UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Establishing a baseline for land use and land cover change in a strategic water management area within the Garden Route Biosphere Reserve

AN HONOURS PROJECT PRESENTED TO THE DEPARTMENT OF GEOGRAPHY, GEOINFORMATICS, AND METEOROLOGY IN CANDIDACY FOR THE DEGREE BSC HONOURS GEOGRAPHY AND ENVIRONMENTAL SCIENCE

By

GABRIELLE GROBLER

UNIVERISTY OF PRETORIA

NOVEMBER 2023

Supervisor: Dr. Kaera Coetzer (UP),

Co-supervisors: Dr. Philemon Tsele (UP), Dr. Bianca Currie (GRBR)

Declaration

I declare that this research project is my own work. It is being submitted for the BSc Honours in Geography and Environmental Science at the University of Pretoria. It has not been submitted before for any degree or examination at any other University.

ABSTRACT

Accurate and up-to-date Land Use and Land Cover (LULC) data are crucial for managing Land Use and Land Cover Changes (LULCC) sustainably in any region. The data play a key role in planning, monitoring, and forecasting changes in LULC, thereby guiding the development of sustainable strategies and policies in terms of managing areas. This study aims to provide a comprehensive analysis of LULCC within the Kaaimans catchment area, a key Strategic Water Source Area within the Garden Route Biosphere Reserve (GRBR), Western Cape, South Africa. This study produces accurate and up-to-date LULC data for this catchment area, by performing an in-depth and quantitative evaluation of historical (1990, 2014, 2018, and 2020), current (2019, 2021, and 2023), and projected future (2030) LULC changes. The study began by analysing the historical LULC of study area by using South African National Land Cover (SANLC) data for the years 1990, 2014, 2018, and 2020. This analysis revealed significant changes in the LULC from 1990 to 2020. Understanding these changes was vital in identifying patterns and the dominating LULC classes within the catchment, which informed the next step of creating current LULC classification maps for the years 2019, 2021, and 2023. LULC classification maps were created using Random Forest (RF) classification models, remotely gathered training data on the priority LULC classes identified with the SANLC maps, and Sentinel-2 satellite imagery for the years 2019, 2021, and 2023. The overall accuracy (OA) and Kappa values were above 85% for each map. The 2023, 2021, and 2019 LULC maps had an OA of 91.31%, 86.83%, and 90.31% with Kappa values of 0.911, 0.870, and 0.907, respectively. Additionally, the Quality Allocation Disagreement Index (QADI) values were used to further assess the accuracies of each map, which showed high to moderate classification accuracy with values of 0.10, 0.14, and 0.12, respectively. These maps demonstrated an increasing trend in plantation forest, built-up areas, and shrubland from 2019 to 2023 which could be a testament to the rapid population increase according to Census data. The study predicted the LULC of the catchment area up to 2030 using two hybrid Markov Chain (MC) models, seven driver variables and the classified LULC maps. The driver variables included static (elevation, slope, aspect, distance from rivers, and distance from main roads) and dynamic (distance from built-up and distance from plantations) variables. The predictive models were the Decision Forest (DF) – MC and Multilayer perceptron (MLP) – MC models within the Land Change Modeller (LCM) tool in the TerrSet 2020 software. The kappa indices (Kno, Klocation, and Kstandard) used to validate the predictive models were all above 80%, with the DF-MC model showing a slightly higher accuracy than the MLP-MC model. The

forecasted expansion of plantation forests into natural forest areas and the growth of built-up areas at the edges of existing urban boundaries, replacing agricultural land, suggest potential environmental and socio-economic implications. These changes could be attributed to the increasing prominence of the forestry industry and urban development in the region.

Acknowledgements

We thank the Garden Route Biosphere Reserve (GRBR) and the Strategic Water Source Areas (SWSA) Network and Working Group operating in the biosphere reserve for their support and interest in this project.

We wish to acknowledge support from the National Research Foundation's (NRF) Global Change Social Sciences Research Programme (GCSSRP V2), Grant UID: 129482 via the broader Project titled "Governance models implemented in biosphere reserves: Progression towards social-ecological systems for sustainability", to which this project is linked (see https://researchbiosphere.org/ [Project lead: Kaera Coetzer]).

Table of Contents

List of Tables

List of Figures

Overview and research problem

I Introduction

Human modification of the Earth's land surface is occurring at an unprecedented rate, scale, and spatial extent (IPBES, 2018). Land-use change, which occurs when the natural landscape is transformed as a result of human activity, is what propels this human-dominated change to the Earth's land surface. It is estimated by Winkler, et al (2021) that the global land-use change since 1960 has altered almost a third of the Earth's land surface. However, due to land use change being a multifaceted environmental and socioeconomic issue, managing and monitoring land use and land cover change (LULCC) needs to be prioritized at global and local scales (Turner, et al., 2007).

In developing countries like South Africa, it is a challenge to sustainably manage land use change due to the rapid expansion of agricultural land and urbanization (Abu Hatab, et al., 2019). As South Africa's population and reliance on the land grows, more changes are brought to the landscape which has implications for the natural land cover and maintenance of ecological function (Mani, et al., 2021; Skowno, et al., 2018). Anthropogenic land use and land cover changes pose a threat to the ecosystem services that the environment provides to humans and introduces the risk of biodiversity and habitat loss (Musetsho, et al., 2021; Weyer, et al., 2015).

Protected areas (PAs) globally continue to suffer both short- and long-term negative effects from land-use change (Guerra, et al., 2019). PAs are the cornerstone of our global and domestic response to protecting natural ecosystems from the negative consequences of land use/land cover change. However, formal PAs can only do so much – and increasingly there is acknowledgement that we need to look beyond PA fences in to the wider landscape to sustainably manage and respond to the escalating pressure of LULCC (Kleijn, et al., 2020). As a result, more recently the Global Biodiversity Framework 2022 places large emphasis on conservation agreements outside of only formal protection so that we can meet the target of conserving 30% of the earth's land and sea.

There are many conflicts surrounding land use and the distribution of natural resources in South Africa which adds unique difficulties for regulating land use and natural resources in the nation especially for PAs (Shumba, et al., 2021; Sinthumule, et al., 2020). To manage these land use conflicts, many scholars argue both the preservation of biodiversity in PAs and the interactions between people and the land resources must be given equal consideration (Kepe, 2008; Mani,

et al., 2021). Therefore, to ensure the long-term protection of critical biodiversity areas in South Africa, PA management strategies and research needs to be centred around their sustainable use in addition to their socioeconomic benefits (Bourne, et al., 2016).

LULC change around PAs is inevitable and involves complex relationships between conservation decision making and the different land uses within the landscape (Asamoah, et al., 2021). These complexities arise from the interplay of various ecological, socio-economic, and policy factors. For instance, conservation decisions may need to consider the livelihoods of local communities that depend on the land for agriculture, forestry, or other economic activities. At the same time, these land uses could potentially lead to environmental degradation, thereby posing challenges to conservation efforts. Therefore, understanding these complex relationships is crucial for developing effective strategies for sustainable land use and conservation. Biosphere reserves (BRs) offer a unique PA management strategy that aims to protect critical biodiversity areas, whilst ensuring that the socio-economic use of the PAs contained within to continue sustainably (Coetzer, et al., 2013). BRs are regions that take an interdisciplinary approach to managing the interactions between people and the ecological systems which focuses on preventing land use conflict while protecting biodiversity in PAs. BRs make use of a zoning system that consists of three interconnected zones, where the interactions between these zones are managed to achieve a sustainable balance of the trade-offs between biodiversity, ecosystem services, and human use (Palliwoda, et al., 2021). There are ten UNESCO designated BRs in South Africa, with the first designated in 1998, the Kogelberg Biosphere Reserve in the Western Cape, and the most recent in 2018, the Marico Biosphere Reserve in the North West province (Pool-Stanvliet & Coetzer, 2020). For this study, the Garden Route Biosphere Reserve (GRBR), established in 2017, within the Western and Eastern Cape is the area of focus. The Garden Route District Municipality has seen unprecedented LULCC in the past decade (Baard & Kraaji, 2014; Guerbois, et al., 2019), with climate change introducing more challenges for the region (Pandy & Rogerson, 2021).

The Outeniqua Strategic Water Source Area contained within the GRBR, that supplies water to George, Oudtshoorn, and the adjacent Garden Route area, is especially vulnerable to the impact of land management because catchment hydrology is sensitive to dynamic changes in land use (Kibena, et al., 2014; Sauka, 2016). The Kaaimans catchment area, that is part of the Outeniqua Strategic Water Source Area, is the study area of this research work. This catchment area needs up to date LULC data to improve the land management strategies around this area, thus this study established a formal baseline of LULC for this catchment area. Furthermore, catchment

management in the face of climate change scenarios and LULCC is crucial in a water-scarce country like South Africa (Mugari & Masundire, 2022; Woyessa, et al., 2006), thus the study has undertaken the initiative to make future LULC forecasts for the study region.

II Problem statement

South Africa's PAs, like most other countries in Africa, face the challenge of sustainably managing LULCC in and around PAs (De Visser, 2016; Wynberg, 2002). BRs are widely recognized as an effective land management strategy that focuses on governing social-ecological systems while preserving biodiversity inside PAs (UNESCO, 2021; Reed, 2019). BR zonation is a spatially structured gradient of acceptable land utilisation to ensure that environmental and ecological function is less heavily impacted by increased intensity of land use (Meganck, 1990). Thus, for BRs to effectively manage these constant exchanges between the trade-offs between strict PAs versus sustainable development priorities, examining the dynamic LULCC in and around these BRs is important.

The analysis of such LULCC as a mechanism to inform appropriate management intervention can only be done effectively with the use of up-to-date LULCC assessments. But, largely due to the science-policy gap between academics, researchers and reserve managers, policy makers, and land use planners, and the lag between management needs and research response (Rose, et al., 2019), some areas like this study's focus area - the Kaaimans catchment area in the GRBR, lack up to date LULC data and LULC forecasts. The science-policy gap explains how research frequently overlooks the needs of policy and practice because there is not enough exchange of knowledge through interdisciplinary partnerships (Cockburn, et al., 2016). Addressing this need can be done through relevant and responsive research from scientists (Godfrey, et al., 2010); a research-practice relationship where research institutions provide timely, appropriately-scaled data for reserve managers. Thus, co-designing LULC modelling research with BRs in South Africa that show a need for better LULCC management strategies and tools, could help fill this gap between researchers and reserve managers. The aim of this study was to provide a baseline for LULCC in a strategic water management area within the GRBR which lacks this critical and up-to-date data to guide the area's land use planning and related policies.

III Research aim and objectives

The main aim of this research is to provide a current baseline of LULCC including future LULC predictions, for an important water resource management area, the Kaaimans catchment area, within the Garden Route Biosphere Reserve, South Africa.

The research objectives were:

Objective 1: Determine the historical LULC for the Kaaimans catchment area using the South African National Land Cover (SANLC) data of 1990, 2014, 2018, and 2020.

Objective 2: Produce current LULC classification maps of the Kaaimans catchment's priority classes using machine learning and remote sensing satellite data from 2019, 2021, and 2023.

Objective 3: Use the current classification maps (in objective 2) with hybrid Markov Chain models to predict future LULC for the Kaaimans catchment to the year 2030.

IV Significance

This research is significant for several reasons: First, the research aims to provide an up-to-date baseline for LULC for an important strategic water resource area within the GRBR, Eastern and Western Cape. This location has seen unprecedented land use changes like agriculture and built-up expansion at the cost of indigenous vegetation and needs up-to-date land use maps to assist with planning and monitoring of the area's land use. This is an urgent research need that has been identified by GRBR stakeholders (Dr. Bianca Currie, Garden Route Biosphere CEO). By mapping the dynamics of LULC from present to future, the data provided by this study will inform LULC planning and policies within this strategic water resource area, thereby contributing significantly to the GRBR's goal of conserving ecosystems while fostering socio-economic development.

