



Remote Sensing and GIS-Based LULC Prediction in Shah Alam: A Strategy for Sustainable Urban Growth and Development

Roslina Idris,^{1*} Mohamad Taufiq Mohamad Saleh²

¹*College of Built Environment, Universiti Teknologi MARA, 40450 Shah Alam, Selangor Darul Ehsan, Malaysia.*

²*Malaysian Space Agency (MYSA), Jalan Tun Ismail, 50480 Kuala Lumpur, Malaysia.*

Received: August 26, 2024, Accepted in revised form: September 10, 2024

Available online: October 21, 2024

ABSTRACT. Land use refers to the development of an area due to human activities, such as agriculture, industry, and residential. Land use and land cover (LULC) dynamics significantly impact agricultural productivity and food security, necessitating a comprehensive understanding and prediction of these changes. This study addresses the importance of accurate land use predictions in the context of rapid urbanisation, population growth, and climate change. This research aims to bridge the gap between land management and food security by utilising remote sensing, machine learning, and GIS technologies to predict dynamic LULC patterns. The transition potential modelling module (MOLUSCE) was used as a plugin in QGIS software to create land use and land cover (LULC) in Shah Alam City using Landsat 8 satellite images and dual sensors from 2014-2023. The ANN model was chosen to predict LULC in 2032. LULC are classified as water, developed, bare land, forests and vegetation. As a result, the developed area was the largest in 2014 and 2023, occupying 51.50% and 63.54% of the total area, respectively. Water was the smallest land use in both years. The study integrated spatial variables such as Digital Elevation Model (DEM) data and road network maps to enhance land use predictions. The findings show a fair agreement (kappa value of 0.619) between predicted and observed land-use changes, highlighting the potential for evidence-based decision-making in sustainable urban development and food security. This research contributes to the field by providing insights into future land use patterns, supporting informed policy decisions, and promoting sustainable agricultural practices for global food security efforts.

Key words: Remote Sensing, GIS Land Use land Cover, Urban Development, Machine Learning

INTRODUCTION

In the context of population increase, urbanisation, and climate change, the ability to monitor and forecast the changes of the LULC cover has become critical in policy formulation for agricultural growth and food security (Gallacher, 2016). The fast development of infrastructures, high population density, and increased urbanisation will reduce agricultural outputs as the portion of used agricultural fields (Indrawati, 2018). Besides, future land use decisions are complicated and depend on many circumstances. It is typically unpredictable what each stakeholder would perceive, intend, and have in mind. Diverse stakeholders oversee land by expressing varying expectations and selecting varying possibilities for the future (Gomes et al., 2019).

*Corresponding author: Tel.: +60 355435644

E-mail address: roslnaidris@uitm.edu.my (Roslina Idris)

Molotoks et al. (2021) and Zhuang et al. (2022) emphasize that accurate LULC change predictions are vital for maintaining food security. It has been proposed that integrating remote sensing, machine learning, and GIS technologies could significantly improve the accuracy and understanding of dynamic LULC predictions and patterns. This approach would also facilitate timely and effective land cover assessments. By using modern technology and different statistical approaches that handle the huge volume of data, this study aims to predict future land use by 2032, considering key factors such as climate change, population growth, and agricultural demand. By addressing existing gaps in land use forecasting, it seeks to provide valuable insights for stakeholders to promote effective land management and ensure food security in a sustainable manner.

It helps in understanding the spatial distributions and existing use of land, the changes in the use of such land and guidance in appropriate agricultural practices and policies. At the same time being able to achieve, the objective of this research is to take full advantage of the available land to increase the effectiveness of the measures used in agriculture and ensure sustainable food security (Alshari & Gawali, 2022). To that extent, this study has sought to examine the interaction between human activities and the environment with a view of understanding the interconnections that exist. Besides, there is the information about mapping of the changes in the land use over the past years, as well as the information about the trends for the further development of the rural environment (Maleki et al., 2020).

