Firdaus HS. Analisis Prediksi Perubahan Penggunaan Lahan dengan Pendekatan Artificial Neural Network dan Regresi Logistik di Kota Balikpapan. Jurnal Geodesi Undip 2021;10:88–97.