Second, by improving the understanding of the relationships between different land uses and land covers function within our study area over time, more effective management strategies can be implemented by the relevant local authorities, environmental agencies, and community stakeholders. LULC maps allow the examination of the landscape characteristics which are important for understanding the condition of land, the extent of certain land uses, and the state of the natural land cover. The aim of BRs is to have an integrated land management strategy which can only be achieved effectively by having the knowledge of the main land uses in and around a BR and how it has changed over time (UNESCO, 2021). The LULC maps produced by this research provide a comprehensive understanding of the current landscape composition to stakeholders and rights holders throughout the study area.

Finally, this research will also contribute to the theoretical understanding of employing machine learning methods and remotely sensed satellite imagery for mapping current and future land use change. The accuracy of the results from the LULC maps and future LULC predictions based on the satellite imagery, predictive models, classification schemes, and software used for this study's specific study area can be a valuable indicator of the advantages and disadvantages of using these techniques for the desired purpose.

Methodology

a. Study area

The GRBR, established by UNESCO in 2017, is located within the Eastern and Western Cape provinces and has a terrestrial area of 6983.63 km², including a marine portion of 324.87 km² (Figure 1). It falls within the biodiversity hotspot, the Cape Floristic Region along the southern coast of South Africa. The Garden Route National Park (GRNP), which is a UNESCO World Heritage Site, serves as the GRBR's core protected zone with a size of 2123.75 km². Around this core zone, is the buffer zone which is the biggest zone of the GRBR with an area of 2880.32 km² following the smallest zone, the transition zone, which is 1979.56 km² (Figure 2).

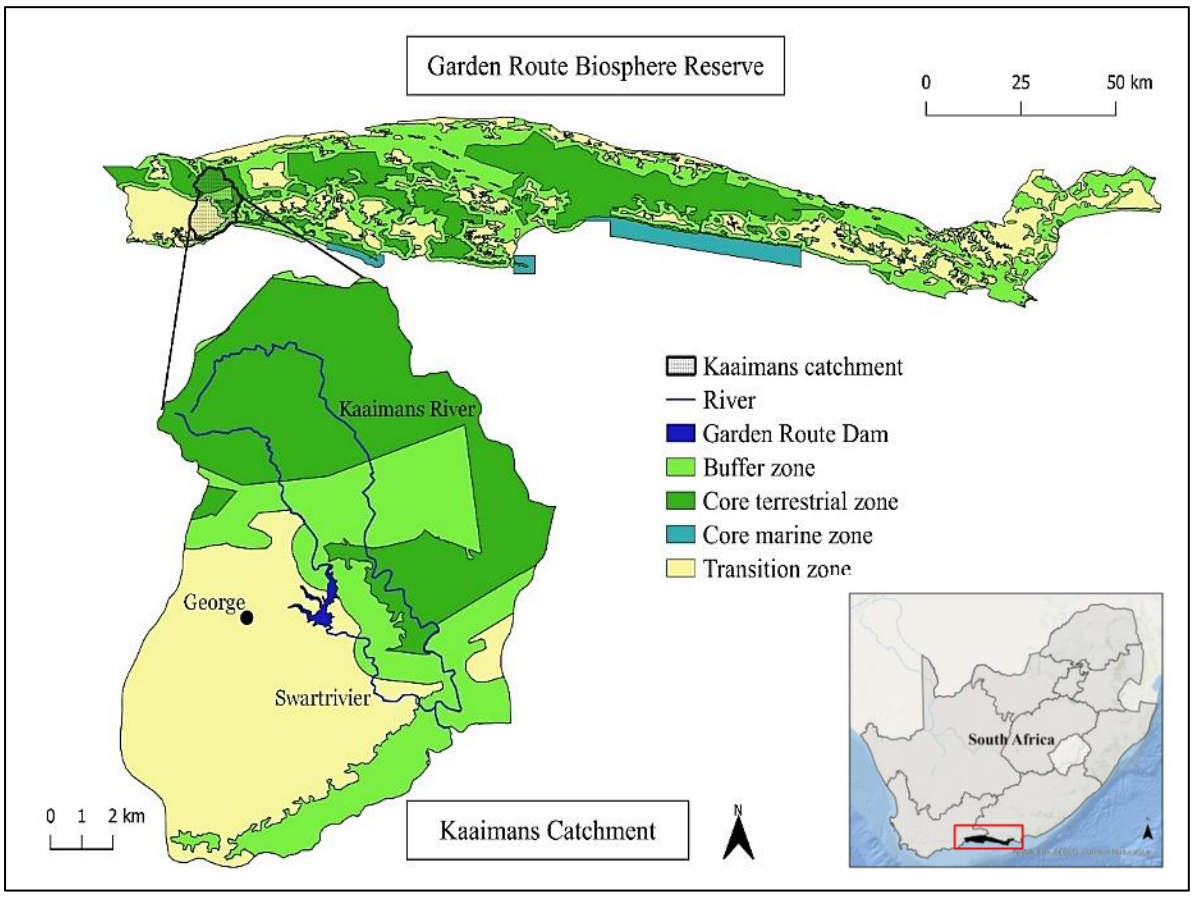

Figure 1: The GRBR and the biosphere zones within the Kaaimans catchment area (quaternary catchment K30C).

The Kaaimans catchment (quaternary catchment K30C; area of 194.78 km²), which is this study's area of interest, has part of The GRNP in the northern and middle eastern regions which forms part of the GRBR's core protected zone. This region is dominated by forests towards the coast and fynbos that occur on the northern most mountainous areas, where the Outeniqua Mountains start. The large transition zone within the study area is where George town is located, which is where urban areas, mining, forestry, agriculture and other human intensive

activities take place. Plantation forests within this quaternary catchment is a large land use type and stretches over the transitional and buffer zones. The plantation forests are mainly composed of the non-native pine tree species called *Pinus elliottii* and *Pinus radiata* that are native to the Southeastern United States and Mexico, respectively (Tuswa, et al., 2019). The *Acacia mearnsii,* known as the Black Wattle, that originate from Australia is a widespread and significant invasive alien tree within this catchment area, and occurs mainly along plantation areas, roads, and river courses (Mukwada & Manatsa, 2017).

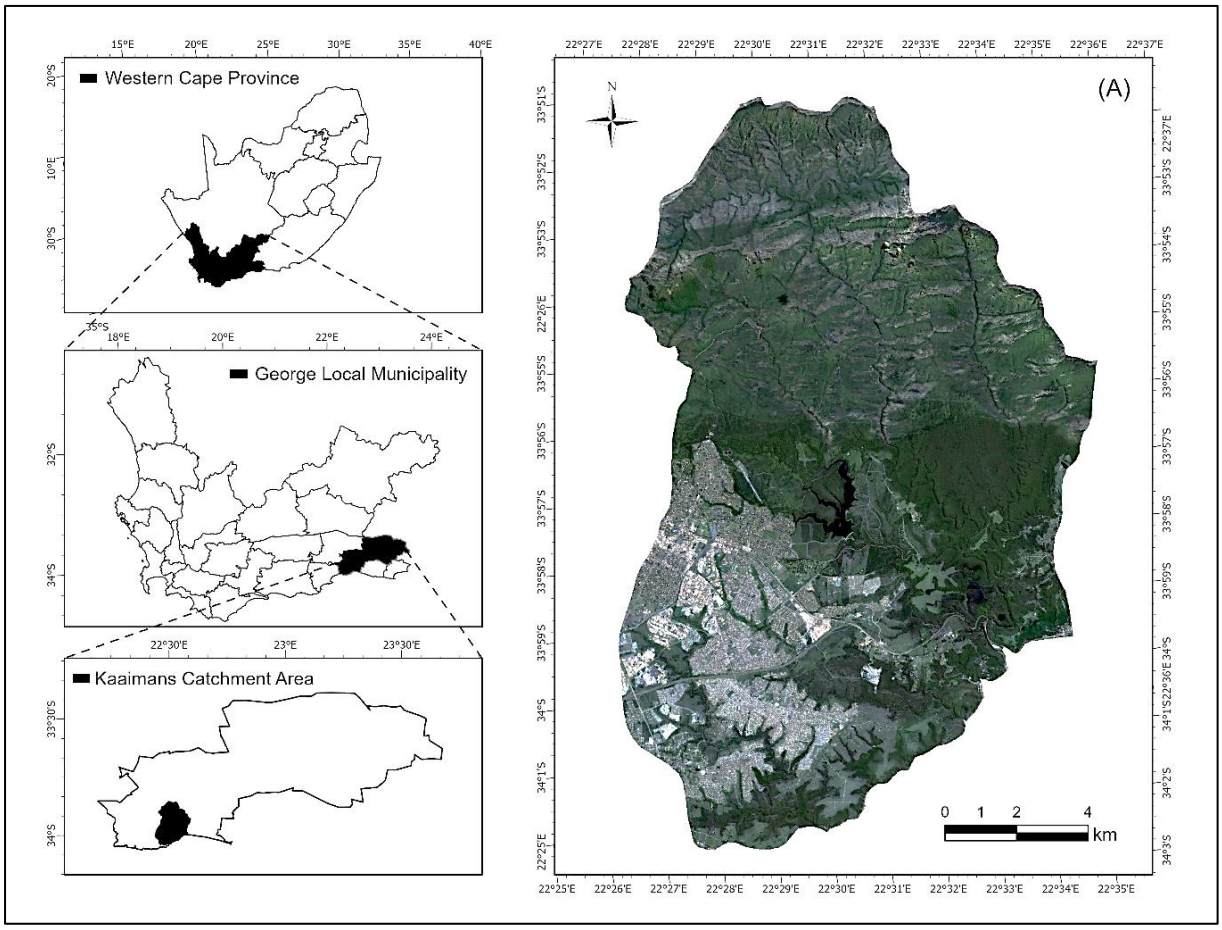

Figure 2: The Kaaimans catchment area within the George Local Municipality in the Western Cape, South Africa. Inset A presents the 2023 true colour Sentinel-2 image of the Kaaimans catchment area.

The GRBR overlaps with the Outeniqua Strategic Water Source Areas that the Kaaimans catchment in the Western Cape forms part of. The Garden Route Dam, which is an important water source for the George Town, and the Kaaimans River Estuary falls within this catchment area. The area has a river system that consists of the Kaaimans river and the Swartrivier.

The George local municipality, one of seven in the Garden Route district, has experienced significant population growth since 2011 to 2023. According to the 2011 Census, the population was just under 200,000. However, the 2022 Census revealed a surprising increase to 294,929 residents. This indicates that since 2011, the population has grown by over 100,000 people, a substantial expansion of 52% (Statistics South Africa, 2023). It is considered by the Department of Cooperative Governance and Traditional Affairs, as the biggest economy in the Garden Route district municipality (2020, p. 21).

This catchment area does not have significant seasonal variation in rainfall and temperature. Annually, the temperature varies between 9 °C and 24 °C (ClimateData.org, 2023). The annual average rainfall for the region ranges between 550 mm and 650 mm (Lakhraj-Govender & Grab, 2019). Seasonally, the most rainfall occurs in the month of November and the lowest amount of rainfall often occurs in the month of June. The Kaaimans catchment's southern region where George Town is located is on a plateau which gives it a relatively flat elevation (based on the 30 m resolution Shuttle Radar Topography Mission), with the northern Outeniqua Mountains that add the variety to the catchment's landscape (Figure 3).