This paper aims at advancing the knowledge of trends in land use and what this entails to the populace during urbanisation as cities expand and grow. This is especially so where large development is expected to occur in future, consideration of resources is very crucial in planning for cities. Previous literatures have also emphasised on the prediction of the future land use (Guidigan et al., 2019; Hakim et al., 2019).

METHODOLOGY

Study Area

The actual case study is set in the commercial town of Shah Alam in the state of Selangor in Malaysia. Shah Alam is the state capital of Selangor and a city-state. Shah Alam being one of the most developed parts of Malaysia, it has gone through several transformations over the years, this is due to the escalating urbanization that is happening in the city. It is significant to recognize and forecast the area of land use in the city because such city grows and improves with time progressively. The map of the study area is depicted in Figure 1.

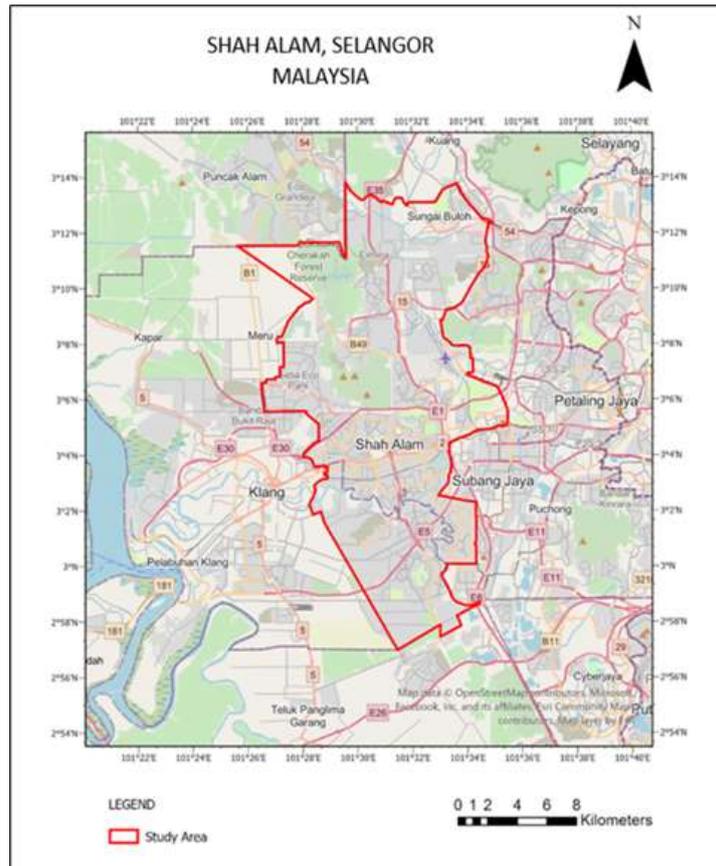


Figure 1. Map of Study Area

Data

The dataset for this study was collected from the United States Geological Survey (USGS) from the Earth Explorer website. Remote sensing data of the study area in the form of Satellite imagery which is Landsat-8 of the years 2021 and 2023 were collected. In order to increase the land use prediction accuracy for the fast-growing city of Shah Alam in the future year of 2032, the study used the Digital Elevation Model (DEM) dataset and road network as spatial variable data layers. The DEM data give a proper and significant detail on the physical features of the land such as the altitude of the region, which is vital in the put and take policy on the allocation of land for production.

Furthermore, the road network map is significant in terms of spatial variables, as it illustrates the transportation dynamics and connectivity within Shah Alam. It is an important agent of 'landscape transformation,' as it determines the accessibility and effectiveness of space for various economic activities and urbanization processes (Zhao et al., 2017; Li et al., 2019). Road networks can significantly affect land use with regions near major roads experiencing higher urbanisation rates due to increased connection and accessibility to services and markets.

With both DEM and road network map data, MOLUSCE can approximate the way land use alters in circumferences of elevation alterations and distance to roads. The simulation's findings offer new insights into how future infrastructure development, urbanisation, and environmental conservation initiatives may affect the city's landscape (Dolui & Sarkar, 2023; Manikandan & Rangarajan, 2021). The integration of spatial variables in MOLUSCE for instance the DEM and road maps fosters the understanding of human activities in relation to physical land factors with view of promoting prosperity of cities in Shah Alam.