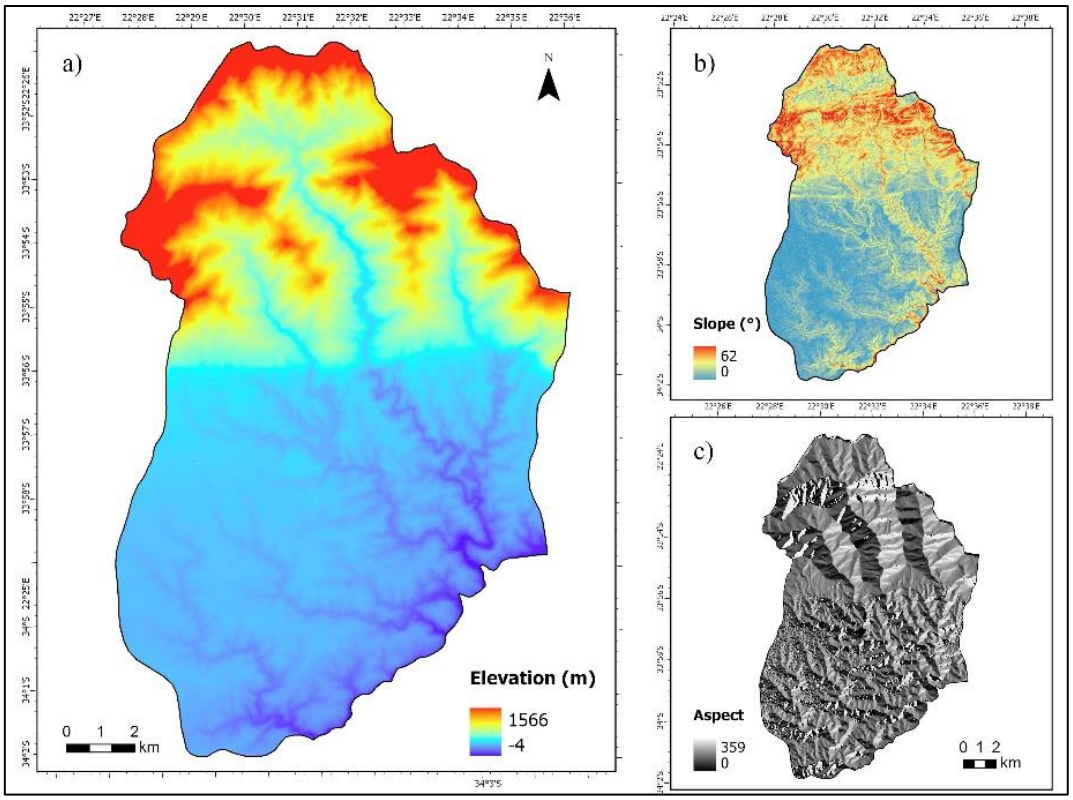

Figure 3: Topographical characterization of the study area presenting: a) elevation in meters, b) terrain slope in degrees, c) the aspect with darker tones indicating south and lighter, north.

b. Data collection and preprocessing

Data for objective 1

For the first objective of this study, the historical LULC for the Kaaimans catchment area was determined using the South African National Land Cover (SANLC) datasets. The datasets used were from the years 1990, 2013/14, 2018, and 2020 (Table 1). The LULC maps for the selected years are created by GEOTERRAIMAGE Pty Ltd and were obtained from the Department of Environmental Affairs GIS (EGIS) website (https://egis.environment.gov.za/data_egis/data_download/current). These datasets were produced by different primary datasets as shown in Table 1. The 1990 and 2013/14 SANLC datasets were produced using Landsat 5 multispectral and Landsat 8 multi-seasonal imagery respectively. These datasets have a 30 m resolution and are ideally suited for approximately 1: 75,000 – 1:250,000 scale mapping and modelling applications (GEOTERRA IMAGE, 2015). The 2018 and 2020 SANLC dataset was produced using 20 m resolution Sentinel-2 satellite imagery. The 1990 SANLC dataset's overall accuracy was not evaluated due to the lack of comparable historical reference data like what was used in the 2013/14 dataset verification. The 2013/14 SANLC dataset overall accuracy based on 6415 sample points is 82.53%. According to the 7 sample points of the 2013/14 SANLC dataset that falls within the study area, the accuracy is 85.71% for the study area. The overall accuracy for the 2018 SANLC dataset is calculated from 6570 reference points and is 90.14%, and based on the 9 reference points within the study area, the accuracy is 88.89%. The overall accuracy for the 2020 SANLC dataset, based on 6835 reference points, is 85.47%, while there is only one reference points in the study area, the specific accuracy for the study area with this dataset cannot be determined.

Table 1: South African National Land-Cover datasets used for the four classification maps of objective one.

Data	Year	Number of Classes	Primary datasets
1990 South African National Land-Cover Dataset (SANLC 1990)	1990	72	Multi-seasonal Landsat 5 Satellite Imagery, acquired between 1990 and 1991.
2013-2014 South African National Land Cover Dataset (SANLC 2013/14)	2013-2014	72	Multi-seasonal Landsat 8 Satellite Imagery, acquired between April 2013 and March 2014.
2018 South African National Land Cover (SANLC 2018)	2018	73	Multi-seasonal Sentinel 2 Satellite Imagery, acquired between January 2018 and December 2018.
2020 South African National Land Cover (SANLC 2020)	2020	73	Multi-seasonal Sentinel 2 Satellite Imagery, acquired between January 2020 and December 2020.

Data for objective 2

The data used for the second objective, which consisted of satellite imagery, training, and validation data points, was crucial for producing current LULC classification maps of the Kaaimans catchment's priority classes. This analysis was carried out using machine learning and remote sensing satellite data from 2019, 2021, and 2023.

Satellite imagery

The data for objective two was the Sentinel-2 Level-2A (L2A) atmospherically corrected Surface Reflectance satellite imagery from the Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus/#/home) which provides free and open access to Sentinel satellite imagery accessed and processed through Google Earth Engine (GEE) (Google, n.d.). Compared to other freely available multispectral satellite data, Sentinel-2 data covers a larger spatial extent, has a higher spatial resolution (up to 10m), and a higher spectral resolution i.e., three red-edge bands and two SWIR bands that make up part of the 13 available spectral bands. Sentinel-2 data also has high temporal resolution (shorter revisit period) of up to 5 days, that allows for more data collection, which is useful for monitoring LULC changes over time (Phiri, et al., 2020). The spectral bands of Sentinel-2 are particularly useful for LULC classification and allows for the discrimination between various features based on their reflectance values (Table 2). Sentinel-2 red-edge bands give useful information for differentiating vegetation types and boosting land cover classification accuracy, especially when there is more than one type of vegetation class (Misra, et al., 2020; Qui, et al., 2017). Bands 9 and 10 were omitted from further analysis due to their lack of utility for LULC classification. Band 9's large pixel size (60m) limits detail, while Band 10, the cirrus band, lacks surface information.

Table 2: The 10 spectral bands of Sentinel-2 L2A imagery used for this study (bands 9 and 10 were excluded).

Spectral band	Name	Centre Wavelength (nm)	Band Width (nm)	Spatial Resolution (m)
B2	Blue	490	65	10
B3	Green	560	35	10
B4	Red	665	30	10
B5	Red-edge 1	705	15	20
B6	Red-edge 2	740	15	20
B7	Red-edge 3	783	20	20
B8	Near infrared	842	115	10
B8A	Near infrared narrow	865	20	20
B11	Shortwave infrared 1	1910	90	20
B12	Shortwave infrared 2	2190	180	20

The satellite images were gathered during the peak seasons of vegetation growth, specifically between November and March, across the years 2019, 2021, and 2023 (Table 3). This is because during these peak seasons, the different land cover types, such as forests, and shrublands are at their most distinctive stages. The basic preprocessing of the Sentinel-2 Level-2A satellite imagery was done in GEE (Google, n.d.). This entailed resampling the images to the same spatial resolution of 10 m, using the default nearest neighbour method as it is recommended that all bands are the same resolution (Richards & Jia, 2006). Cloud masking was done using the QA60 band to have cloudless imagery for classification. To reduce the effects of shadows on the imagery, especially at the northern mountainous areas of the study area, satellite images of an illumination zenith angle of less than thirty was selected.

Table 3: Sentinel-2 L2A satellite imagery collected from the Copernicus Open Access Hub via GEE.

Satellite imagery	Cloud cover (%)	Date	Name
Sentinel-2 Level-2A	0.06	7 January 2023	S2A_MSIL2A_20230107T081321_N0509_R078_T34HFH_20230107T120151
	0.07	3 November 2021	S2A_MSIL2A_20211103T081111_N0301_R078_T34HFH_20211103T103304
	0.03	7 February 2019	S2A_MSIL2A_20190207T081111_N0211_R078_T34HFH_20190207T111505

The priority classes will reflect the dominant LULC classes within the study area. The training data collection for the classifications consisted of the seven LULC classes (Table 4).

Table 4: Description of the priority classes identified and used for LULC classification of the Kaaimans catchment area.

Class no.	Class name	Description
1	**Plantation forest**	These are areas where trees have been planted for commercial use, and are non-native species.
2	**Natural forest**	These are areas covered by indigenous trees and vegetation.
3	**Water**	This includes bodies of water such as dams, rivers, and estuaries.
4	**Built-up**	These are areas with human-made structures and surfaces, including residential, commercial, and industrial buildings.
5	**Bare land**	This refers to areas of land with little to no vegetation or human-made structures.
6	**Agriculture**	These are areas used for farming and agricultural activities.
7	**Shrubland**	This refers to areas dominated by shrubs and small trees, mostly native species.

Training data collection

The quality of the training datasets is important for supervised classification accuracy and each class defined spectrally for accurate classification (Foody, et al., 2016; Kavzoglu, 2009). The size of the samples per class and the sampling strategy are determinants of the quality of the training data. The training data collection consisted of georeferenced points collected from Google Earth Pro, and the National Geo-spatial Information (NGI) aerial imagery that were grouped based on the priority classes identified (Table 4). The training datasets used for the three RF classifications (years 2019, 2021, and 2023) were randomly distributed and the size of the training points per LULC class was representative of actual class proportions in the landscape. To ensure that the training data points for each class fully represented most of the LULC class variation, core and transitional training points were collected per class. The core training points were collected in the middle or homogenous areas of a specific class, while the transition training points were collected on the borders between classes. This approach has shown to help improve the accuracy of the classification model (Foody, 1990). Thus, stratified random sampling of the seven classes for each image was done with the focus of gathering both transition and core training points. A total of 1099, 985, and 997 training samples were created for the training and testing samples for the 2019, 2021, and 2023 satellite images respectively. The training data points of each satellite image were divided into 70% for training, and 30% for testing using stratified random sampling in RStudio (Table 5, Table 6, and Table 7). This method prevents model overfitting and ensures good performance on new data. Stratified random sampling guarantees each land cover class is well-represented in both training and testing datasets, enhancing the reliability of the LULC mapping.

Table 5: The training sample split of the 2023 Sentinel-2 L2A satellite image for classification.

Class no.	Class name	Training samples	Testing samples	Total samples:
1	**Plantation forest**	175	75	**250**
2	**Natural forest**	106	45	**151**
3	**Water**	33	14	**47**
4	**Built-up**	189	80	**269**
5	**Bare land**	122	52	**174**
6	**Agriculture**	61	26	**87**
7	**Shrubland**	85	36	**121**
	Total samples:	**771**	**328**	**1099**

Table 6: The training sample split of the 2021 Sentinel-2 L2A satellite image for classification.

Class no.	Class name	Training samples	Testing samples	Total samples:
1	**Plantation forest**	126	54	**180**
2	**Natural forest**	107	45	**152**
3	**Water**	33	13	**46**
4	**Built-up**	178	76	**254**
5	**Bare land**	88	37	**125**
6	**Agriculture**	80	33	**113**
7	**Shrubland**	81	34	**115**
	Total samples:	**693**	**292**	**985**

Table 7: The training sample split of the 2023 Sentinel-2 L2A satellite image for classification.

Class no.	Class name	Training samples	Testing samples	Total samples:
1	**Plantation forest**	161	69	**230**
2	**Natural forest**	119	51	**170**
3	**Water**	28	12	**40**
4	**Built-up**	164	69	**233**
5	**Bare land**	63	27	**90**
6	**Agriculture**	101	43	**144**
7	**Shrubland**	63	27	**90**
	Total samples:	**699**	**298**	**997**

Furthermore, analysing the spectral reflectance curves of the mean extracted reflectance values for each model's sample points is crucial for several reasons. Firstly, these spectral curves provide a visual representation of how different LULC classes reflect light across various wavelengths (Sentinel-2 bands), which can help in identifying and distinguishing between these classes. Secondly, by comparing the spectral reflectance curves of different sample points, we can gain insights into the variability within and between different classes of LULC. By comparing these curves with the average spectral signatures that are expected of each class at certain wavelengths, we can make sure that the sample points collected for each model year makes sense. This comparative analysis not only ensures the logical consistency of the sample data but also enhances the reliability of each model's predictive capabilities. To make sure the training datasets collected for each of the three satellite image dates are representative of each class, the spectral signature curves were plotted based on the mean reflectance values extracted from the corresponding satellite image in RStudio (Figure 4; Figure 5; Figure 6). This data analysis process provided insights into the variability within each class and between different classes. Using these spectral profiles of the training points that signify each LULC class, it was determined that the training data collected represented distinct spectral responses for each class in the three datasets and was suitable for classification.

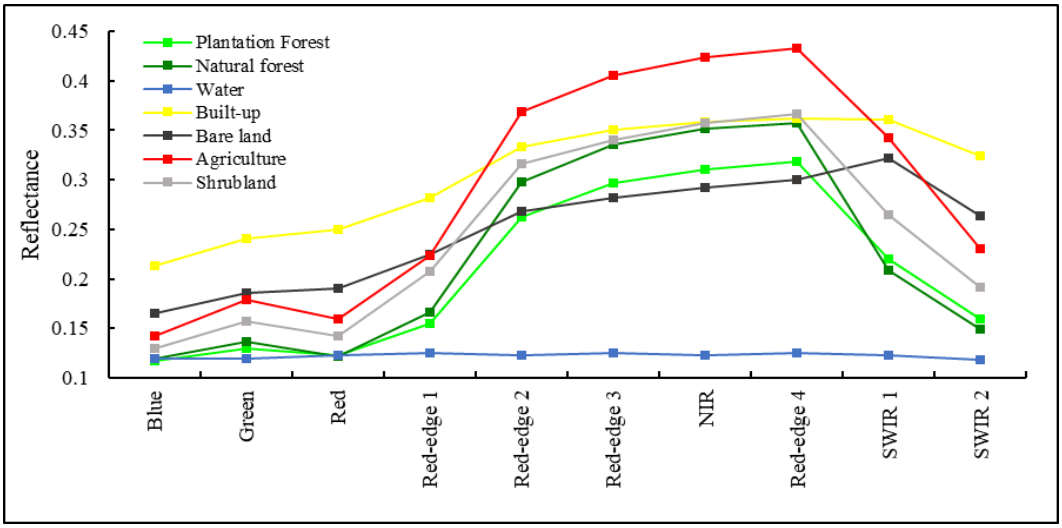

Figure 4: The spectral signature curves of each class for the 2023 training data points.

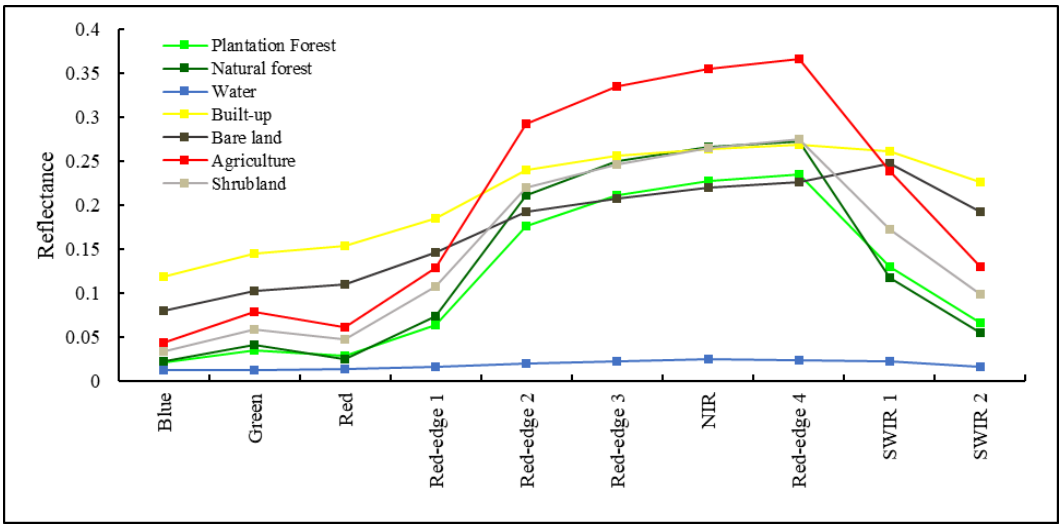

Figure 5: The spectral signature curves of each class for the 2021 training data points.

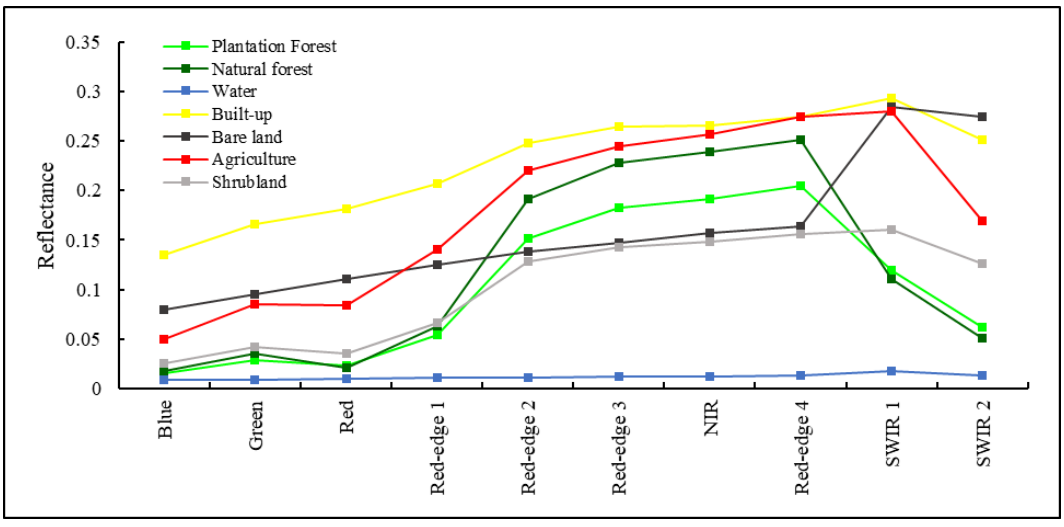

Figure 6: The spectral signature curves of each class for the 2019 training data points.

Data for objective 3

The data used for objective 3 consisted of three LULC classification maps, and seven driver variables which will be used for 2030 LULC predictions of the Kaaimans catchment area.

Classification maps

The data used for objective 3 (future LULC predictions for the Kaaimans catchment area) were based on the three LULC classification images (2019, 2021, and 2023) from the results of objective 2 (supervised LULC classification).

Driver variables dataset

The static and dynamic driver variables which were used for the transition model consist of topographical and land use driving factors. The most common variables used for predicting future LULC within the Land Change Modeller (LCM) prediction tool in TerrSet 2020 software are slope, elevation, aspect, and distance from main roads, and distance from built-up areas (Clark Labs, 2022; Kim, et al., 2020). Based on the LULC dynamics and the author's knowledge of the study area, the distance from plantations and distance from rivers (which includes water bodies) were also included as potential driving factors. The distance from built-up and the distance from plantations were used as dynamic variables whereas the rest were used as static variables. The slope and aspect variables were created in ArcGIS Pro using the digital elevation model (DEM) that was extracted from the NASA Shuttle Radar Topography Mission (SRTM) 30 m resolution DEM (Farr, et al., 2007). The built-up area, main roads, and rivers were made into vector layers using the 2023 satellite image. The distance from plantation forest vector layer was created by using the Google Earth Pro historical imagery and the 2023 satellite image to locate plantation lands within the study area. These vector layers were transformed to raster and changed to distance maps in ArcGIS Pro using the "Euclidean distance" tool.

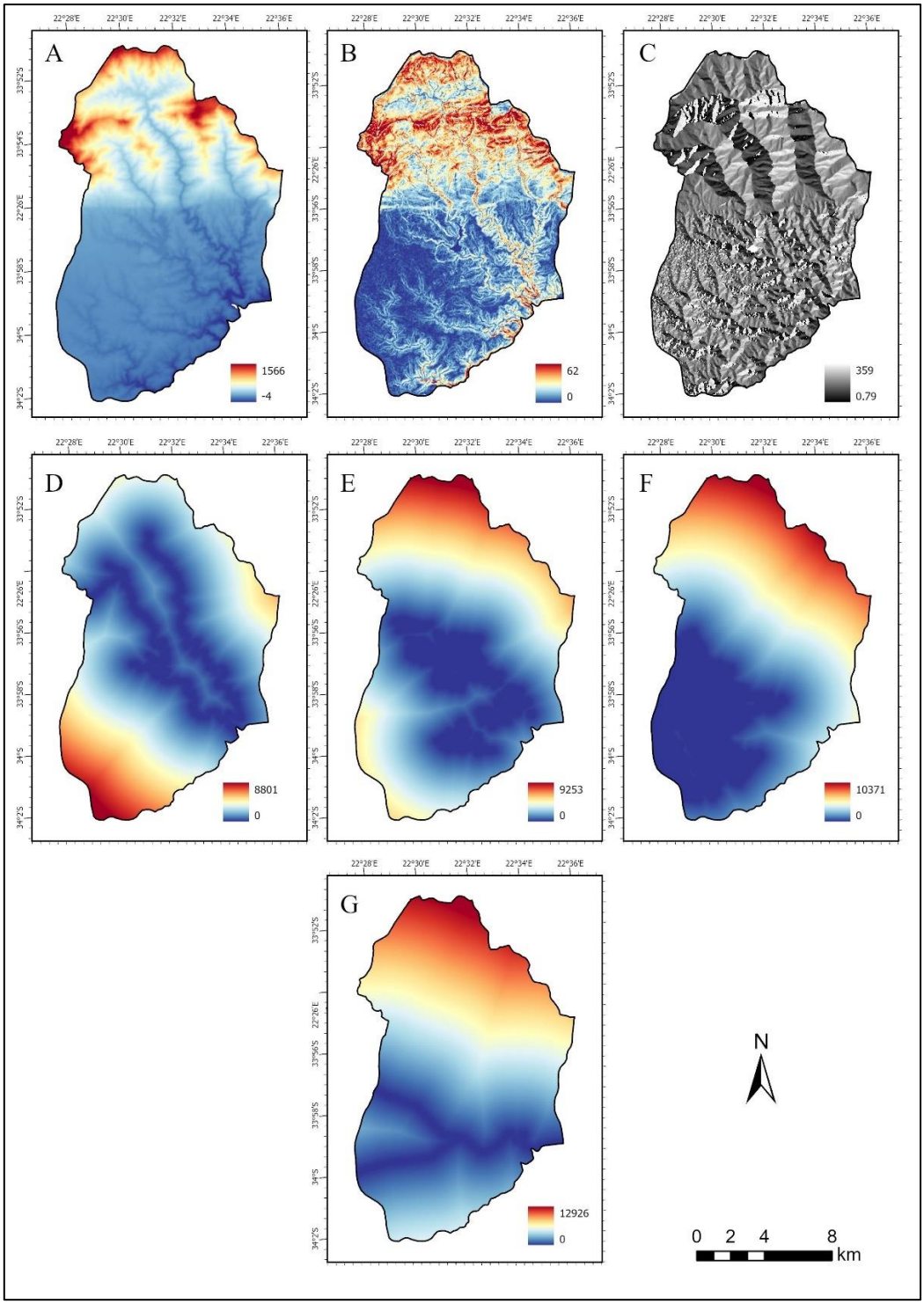

Figure 7: The static and dynamic driver variables as input to the LCM: A is elevation; B is slope; C is aspect; D is distance from rivers (including water bodies); E is distance from plantations; F is distance from built-up; D is distance from main roads.