Method Design

The method of the case study contains four (4) stages. It provides a detailed explanation of the flowchart illustrating each step of acquiring the land use prediction using the MOLUSCE plugin in QGIS Software (Figure. 2).



Figure 2. Case Study Flowchart

Data Acquisition of Landsat Images of 2014 and 2023

The study obtained satellite images for the study area in 2014 and 2023 from the United States Geological Survey (<http://usgs.com>). Landsat 8 was selected due to its two sensors: the Operational Land Imager (OLI), with nine spectral bands (Band 1 to Band 9) and the Thermal Infrared Sensor (TIRS) with two spectral bands (Band 10 and Band 11). The spatial resolution is also high with panchromatic imagery at 15 meters and multispectral at 30 meters of Landsat 8. For accurately identifying and predicting any changes in the land use/land cover of the study area, the 15-meter resolution is considered sufficient to generate a very useful Land use/ Land cover map, aligning with the focus of this case study on identifying and predicting land use/land cover changes in the study area.

Pre-processing of Landsat 8 Images of the Year 2014 and 2023

The Landsat 8 images of 2014 and 2023 were atmosphere corrected using ENVI software and then pre-processed using ArcGIS Version 10.8.3 (Esri, 2023). The study area was defined and set for Processing using ‘Clip Raster’ geoprocessing tool. As it has been stated by Esri (2023), this tool is applied for the purposes of generating a subset of a raster dataset, image service layer or mosaic dataset with the help of a rectangular area of interest. This area is defined either by minimum and maximum x and y values or by an output extent file. In this case, the limits of Shah Alam city are utilized as the output limit. In this stage the output extent is bounded at the city limit of Shah Alam.

Land Use and Land Cover (LULC) Classification of the Study Area.

There are different ways through which satellite data can be employed to classify the land use/land cover. Farid et al. (2022) described these methods, as follows; The first method is the supervised and unsupervised image classification. The second one is the machine learning algorithm. The third one is the Object-based image classification algorithms (OBIA) In attempting to classify the amount of land use/ land cover to be used in this case study, the supervised classification method was adopted. Supervised classification involves defining the signature of each class in the raster image to enable comparison of each pixel with the determined signatures then label them with the similar values digitally (Ibrahim et l., 2016).

Previous studies have shown various successful land use / land cover classification results (Jadraque Gago et al., 2020; Latif et al., 2017; Moazzam et al., 2022). For 2014 and 2023 land use/land cover classes are Water, Developed Area, Bare land, Forest and Vegetation Area. The United Nation’s Committee on World Food Security has linked food security to three (3) Sustainable Development Goals (SDGs), namely: (i) SDG 1 aims to end poverty; (ii) SDG 2 aims to ensure food security, improve nutrition, and promote sustainable agriculture, and (iii) SDG 3 focuses on ensuring healthy lifestyles and well-being across all age groups. Therefore, it may be interesting to stress the relevance of the three above-mentioned SDGs as a way to achieve the broad goal of food security whereby a country’s policies or legal provisions should be made to conform to these three Sustainable Development Goals.

Land Use/Land Cover (LULC) Prediction using Module of 2032

MOLUSCE or Modules for Land Use Change Evaluation, is intended for modelling and simulation of land use changes. Tadese et al. (2021) used this tool to assess the changes in land use and forest cover which is potential for modelling of land use change or areas vulnerable to deforestation and for the simulation of changes in LULC and forest cover (Smith et al., 2018; Johnson et al., 2018). MOLUSCE has been designed as a user-friendly plugin for the QGIS software and is able to efficiently perform LULC change calculation. It can predict depending on two different years of the land use/land cover maps, integrate the other spatial factors such as elevation, slope, distance to roads, population density etc. MOLUSCE enables the user to estimate the Transition Potential Model by different methodologies: Artificial Neuronal Network (ANN), Weights of Evidence (WoE), Logistic Regression (LR) and

Multiple Criteria Evaluation (MCE).