Chapter 1: Objective 1

1.1 Introduction

This chapter focuses on the first objective of examining the historical LULC for the Kaaimans catchment area, in the Western Cape, South Africa. The analysis is based on the South African National Land Cover datasets from the years 1990, 2013/14, 2018, and 2020.

The aim of this chapter is to understand the historical changes in LULC within the Kaaimans catchment area over these time periods. This analysis will provide a comprehensive overview of how LULC have evolved over three decades, offering valuable insights into patterns and trends that have shaped the current state of the catchment area. In the subsequent sections, the methods used to analyse the SANLC datasets, and the results are explained. The dominant LULC classes identified from the results of this chapter were used for the classification of the study area's current LULC.

1.2 Methods

Using QGIS software with the Semi-Automatic Classification Plugin, post-processing of the SANLC datasets was done to standardize the classification schemes of each dataset (QGIS Development Team, 2002; Table 8). Resampling to a spatial resolution of 10 m was done in ArcGIS Pro software for each classification image to harmonize the datasets (ESRI, 2023).

Table 8: The reclassification of each of the SANLC datasets into 13 LULC classes.

Class Name	New Classes	NLC 1990	NLC 2013/14	NLC 2018	NLC 2020
Indigenous forest	1	4	4	1; 2	1; 2
Woodland	2	6	6	3; 4	3; 4
Forestry	3	32; 33; 34	32; 33; 34	5; 6; 7	5; 6; 7
Low shrubland	4	8; 9	8; 9	8; 9	9
Grassland	5	7	7	13	13
Waterbodies	6	1; 2	1; 2	15; 16; 18; 19; 21	14; 15; 16; 18; 19; 21
Wetland	7	3	3	22; 23	22; 23
Bare land	8	41	41	25; 26; 29; 30; 31	25; 26; 29; 30; 31
Commercial agriculture	9	10; 11; 12	10; 11; 12	32; 33; 39; 40	32; 33; 39; 40
Fallow land	10	29; 27	29; 27	42; 43; 44; 45; 46; 73	42; 43; 44; 45; 46; 73
Built-up area	11	42; 43; 44; 46; 48; 49; 50; 51; 52; 57; 58; 59; 61; 62; 63; 64; 69; 70; 71; 72	42; 43; 44; 46; 47; 48; 49; 50; 51; 52; 57; 58; 59; 60; 61; 62; 63; 64; 69; 70; 71; 72	47; 48; 49; 50; 51; 53; 54; 55; 56; 61; 63; 65; 66; 67	47; 48; 49; 50; 51; 53; 54; 55; 56; 57; 58; 59; 60; 61; 63; 65; 67; 72
Non-commercial agriculture	12	23; 24; 25; 53; 54; 55; 56	23; 24; 25; 53; 54; 55; 56	57; 58; 59; 60	57; 58; 59; 60
Mine site	13	35; 36; 38; 39	35; 36; 38; 39	69	69
Land-Cover classes not within study area		13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 26; 28; 30; 31; 37; 40; 45; 47; 60; 65; 66; 67; 68; 72	13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 26; 28; 30; 31; 37; 40; 45; 65; 66; 67; 68	10; 11; 12; 14; 17; 20; 24; 27; 28; 34; 35; 36; 37; 38; 41; 62; 64; 68; 70; 71; 72	11; 41; 8; 10; 12; 17; 20; 24; 27; 28; 34; 35; 36; 37; 38; 52; 62; 64; 68; 70; 72

The classes used for the reclassification of the SANLC datasets were selected based on their suitability for the study region and comparability across the four datasets (Table 8 and 9). Following standardisation and reclassification of the four SANLC maps, quantitative change analysis was done based on the area percentage and area in square kilometres.

1.3 Results

The 1990, 2013/14, 2018, and 2020 land-use maps were reclassified into 13 classes (Table 9) to attain a standardized classification scheme across the four datasets.

Table 9: Description of land-cover classes used in the study (adapted from DEA E1434 Land-Cover: South African National Land-Cover 2018 Report & Accuracy Assessment., 2019).

Class Name	Classes	Description of Class
Indigenous forest	1	Natural or semi-natural indigenous forest characterized by towering trees, with tree canopy densities normally greater than 75% and canopy heights typically greater than 5 m, connected with various understory vegetation canopies.
Woodland	2	Dense bush, thicket, woodland and open bush, thicket, woodland vegetation with canopy heights around 2.5 m.
Forestry	3	Planted tree forests that consist primarily of exotic pine, eucalyptus, and wattle tree types. Represents a combination of young, temporary, and mature, clear-felled stands.
Low shrubland	4	Natural, low vegetation that can consist of fynbos and other regions of woody shrublands.
Grassland	5	Natural / semi-natural vegetation areas characterized by grass as the dominant type, with bush and/or tree densities ranging from 20% to 40%, depending on canopy heights.
Waterbodies	6	Open surface water areas. Water bodies, whether man-made or natural, can be flowing, static, fresh water, or salty.
Wetland	7	Natural or semi-natural wetlands that include both riparian and emergent aquatic vegetation in pans. The wetlands can either be covered by permanent or seasonal herbaceous vegetation.
Bare land	8	Consists of natural rock surfaces of natural non-vegetated areas, but also includes and not limited to dry pans, coastal sand dunes, beach sand, and bare riverbeds.
Commercial agriculture	9	Lands used for producing commercial large scale agricultural crops, which included but is not limited to permanent orchards, permanent vines, annual crops that are non-pivot irrigated, and annual rain-fed crops.
Fallow land	10	Consists of long term non-active old fields that are overgrown by trees, woody, grass, bare, and low scrub cover.
Built-up area	11	Includes all built-up areas. This consists of formal and informal residential areas, villages, recreational, commercial, and industrial infrastructure along with roads.
Non-commercial agriculture	12	Peri-urban areas where agricultural fields are located between smallholdings, either dominated by trees, bushes, low vegetation, or grass.
Mine site	13	Extraction sites such as open cast mines and quarries, which are non-vegetated and surface-based mining activities.

The spatial temporal analysis of historical land-use patterns of the Kaaimans catchment can be seen in the four SANLC classification maps below (Figure 5). Based on the visual comparison of the maps, the historical spatial distribution of the LULC classes in the catchment's northern half remained relatively consistent between 1990 and 2020, with this area being mostly dominated by indigenous forest (22.79%), woodland (15.65%), shrubland (18.88%), and grassland (5.74%) (Table 10). This visual observation of no significant change in the northern region between those four classes is supported statistically by an ANOVA test ($F[3, 12] = 3.33$; $p = 0.07$). This specific area forms the core terrestrial zone of the GRBR and part of GRNP which is a protected area made up of primarily natural vegetation (Figure 1). The middle and southern half are predominantly occupied by commercial forestry (8.41%) (consistently around the Garden Route Dam), urban areas (16.31%), and agriculture fields (3.19%) (Figure 9; Table 10). There are significance changes present in this southern half region that is dominated by human driven land uses ($F[2, 9] = 27.93$; $p = 0.00$).

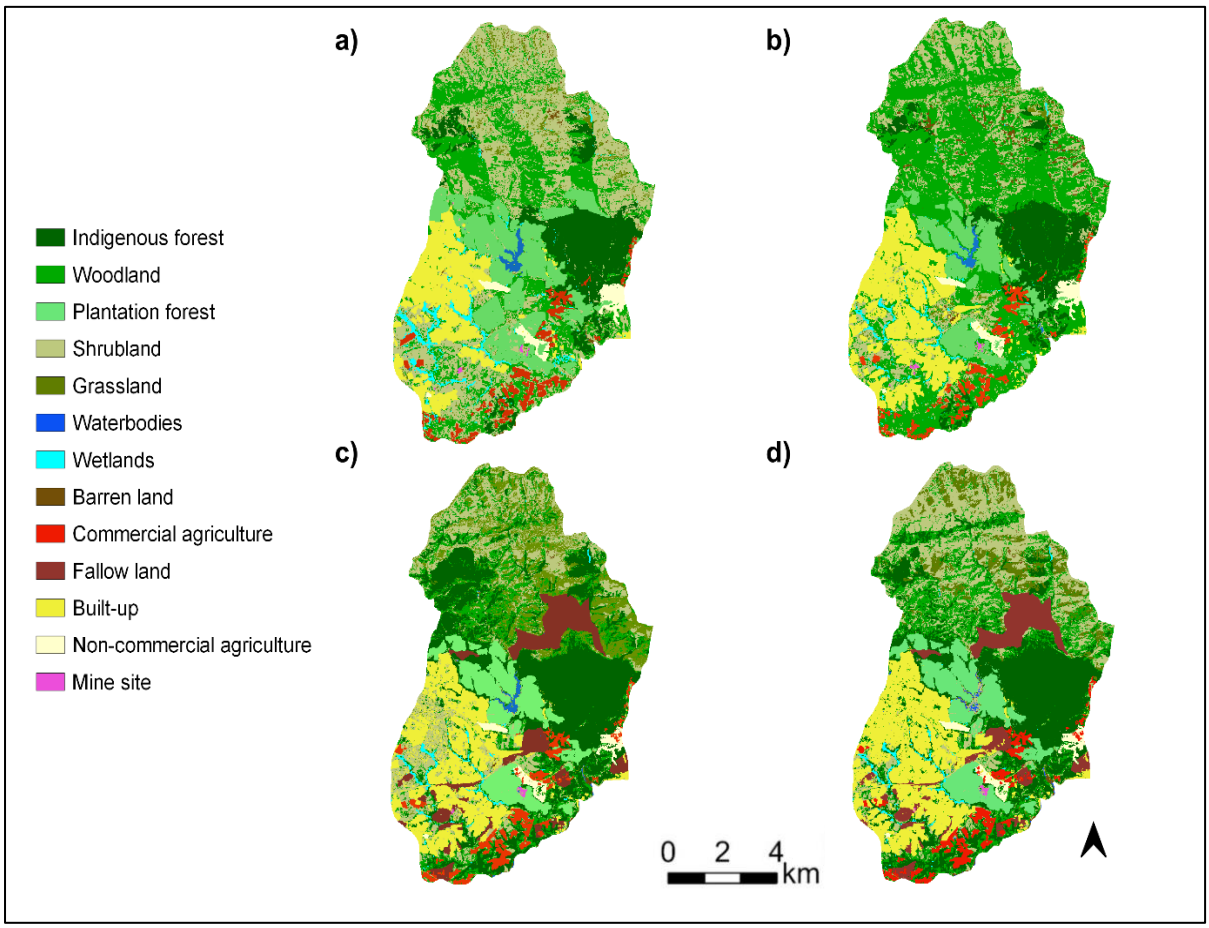

Figure 8: Spatial temporal distribution of the historical LULC in the Kaaimans catchment based on the SANLC 1990 (a), 2013/14 (b), 2018 (c), and 2020 (d) datasets.

It is noted that the Fallow land class was not included within the accuracy assessment for the 2018 and 2020 datasets because the class is based on historical field boundaries which is not visible on the imagery used (SANLC 2018 Accuracy Assessment Report, 2021). Therefore, it is uncertain if the large areas indicated as fallow land in the 2018 and 2020 image are accurate. Between 1990 and 2020, the notable visible changes in LULC classes were the increase of indigenous forest and built-up areas (Figure 9; Table 10). The shrubland class has seen a decrease in that time. The plantation forest, commercial and non-commercial agriculture LULC classes remained relatively constant, with slight changes in their spatial distribution. Overall, the LULC of the Kaaimans catchment from 1990 to 2020 shows a decline in shrubland and woodland, with an increase of indigenous forest, built-up, and grassland LULCs.