MOLUSCE, when employed to predict land use/land cover (LULC) maps, effectively captures the influence of neighboring factors on land use decisions by incorporating various geographic layers of the environment. It also replicates the relational patterns that exist among different locations. Within a specific region, geographic elements such as proximity to metropolitan centers, availability of transportation networks, soil quality, and types of vegetation are crucial in determining land suitability. By integrating multiple criteria through Multi-Criteria Evaluation, MOLUSCE provides a comprehensive perspective on land use suitability and accessibility, offering insights from various dimensions. These spatial factors are included in the MOLUSCE simulation model so that the results can help the planners, researchers, as well as the policymakers in the sustainable spatial planning. Besides, the Kappa statistic created during the analysis can be used to verify the degree of accuracy of the land use predictions made. The kappa statistic is frequently used to validate the accuracy of classified images, such as land cover/land use types derived from satellite imagery (Smith & Doe, 2018; Williams & Thompson, 2020). These solutions appear to be valuable, as numerous studies have demonstrated that the application of Artificial Neural Network (ANN) models, along with this module, yields better outcomes (Ahmad et al., 2023).

According to Gharaibeh et al. (2020) suggest that Artificial Neural Networks (ANN) have the capabilities to capture complex patterns in the large and growing landscapes of the cities as well as in the changing uses of the lands. The understanding of interactions among various factors is enhanced by these models, allowing for the identification of trends in land use categories. Data illustrating the behavioral dynamics of land-use phenomena are provided, enabling a more nuanced analysis of how different influences affect land utilization. Through implied driving forces, the ANN model can also identify possible relations of interdependence. Additionally, adopting an Artificial Neural Network (ANN) model is important because it minimizes confusion regarding which aspects have the most influence on land change. It effectively illustrates the impacts of each driving factor used in the analysis (Kiprotich et al., 2021). For the current research, ANN model was selected to forecast the LULC of Shah Alam City by the year 2032. As in the case of ANNs model, iterations are also important in predicting the LULC. It is commonly known as passes through the training data and corresponds to the number of training cycles that the algorithm goes through in the training process. The number of iterations in this study was set at 5000 in a way that the ANN can analyse the data adequately.

In the process of predicting the LULC map, MOLUSCE can study the influence of surrounding factors for the decisions made on the land use since the method can incorporate geographic data layers that embody spatial characteristics of the area. It can also develop topological relations between different places and centers, which is also more effective than graphic outlines of locations. Gharaibeh et al. (2020) also went on to explain that spatial covariates that essentially influences LULC change prediction are slope, distance to roads as well as urban centers, fertility of the soil and elevation. When forecasting the land use of Shah Alam in the year 2032 the two spatial variables chosen were to enhance the prediction simulation were elevation and the road network.

RESULTS AND DISCUSSION

This section consists of three subsections, each presenting the results and analysis of the case study. To this, it has presented the LULC maps of the year 2014 and 2023 and an expected LULC map of the year 2032.

Land Use and Land Cover (LULC) Map of Shah Alam in 2014 and 2023.

In this process, fifteen training samples of each class of water, the developed area, bare land, forest and vegetated area were generated in order to identify the land use of the study area. Pixel-based classification of images is based on the fact that the classification is done on an individual pixel of the image. The values for each pixel are then represented in spectral space up on which the pixel is assigned to a given class (Esri, 2023). Figure 3 represent the LULC of the year 2014 and 2023, respectively.