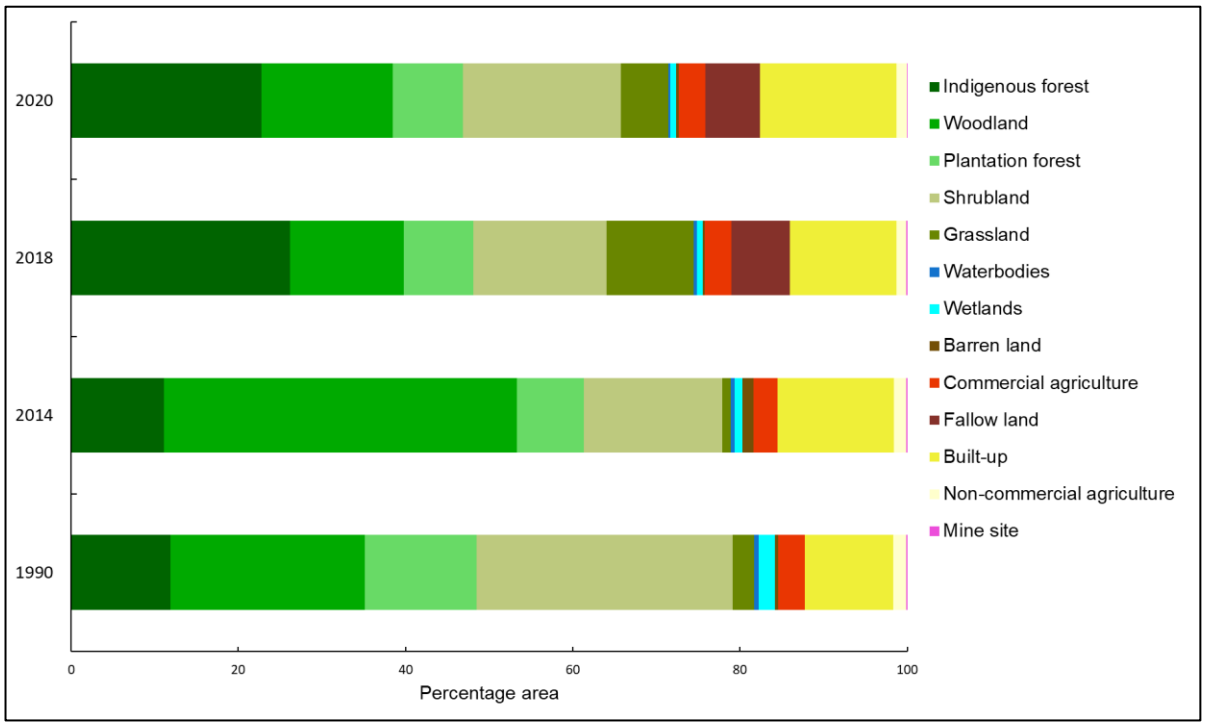

Figure 9: The percentage area change of the 13 LULC classes of the 1990, 2013/14, 2018, and 2020 SANLC datasets for the Kaaimans catchment area.

The dominant LULC class types identified within the Kaaimans catchment is the indigenous forest, woodland, plantation forest, shrubland, and built-up with areal extents of 22.79%, 15.65%, 8.41%, 18.88%, and 16.31% of the total study area based on the 2020 SANLC dataset, respectively (Table 10).

Table 10: The area statistics for the historical LULC in the Kaaimans catchment based on the four SANLC datasets.

Class Name	SANLC 1990		SANLC 2014		SANLC 2018		SANLC 2020	
	Area (%)	Area (km²)	Area (%)	Area (km²)	Area (%)	Area (km²)	Area (%)	Area (km²)
Indigenous forest	11.87	22.64	11.10	21.19	26.22	49.84	22.79	43.32
Woodland	23.28	44.41	42.20	80.51	13.56	25.77	15.65	29.74
Plantation forest	13.37	25.50	8.02	15.31	8.36	15.89	8.41	15.98
Shrubland	30.62	58.43	16.59	31.65	15.89	30.21	18.88	35.89
Grassland	2.56	4.88	1.01	1.93	10.42	19.80	5.74	10.91
Waterbodies	0.54	1.03	0.40	0.77	0.40	0.77	0.19	0.37
Wetlands	1.98	3.77	0.97	1.84	0.73	1.39	0.73	1.39
Barren land	0.33	0.62	1.33	2.54	0.19	0.35	0.27	0.51
Commercial agriculture	3.24	6.19	2.86	5.46	3.20	6.09	3.19	6.06
Fallow land	0.00	0.00	0.00	0.00	7.02	13.35	6.57	12.50
Built-up	10.53	20.09	13.91	26.54	12.72	24.18	16.31	31.00
Non-commercial agriculture	1.61	3.08	1.51	2.87	1.21	2.30	1.20	2.29
Mine site	0.08	0.15	0.09	0.17	0.09	0.16	0.08	0.15
Total	**100.00**	**190.79**	**100.00**	**190.79**	**100**	**190.11**	**100**	**190.11**

According to the annual rate of change of the LULC classes of the SANLC datasets, the catchment area undergoes various spatial temporal LULC changes annually (Table 11). Indigenous forest and grassland exhibited yearly reductions during the period from 2018 to 2020, however, these decreases were not significant ($F[1, 2] = 4.92$; $p = 0.27$). Built-up area, shrubland, and woodland class types show a significant increase for this period ($F[2, 3] = 148.51$; $p = 0.01$). Overall, the catchment area has undergone significant changes in its LULC from 1990 to 2020. Based on the standard deviations of the annual rate of change periods as shown in Table 11, the high standard deviation (4.18 %) of the 2014 to 2018 period indicates that there were significant variations in the changes in the LULC during this period. The low standard deviation of 0.38 % for the 1990 to 2013/14 period suggests that the changes in LULC were relatively uniform. For the 2018 to 2020 period, the standard deviation of 1.56 % suggests moderate variations in the changes in LULC.

Table 11: The annual rate of change (%) of the 13 LULC classes in the catchment area for the 1990 to 2020 SANLC datasets.

	1990-2013/14	2014-2018	2018-2020
Indigenous forest	-0,06	7,16	-3,26
Woodland	1,50	-13,69	1,98
Plantation forest	-0,42	0,15	0,05
Shrubland	-1,12	-0,36	2,84
Grassland	-0,12	4,47	-4,44
Waterbodies	-0,01	0,00	-0,20
Wetlands	-0,08	-0,11	0,00
Barren land	0,08	-0,55	0,08
Commercial agriculture	-0,03	0,16	-0,01
Fallow land	0,00	3,34	-0,42
Built-up	0,27	-0,59	3,41
Non-commercial agriculture	-0,01	-0,14	0,00
Mine site	0,00	0,00	-0,01
Standard deviation:	0.38	4.18	1.56

1.4 Discussion Chapter 1

This chapter analysed the spatial-temporal change of LULC in the Kaaimans catchment using the SANLC maps of the years 1990, 2013/14, 2018, and 2020. The results provided a clear understanding of the historical LULC trends of the study area. It was observed that the period from 1990 to 2020 experienced significant LULC changes, primarily due to the transitions among different LULC classes, as depicted in Figures 8 and 9. It was interesting to find that, in terms of spatial arrangement, the study area shows a stark contrast between the natural and human-influenced areas within the Kaaimans catchment. Approximately half of the study area is dominated by natural vegetation, such as indigenous forest, woodland, shrubland, and grassland. This area, forming the core terrestrial zone of the GRBR and part of the GRNP, serves as a testament to the effectiveness of conservation efforts. Thus, this region is expected to remain naturally vegetated by mainly shrublands and indigenous forest when assessing the most recent LULC in the following chapters. On the other hand, the southern half of the study area shows the trends of expanding human activity. Dominated by commercial forestry, urban areas, and agriculture fields, this region reflected the dynamic interplay between development and land-use change. The coexistence of these contrasting landscapes within the same catchment area provided a unique opportunity to study the interactions between natural ecosystems and human-induced changes, offering valuable insights for sustainable land-use planning and management strategies.

A critical reflection on the SANLC datasets from 1990, 2013/14, 2018, and 2020 reveals that the different methodological approaches used in these datasets may have led to some noise in the absolute changes observed. The methods change through the years and although standardisation was applied for the assessment here, some of those methodological artefacts likely remain. For example, the inclusion of the fallow land class only in the 2018 and 2020 datasets introduces an additional layer of uncertainty, as the classification of this land type in previous datasets is unclear. Moreover, each dataset does not contain a significant number of validation data points in the study area to validate the maps and determine the accuracy of these datasets specific to the study area. This suggests that these datasets are more suited for small-scale assessments covering larger areas rather than smaller areas. However, despite these limitations, the SANLC datasets provide a valuable overall understanding of the LULC changes, serving as a useful starting point for more detailed, localised studies in South Africa.

When looking at the northern half of the study area, the results showed that the major land cover transitions were between bare land and the natural vegetation classes. This decrease and

increase of bare land and natural vegetated areas in that region suggests likely, natural processes play an important role in driving those changes. For example, the catchment's location in the Western Cape province of South Africa, is known to be a fire-prone environment due to the dominance of fynbos vegetation type in the area (Msweli, et al., 2020). Therefore, the observed interchange between bare land and natural vegetation could potentially be attributed to the occurrence of both unplanned wildfires and strategically implemented controlled burnings. Controlled burnings by South African National Parks (SANParks) are strategic fire management methods to enhance biodiversity and mitigate wildfire risks. They rejuvenate the landscape, improve soil quality, and make the area more resistant to dangerous wildfires (Rogers, 2019). These could explain the observed interchange between bare land and natural vegetation. However, it is important to consider other factors as well. While fires may play a role, the changes could also be due to human-driven activities such as clearing (for example, alien plant clearing). Numerous studies conducted in the region have highlighted the significant impact of invasive alien plant species on water resources and these species are known to consume substantial volumes of water, thereby reducing its availability (Le Maitre, et al., 2014; Le Maitre, et al., 2016; van Wilgen, et al., 2016). This is particularly concerning for a water-scarce country like South Africa, where maintaining water resources is of paramount importance (Tanner, et al., 2022). In this case, while revegetation post-clearing may be a natural process, the creation of the bare patches is a result of anthropogenic intervention. It is crucial to note that without definitive evidence, we must consider all possible causes for these transitions. Therefore, both natural processes and human activities are likely contributing to the observed land cover changes.

The middle and southern half region is predominantly occupied by LULC classes associated with anthropogenic land use and the changes between these classes are primarily driven by human activities e.g., built-up, commercial, and non-commercial agriculture, and mining. The area used for commercial and subsistence agriculture was expected to increase to sustain the expanding urban area and population growth of George Town (Tizora, 2016), but the area used for agriculture stayed relatively stable across the analysis period. This could be attributed to the fact that the agricultural sector in the Garden Route is entirely reliant on water resources (Steyn, 2013), which may have limited significant expansion in this area.

The plantation forest land serves as a dividing line between the natural and urban halves of the study area, occupying a substantial portion of the land (8.49% in 2020). There has been a noticeable increase in the area covered by this land use class from 2014 to 2020, and this trend

is expected to persist. Plantation forests are a big industry for the area, and it provides a significant source of income and employment for the local population (Department of Agriculture, Forestry and Fisheries, 2023). However, the presence of plantation forests adjacent to indigenous forests in this region potentially exacerbates the risk and severity of fire penetration, compared to if the indigenous forests were next to native shrublands (Gidday, et al., 2021). This is because plantation forests through increasing artificial aboveground biomass, provide more fuel for veld fires, and the encroachment of the plantation exotic species into the shrubland, woodland, and indigenous forest areas increase the risks of more frequent and severe veldfires (Kraaij, et al., 2018). Thus, the expansion of plantation forests and the encroachment of the exotic species introduced by these plantations into the natural vegetation is important to monitor in the catchment area.