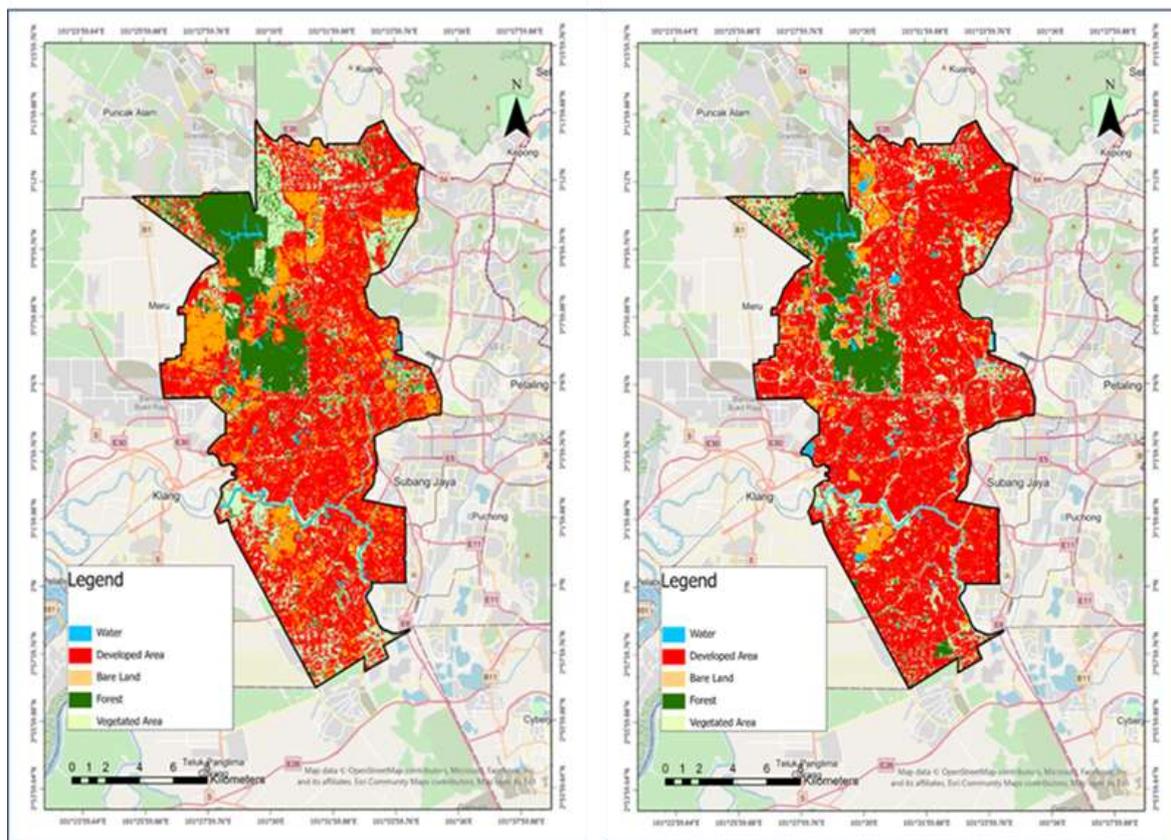


Figure 3. LULC Map of Shah Alam in 2014 (left) and LULC Map of Shah Alam in 2023 (right).

Based on calculations (Table 1), the area of Shah Alam is 302.68 square kilometers. Table 1 shows the land use by five classes, with the developed area being the largest in 2014 and 2023, occupying 51.50% (155.88 km²) and increasing to 63.54% (192.31 km²) of the total area, respectively. Vegetated Area, covering vegetated areas excluding forests such as grasslands and shrubbery, was the second-largest land use in both years, with 17.60% (53.27 km²) and 17.11% (51.79 km²), respectively. In 2014, bare land covered 15.74% (47.65 km²) and forests accounted for 12.72%

(38.50 km²), while in 2023, forests covered 10.87% (32.90 km²) and bare land covered 6.03% (18.25 km²). Water was the smallest land use in both years, occupying 2.44% (7.39 km²) in 2014 and 2.46% (7.43 km²) in 2023.

Table 1. The Area of LULC in 2014 and 2023

Land Use Class	Area sq. km ²		Percentage of Area	
	2014	2023	2014	2023
Water	7.39	7.43	2.44	2.46
Developed Area	155.88	192.31	51.50	63.54
Bare Land	47.65	18.25	15.74	6.03
Forest	38.50	32.90	12.72	10.87
Vegetated Area	53.27	51.79	17.60	17.11

Prediction of Land Use Area using MOLUSCE Plugin.

Based in this study, the LULC for the study area in the year 2032 was modelled by the application of Artificial Neural Networks (ANN). The prediction of the Land Use/Land Cover Map of Shah Alam of year 2032 is shown in Figure 4.

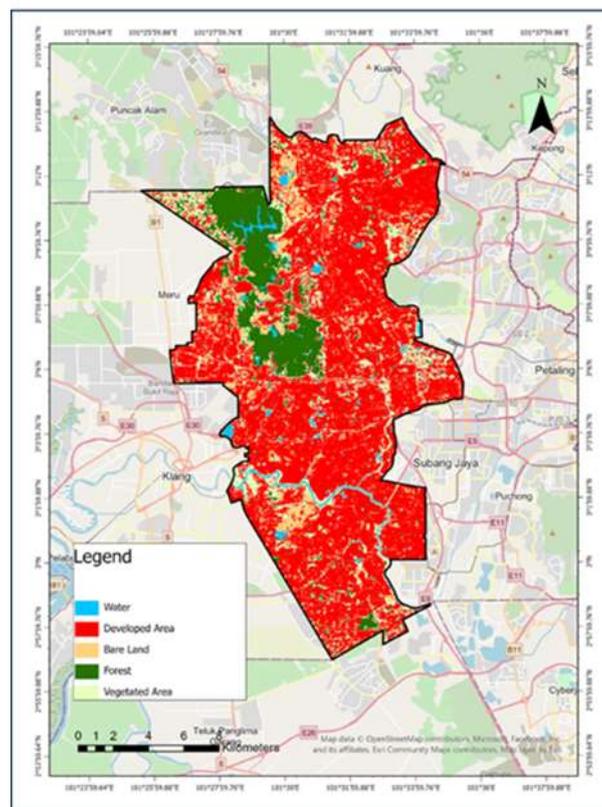


Figure 4. LULC Predicted LULC Cover Map of Shah Alam in 2032

To enhance the prediction of the 2032 Land Use/Land Cover (LULC) map, various spatial variables, including the Digital Elevation Model (DEM) and the road map of the study area, have been incorporated into the MOLUSCE model. The Artificial Neural Network (ANN) was trained using a dataset with a maximum of 5,000 training cycles, which dictates the number of epochs the training algorithm processes the dataset. The ANN configuration included 10 hidden layers and a momentum of 0.05. The inter-observer reliability for the LULC map prediction for 2032 was measured at 0.619, indicating moderate agreement. Furthermore, the minimum validation error for the LULC samples in the validation set for 2023 was 0.05, quantifying the extent of information error achieved.

The MOLUSCE model effectively analyzes the influence of surrounding factors on land use decisions by incorporating geographic data layers that represent the spatial characteristics of the area. This capability allows it to establish topological relationships between different locations and centers, making it more effective than traditional graphic outlines. Key spatial covariates identified by Gharaibeh et al. (2020), such as slope, distance to roads and urban centers, soil fertility, and elevation, are crucial for predicting LULC changes. For the forecast of land use in Shah Alam in 2023, elevation and the road network were chosen as the two primary spatial variables to enhance prediction accuracy.

To improve the reliability of these predictions, it is essential to integrate additional spatial and non-spatial factors, such as socioeconomic data, land ownership patterns, and legal constraints. By refining the model with these comprehensive datasets, the accuracy of the LULC predictions can be significantly enhanced, providing better insights for urban planning and environmental management. While the model achieved a kappa value of 0.619, suggesting it performs better than random guessing. Exploring different settings for iteration numbers and spatial data resolution will also contribute to achieving more precise forecasts.

CONCLUSION

In conclusion, therefore, the study has a kappa value of 0.619 derived from the MOLUSCE model with spatial variables of the DEM, road map at 5000 iterations reveals a rather satisfactory level of goodness of fit between predicted and observed land use changes. This means that despite the model performance is certainly not random, there is great potential for improvement. Concerning the influence on the model's quality, further characteristics, including socio-economic factors, ownership, environmental constraints, and legal features, should be introduced as well. In addition, other factors such as the number of iterations and the level of granularity of existing spatial data are also other factors for model improvement. Such results indicate that the employment of spatial variables in the prediction of changes in land uses has a good start, but it also denotes the fact that there is more improvement that is necessary in the model and future study. By doing so it is possible to obtain more accurate forecasts of land use changes, which is critically important for land use planning, management of the environment, as well as for path planning in various spheres.