1.5 Conclusion Chapter 1

This chapter successfully achieved its objective of identifying the dominant LULC classes within the study area using existing and readily available datasets of national landcover. It has provided an in-depth understanding of the potential LULC trends in the Kaaimans catchment, as inferred from the historical SANLC datasets. The dominant classes that were identified include indigenous forest, water, agriculture, bare land, plantation forest, shrubland, and built-up areas. These classes emerged as the most significant due to their prevalence and the substantial changes observed in these categories over the 30-year analysis period. Furthermore, these classes occupy the largest areas, further emphasizing their dominance and importance in the landscape.

This chapter provides a detailed analysis of the Land Use and Land Cover (LULC) changes in the Kaaimans catchment from 1990 to 2020. The northern half of the area, dominated by natural vegetation, reflects effective conservation efforts, while the southern half shows the expansion of human activity, including commercial forestry, urban areas, and agricultural fields. These dominant changes that occur in the southern and northern regions of the study area correlates with the BR's transition and core zones, respectively. This alignment is consistent with the graduated land use gradient that BR management strives to apply. This gradient transitions from the southern region, marked by significant human activity, to the protected areas in the north of the study area, where human activity is more controlled and limited. The study suggests that natural processes, possibly veldfires, drive the major land cover transitions in the northern half, while human activities primarily drive changes in the middle and southern regions. The presence of plantation forests adjacent to indigenous forests potentially exacerbates the risk and severity of fire penetration. These findings underscore the importance of sustainable land-use planning and continuous monitoring of LULC changes.

Chapter 2: Objective 2

2.1 Introduction

The SANLC datasets provide valuable historical data on LULC changes. However, the SANLC datasets only cover LULC up to the year 2020. This limitation is only part of the reason for creating new LULC maps for the Kaaimans catchment area. The SANLC datasets were used to gain a quick understanding of LULC dynamics based the available data. This led to a comprehensive understanding of the study area, which was then applied to a more detailed LULC assessment. To understand the more recent changes in the Kaaimans catchment area, it is necessary to create current LULC classification maps for the years 2019, 2021, and 2023. This chapter, therefore, focuses on the second objective of creating these current LULC classification maps using Sentinel-2 satellite imagery and supervised machine learning processes (Random Forest). This approach will address numerous uncertainties inherent in the SANLC methodology, such as the inconsistencies observed in the SANLC maps. Instead of relying on the small-scale SANLC datasets, this study will provide a large-scale classification that is focused on the specific study area rather than the entirety of South Africa. This will ensure a more precise and relevant representation of the LULC in the catchment area.

The subsequent sections of this chapter will detail the methodology employed for the analysis, followed by a discussion of the results. The dominant LULC classes used for the classification maps were identified in chapter 1 and these included indigenous forest, water, agriculture, bare land, plantation forest, shrubland, and built-up areas.

2.2 Methods

Random forest classifier

The three LULC classification maps were created using the Random Forest (RF) machine learning algorithm in the RStudio programming software (R Core Team, 2021). RF, which was developed by Breiman (Breiman, 2001), is a popular machine learning method used for LULC classification due to the excellent classification results (Rodriguez-Galiano, et al., 2012). The RF algorithm is a supervised machine learning process based on ensemble learning, which is based on using multiple classifiers, in this case, decision trees, to solve problems and improve model performance. The RF algorithm therefore consists of a "forest" of these decision trees, which are trained by the bagging method. In the context of RF, the bagging method applies when predictions are taken from each tree, rather than just one decision tree. The final output is predicted based on the majority votes of predictions. The number of trees for the RF model and the number of splits at each node are the hyperparameters. The number of splits at each node introduces variability in the decision-making process, adding randomness to the model. This leads to a diverse set of decision trees in the random forest, preventing overfitting and increasing accuracy. The "randomForest" function in RStudio used for this study, from the "randomForest" package automatically determines the "split" values and "mtry" values. The "mtry" value is the parameter that specifies the number of randomly sampled variables at each split when building a tree. This function determines the split values automatically based on the greatest decrease in the Gini impurity, a measure that is used to decide the best way to split the data at each node while building the decision trees.

The variable importance, an additional function of RF, measures the correlation between the variables and ranks each one based on its impact on the model's accuracy. The mean decrease in accuracy is calculated automatically by the function and the variable importance plot is visualized using the "varImpPlot" function in RStudio. Variable importance is important in a RF model for identifying key variables with high importance that are key drivers of the classification.

Variable generation

The variables that were used for each model were the ten Sentinel-2 spectral bands in Table 2 and the vegetation indices (VIs) named in Table 12 below. The VIs were calculated using the "RStoolbox" package in RStudio and the coefficient values of the spectral indices were kept as defaults (Leutner, 2023).

Table 12: The vegetation indices selected and used for the models.

Name	Relevance	Formula	Reference
NDVI (normalized difference vegetation index)	NDVI measures photosynthetically active biomass in plants. It is the most highly used index to monitor plant development dynamics	$\frac{NIR - Red}{NIR + Red}$	(Rouse, et al., 1973)
SAVI (soil adjusted vegetation index)	SAVI is used to mitigate the impact of soil brightness to correct NDVI in areas where vegetative cover is low.	$\frac{NIR - Red}{(NIR + R\,ED + 0.5) \times 1.5}$	(Huete, 1998)
EVI (Enhanced Vegetation Index)	EVI complements NDVI and is more sensitive to changes in the areas of high biomass, canopy structure, and plant phenology.	$\frac{2.5 * (\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + 6 * \mathrm{Red} - 7.5 * \mathrm{Blue} + 1)}$	(Huete, et al., 2002)
RVI (Ratio Vegetation Index)	A simple ratio-based vegetation index that indicates the amount of vegetation while reducing atmospheric and topography effects.	$\frac{NIR}{Red}$	
CLRE (Red-edge-band Chlorophyll Index Corrected)	This index is used to estimate chlorophyll content of leaves which helps improve the accuracy of vegetation classification.	$\frac{redEdge3}{redEdge1} - 1$	(Gitelson, et al., 2003)
NDREI1 (Difference Red-edge Index 1 normalized)	This index is sensitive to vegetation chlorophyll content and improves classification accuracy when combined with Vis like NDVI.	$\frac{(redEdge2 - redEdge1)}{(redEdge2 + redEdge1)}$	(Gitelson & Merzlyak, 1994)
NDREI2 (Difference Red-edge Index 2 normalized)	This index is sensitive to vegetation chlorophyll content and plant phenology which improves classification accuracy when combined with Vis like NDVI.	$\frac{(redEdge3 - redEdge1)}{(redEdge3 + redEdge1)}$	(Barnes, et al., 2000)

Accuracy assessment

The Overall Accuracy (OA), the Kappa statistic, Quality Allocation Disagreement Index (QADI), Out-of-Bag Error (OOB), and the user- (UA) producer (PA) accuracy, were computed to perform a comprehensive accuracy assessment of each RF model (for the 2019, 2021, and 2023 models). Evaluating the classification accuracy offers a degree of confidence in the outcomes and the subsequent identification of LULC changes (Cheng, et al., 2021). The OA is the probability of a randomly selected point on the map being classified correctly and it is calculated as follows:

$$OA = \frac{Total\ correctly\ classified\ samples}{Total\ number\ of\ samples} \quad \text{(Eq. 1)}$$

The Kappa statistic, introduced by Cohen (Cohen, 1960), is a popular metric used in assessing the performance of RF models due to its ability to give a more representative measure of model performance than OA. The Kappa statistic shows the difference between the observed

agreement and the agreement expected by chance and a higher Kappa value (>0.75) means that the observed agreement is better than the probability of just random agreement. It is calculated as follows:

$$Kappa = \frac{Observed\ accuracy - By\ chance\ accuracy}{1 - By\ chance\ accuracy} \quad \text{(Eq. 2)}$$

However, there are criticism on the use of the Kappa statistic for assessing LULC classification models due to its sensitivity to the asymmetric distributions of samples across different classes, its inadequate measure of the inter-rater agreement, and that the random accuracy baseline which Kappa uses for comparison to the actual accuracy is often not reliable (Feizizadeh, et al., 2022; Foody, 2020). Therefore, to capture the complexity of classification accuracy, multiple accuracy measures should be used for a comprehensive assessment (Cheng, et al., 2021). To address some of the Kappa statistic's critiques mentioned, the QADI will also be used for accuracy assessment, which solves the Kappa statistic's difficulties (Feizizadeh, et al., 2022). The QADI is calculated from the confusion matrices as follows:

$$QADI = \sqrt{(\frac{A}{N})^2 + (\frac{Q}{N})^2} \quad \text{(Eq. 3)}$$

Where, A, is the allocation of disagreement between classes; N, the total sampled pixels; Q, the quantity of disagreement between classes. Based on the QADI scale, values between 0.00 and 0.12 indicate low disagreement and a high confidence, while values between 0.20 and 1 show high disagreement and a low confidence, and the moderate confidence values are between 0.12 and 0.20 (Feizizadeh, et al., 2022).

OOB error is used to give an unbiased estimate of a model's prediction error, when the model uses the bagging process like RF. The bagging process, used in models like RF, involves creating multiple subsets of the original data, building a separate prediction model for each subset, and then combining the results. The OOB error is a complex formula and it involves giving the mean prediction error of multiple samples and multiple trees.

The performance of each model was further assessed using confusion matrices for both the training and testing stages. The confusion matrix's diagonal elements represent the number of points for which the predicted class is equal to the reference class (correct classifications), while off-diagonal elements are those that are mislabelled by the classifier (misclassifications). From these confusion matrices, the UA and PA were computed. These metrics provide insights into the model's class-specific performance.

Image postprocessing

To reduce classification errors, the classified images will be manually edited and incorrectly classified pixels will be replaced in ArcGIS Pro software (ESRI, 2023). The goal of the image postprocessing is to prevent error accumulation before the classified images will be used for future LULC predictions of the study area (Figure 11).

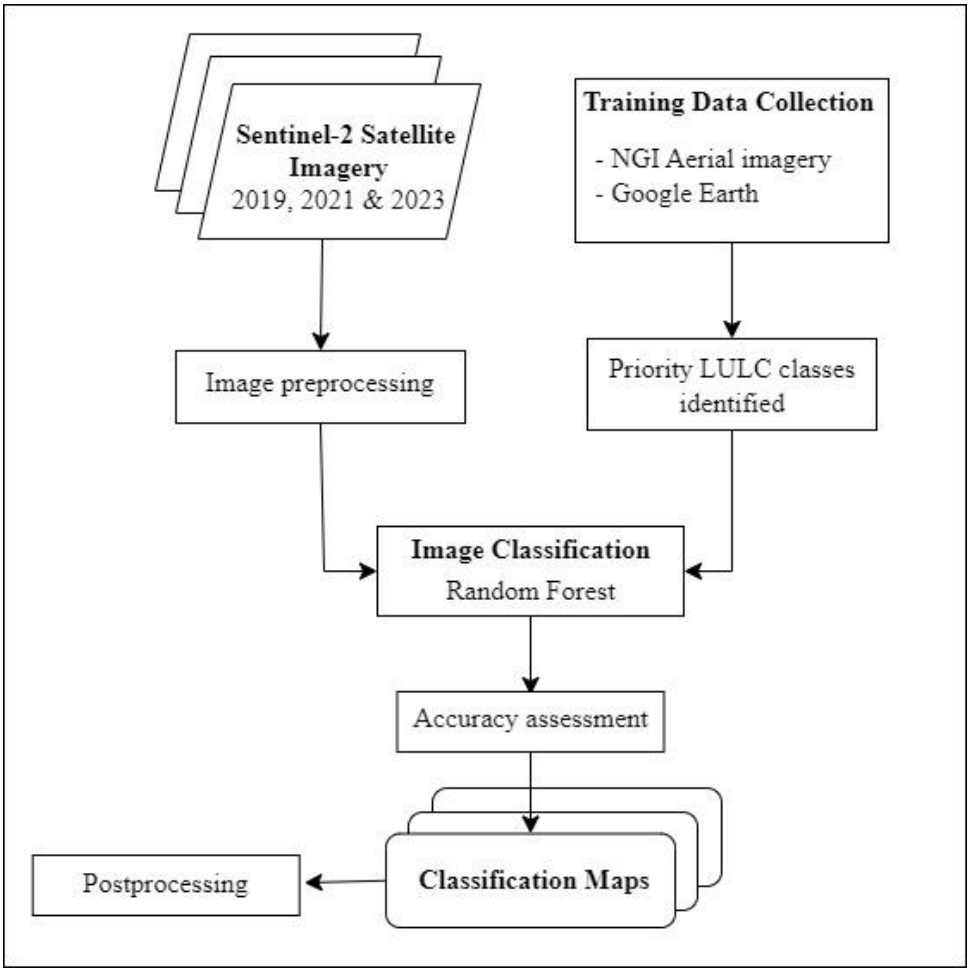

Figure 10: Flowchart of the methods for completing objective 2.