ACKNOWLEDGMENTS

The authors sincerely thank the Government of Malaysia, Shah Alam City Council (MBSA) for helping in this research work through data supply.

AUTHOR CONTRIBUTIONS

Roslina Idris: Main conceptualization, methodology, data analysis, writing-original draft, writing-review and editing.

Mohamad Taufiq Mohamad Saleh: Methodology, data collection and data analysis

FUNDINGS

This research did not receive any financial funding.

DATA AVAILABILITY

Not applicable

COMPETING INTEREST

The authors declare that there are no competing interests.

REFERENCES

- Ahmad, M.N., Shao, Z., & Javed, A. (2023). Modelling land use/land cover (LULC) change dynamics, future prospects, and its environmental impacts based on geospatial data models and remote sensing data. *Environmental Science and Pollution Research*, 30(12), 32985-33001.
- Alshari, E.A., & Gawali, B.W. (2022). Modeling land use change in sana'a city of Yemen with MOLUSCE. *Journal of Sensors*, 2022(1), 7419031.
- Dolui, S., & Sarkar, S. (2023). Modelling landuse dynamics of ecologically sensitive peri-urban space by incorporating an ANN cellular automata-Markov model for Siliguri urban agglomeration, India. *Modeling Earth Systems and Environment*, 10, 167–199.
- El-Tantawi, A.M., Bao, A., Chang, C., & Liu, Y. (2019). Monitoring and predicting land use/cover changes in the Aksu-Tarim River Basin, Xinjiang-China (1990–2030). *Environmental Monitoring and Assessment*, 191, 1-18.
- ESRI. (2023). Clip Raster (Data Management)—ArcGIS Pro | Documentation. <https://pro.arcgis.com/en/pro-app/3.0/tool-reference/data-management/clip.htm>. Accessed on 13 March 2023
- ESRI. (2023). Image Classification Wizard—ArcGIS Pro | Documentation. <https://pro.arcgis.com/en/pro-app/3.0/help/analysis/image-analyst/the-image-classification-wizard.htm>. Accessed on 13 March 2023
- Farid, N., Moazzam, M. F. U., Ahmad, S. R., Coluzzi, R., & Lanfredi, M. (2022). Monitoring the impact of rapid urbanization on land surface temperature and assessment of surface urban heat island using landsat in megacity (Lahore) of Pakistan. *Frontiers in Remote Sensing*, 3, 897397.
- Gallacher, D. (2016). Drone applications for environmental management in urban spaces: A review. *International Journal of Sustainable Land Use and Urban Planning*, 3(4), 1-14.
- Gharaibeh, A., Shaamala, A., Obeidat, R., & Al-Kofahi, S. (2020). Improving land-use change modeling by integrating ANN with Cellular Automata-Markov Chain model. *Heliyon*, 6(9), e05092.

- Gomes, E., Abrantes, P., Banos, A., Rocha, J., & Buxton, M. (2019). Farming under urban pressure: Farmers' land use and land cover change intentions. *Applied Geography*, *102*, 58-70.
- Guidigan, M. L. G., Sanou, C. L., Ragatoa, D. S., Fafa, C. O., & Mishra, V. N. (2019). Assessing land use/land cover dynamic and its impact in Benin Republic using land change model and CCI-LC products. *Earth Systems and Environment*, *3*(1), 127-137.
- Hakim, A. M. Y., Baja, S., Rampisela, D. A., & Arif, S. (2019). Spatial dynamic prediction of landuse/landcover change (case study: tamalanrea sub-district, makassar city). *IOP Conference Series: Earth and Environmental Science*, *280*, 012023.
- Ibrahim, I., Abu Samah, A., Fauzi, R., & Noor, N. M. (2016). The land surface temperature impact to land cover types. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *41*, 871-876.
- Indrawati, E. (2018). Urban farming model in South Jakarta. *IOP Conference Series: Earth and Environmental Science*, *106*, 012052.
- Jadraque Gago, E., Etxebarria Berrizbeitia, S., Pacheco Torres, R., & Muneer, T. (2020). Effect of land use/cover changes on urban cool island phenomenon in Seville, Spain. *Energies*, *13*(12), 3040.
- Johnson, L.M., & Lee, C.Y. (2019). Simulating future land cover changes with MOLUSCE. In *Proceedings of the International Conference on Environmental Modelling and Software* (pp. 134-142). Elsevier.
- Kiprotich, P., Wei, X., Zhang, Z., Ngigi, T., Qiu, F., & Wang, L. (2021). Assessing the impact of land use and climate change on surface runoff response using gridded observations and swat+. *Hydrology*, *8*(1), 48.
- Latif, Z.A., & Kamsan, M.E.S. (2017). Assessing the relationship of land use land cover on surface temperature in city of Shah Alam, Malaysia using landsat-8 oli. *Journal of Fundamental and Applied Sciences*, *9*(5S), 514-525.
- Li, Y., & Jiang, N. (2019). Urban road network expansion and its driving variables: A case study of Nanjing City. *International Journal of Environmental Research and Public Health*, *16*(13), 2318.
- Maleki, M., Van Genderen, J.L., Tavakkoli-Sabour, S.M., Saleh, S.S., & Babae, E. (2020). Land use/cover change in Dinevar rural area of West Iran during 2000–2018 and its prediction for 2024 and 2030. *Geographia Technica*, *15*, 93-105.
- Manikandan, K., & Rangarajan, S. (2021). Predicting the future land use and land cover changes for Bhavani Basin, Tamil Nadu, India using QGIS MOLUSCE plugin. *Environmental Science and Pollution Research*, *29*, 86337–86348.
- Moazzam, M.F.U., Doh, Y.H., & Lee, B.G. (2022). Impact of urbanization on land surface temperature and surface urban heat Island using optical remote sensing data: A case study of Jeju Island, Republic of Korea. *Building and Environment*, *222*, 109368.
- Molotoks, A., Smith, P., & Dawson, T.P. (2021). Impacts of land use, population, and climate change on global food security. *Food and Energy Security*, *10*(1), e261.
- Roy, S.K., Alam, M.T., Mojumder, P., Mondal, I., Kafy, A.A., Dutta, M., ... & Mahtab, S.B. (2024). Dynamic assessment and prediction of land use alterations influence on ecosystem service value: A pathway to environmental sustainability. *Environmental and Sustainability Indicators*, *21*, 100319.
- Smith, J. A., & Doe, R.B. (2018). Application of the Kappa statistic in land use prediction and classification accuracy assessment. *Journal of Remote Sensing*, *45*(2), 123-134.
- Smith, J. A., & Doe, R.B. (2018). Utilizing MOLUSCE for predicting deforestation patterns in Southeast Asia. *Journal of Land Use Science*, *15*(3), 245-261.

Tadese, S., Soromessa, T., & Bekele, T. (2021). Analysis of the current and future prediction of land use/land cover change using remote sensing and the CA-Markov model in Majang forest biosphere reserves of Gambella, Southwestern Ethiopia. *The Scientific World Journal*, 2021(1), 6685045.

Williams, P. L., & Thompson, N. H. (2020). Evaluating land use change predictions with Kappa statistic: A case study in urban development. In *Proceedings of the International Conference on Geospatial Data Analysis* (pp. 56-67).

Zhao, G., Zheng, X., Yuan, Z., & Zhang, L. (2017). Spatial and temporal characteristics of road networks and urban expansion. *Land*, 6(2), 30

Zhuang, D., Abbas, J., Al-Sulaiti, K., Fahlevi, M., Aljuaid, M., & Saniuk, S. (2022). Land-use and food security in energy transition: Role of food supply. *Frontiers in Sustainable Food Systems*, 6, 1053031.