LULC change analysis

The analysis of LULC changes was conducted using the LCM in TerrSet (Clark Labs, 2022). This model facilitated the computation of gains, losses, and net changes for each class during two periods: 2019 to 2021 and 2021 to 2023. Additionally, the percentage area and the area in square kilometres for each class were calculated and tabulated for each year. Utilizing these values, the rate and percentage of change were determined for both periods (2019 – 2021 and 2021 - 2023) using the equations provided below (Leta, et al., 2021):

$$Percantage\ change = \frac{Area\ of\ later\ image - Area\ of\ earlier\ image}{Area\ of\ earlier\ image} \times 100 \quad \text{(Eq. 3)}$$

$$Rate\ of\ change = \frac{(Area\ of\ later\ image - Area\ of\ earlier\ image)}{Number\ of\ years\ between\ each\ image} \quad \text{(Eq. 4)}$$

Furthermore, the contributors to the net change experienced by all the LULC classes within the catchment area were also looked at using the change analysis tool in LCM. Change maps were created of the gains, losses, and persistence of the main LULC that underwent the most changes from 2019 to 2023 based on the above change analysis. Change maps allow us to analyse the spatial dynamics of LULC changes.

2.3 Results

Overall accuracy assessment

All three models performed well, with a high overall classification accuracy (above 85%). These accuracy assessment results indicate that all three models performed well in both training and testing phases, with slight variations in their accuracies and Kappa statistics (Table 13). The QADI values of the training set for each model shows that the 2023 (0.10) and the 2019 (0.12) models have high confidence in their classification accuracy with low disagreement, while the 2021 (0.14) model has a moderate confidence and disagreement (Table 13).

Table 13: Model hyperparameters and accuracy assessment metrics for evaluating each of the RF model's performances.

Hyperparameters			Classification accuracy			
RF Model	**Mtry**	**Split**	**Overall accuracy (%)**	**Out-of-bag acc. (%)**	**Kappa**	**QADI**
2023	500	9	91.31	91.31	0.911	0.10
2021	500	13	86.83	87.65	0.870	0.14
2019	500	9	90.13	90.26	0.907	0.12

Class specific accuracy assessment

The confusion matrices of the 2023 RF model shows that the most misclassifications were the bare land class (class no. 5) with the built-up area class (class no. 4) which is why the bare land class has the lowest user's accuracy percentage. The overall accuracy for the training (91.31%) and testing (92.68%) confusion matrices for this model also suggests that the model performed well in classifying the 2023 satellite image (Table 14; Table 15).

Table 14: The training confusion matrix and statistical measures based on the training dataset for the 2023 RF classification model.

Class no.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**	**User's accuracy (%)**
1	168	5	0	0	0	0	2	175	**96**
2	9	94	0	0	0	0	3	106	**88.68**
3	0	0	33	0	0	0	0	33	**100**
4	0	0	0	173	8	4	4	189	**93.01**
5	1	0	0	19	99	1	2	122	**81.15**
6	0	0	0	3	0	57	1	61	**93.44**
7	1	0	0	0	2	2	80	85	**94.11**
Total	179	99	33	195	109	64	92	771	
Producer's accuracy (%)	**93.85**	**94.95**	**100**	**88.72**	**90.83**	**89.06**	**86.96**		
Overall accuracy (%)	**91.31**								
QADI	**0.10**								

Table 15: The validation confusion matrix and statistical measures based on the testing dataset for the 2023 RF classification.

Class no.	1	2	3	4	5	6	7	Total	User's accuracy (%)
1	74	6	0	0	0	0	0	80	**92.5**
2	1	39	0	1	0	0	0	41	**95.12**
3	0	0	14	0	0	0	0	14	**100**
4	0	0	0	73	7	1	1	82	**89.02**
5	0	0	0	5	45	0	1	51	**88.24**
6	0	0	0	0	0	25	0	25	**100**
7	0	0	0	1	0	0	34	35	**35**
Total	75	45	14	80	52	26	36	328	
Producer's accuracy (%)	**98.67**	**86.67**	**100**	**91.25**	**86.54**	**96.15**	**94.44**		
Overall accuracy (%)	**92.68**								
QADI	**0.07**								

The confusion matrices of the 2021 RF model shown in Table 16 and Table 17 in that the most misclassifications were again the bare land class (class no. 5) getting confused with the built-up area class (class no. 4). There are misclassifications between the plantation (class no. 1) and natural forest (class no. 2) classes. However, the overall accuracy for the training (86.83%) and testing (86.30%) confusion matrices demonstrates that the model performed well (Table 16 and Table 17).

Table 16: The training confusion matrix and statistical measures based on the training dataset for the 2021 RF classification model.

Class no.	1	2	3	4	5	6	7	Total	User's accuracy (%)
1	112	10	0	0	0	1	2	125	**89.60**
2	12	92	0	0	0	0	3	107	**85.98**
3	2	0	31	0	0	0	0	33	**93.94**
4	0	0	1	150	13	9	4	177	**84.75**
5	0	0	0	19	67	1	1	88	**76.14**
6	0	0	0	3	1	75	1	80	**93.75**
7	4	0	0	1	2	1	73	81	**90.12**
Total	130	102	32	173	83	87	84	691	
Producer's accuracy (%)	**86.15**	**90.20**	**96.88**	**86.71**	**80.72**	**86.21**	**86.90**		
Overall accuracy (%)	**86.83**								
QADI	**0.14**								

Table 17: The validation confusion matrix and statistical measures based on the testing dataset for the 2021 RF classification.

Class no.	1	2	3	4	5	6	7	Total	User's accuracy (%)
1	48	4	0	0	0	0	1	53	**90.57**
2	3	41	0	0	0	0	0	44	**93.18**
3	0	0	13	0	0	0	0	13	**100.00**
4	0	0	0	64	11	1	0	76	**84.21**
5	0	0	0	8	26	2	0	36	**72.22**
6	0	0	0	2	0	29	2	33	**87.87**
7	3	0	0	2	0	1	31	37	**83.78**
Total	54	45	13	76	37	33	34	292	
Producer's accuracy (%)	**88.89**	**91.11**	**100.00**	**84.21**	**70.72**	**87.87**	**91.98**		
Overall accuracy (%)	**86.30**								
QADI	**0.14**								

Lastly, the 2019 RF model's confusion matrices shows that the most misclassifications took place between the plantation (class no. 1) and the natural forest (class no. 2) classes (Table 18 and Table 19). Though these misclassifications were once again minimal and overall, the 2019 model performed well with accurately classifying each class.

Table 18: The training confusion matrix and statistical measures based on the training dataset for the 2019 RF classification.

Class no.	1	2	3	4	5	6	7	Total	User's accuracy (%)
1	146	7	0	2	0	0	6	161	**90.68**
2	18	101	0	0	0	0	0	119	**84.47**
3	0	0	28	0	0	0	0	28	**100**
4	3	0	0	145	6	9	0	163	**88.96**
5	0	0	0	3	59	1	0	63	**93.65**
6	0	0	0	4	1	96	1	102	**94.12**
7	8	0	0	0	0	0	55	63	**87.30**
Total	175	108	28	154	66	106	62	699	
Producer's accuracy (%)	**83.43**	**93.52**	**100**	**94.16**	**89.39**	**90.57**	**88.71**		
Overall accuracy (%)	**90.13**								
QADI	**0.12**								

Table 19: The validation confusion matrix and statistical measures based on the testing dataset for the 2019 RF classification.

Class no.	1	2	3	4	5	6	7	Total	User's accuracy (%)
1	64	9	0	0	0	0	2	75	**85.33**
2	1	42	0	0	0	0	0	43	**97.67**
3	0	0	12	0	0	0	0	12	**100**
4	0	0	0	63	1	0	0	64	**98.44**
5	1	0	0	2	26	0	0	29	**89.66**
6	0	0	0	4	0	43	0	47	**91.49**
7	4	0	0	0	0	0	25	29	**86.21**
Total	70	51	12	69	27	43	27	299	
Producer's accuracy (%)	**91.43**	**82.35**	**100**	**91.30**	**96.30**	**100**	**92.59**		
Overall accuracy (%)	**91.97**								
QADI	**0.12**								

Variable importance

For the 2023 RF model the RedEdge1 band (Band 5) contributes the highest importance (Figure 12). Given the high vegetation density of the study area and the distinct nature of the two vegetation classes, the RedEdge1 band is recognized for enhancing model accuracy under such conditions (Ghosh & Saini, 2019; Nasiri, et al., 2022). This pattern is also observed in the other two models, where the RedEdge bands consistently rank among the top variables in terms of importance. In both the 2021 and 2019 models, the red band (Band 4) holds the highest variable importance (Figure 12 and Figure 13). The SWIR2 band (Band 12) consistently ranks among the top variables across all three models, underscoring its significant role in these predictive models. The SWIR2 band is known for providing good contrast between different vegetation types due to its sensitivity to moisture content in vegetation (Wong, et al., 2018).

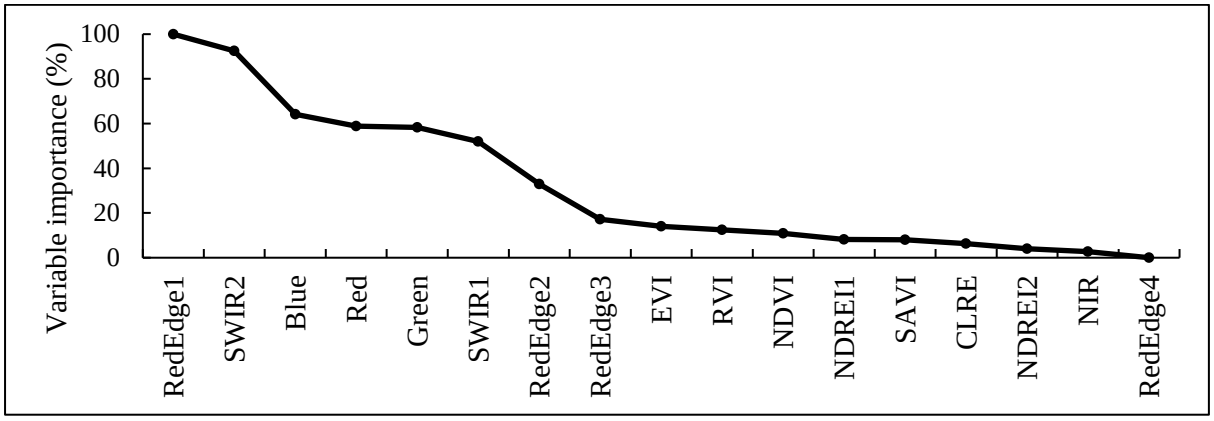

Figure 11: Variable importance (%) of the 2023 RF classification model.

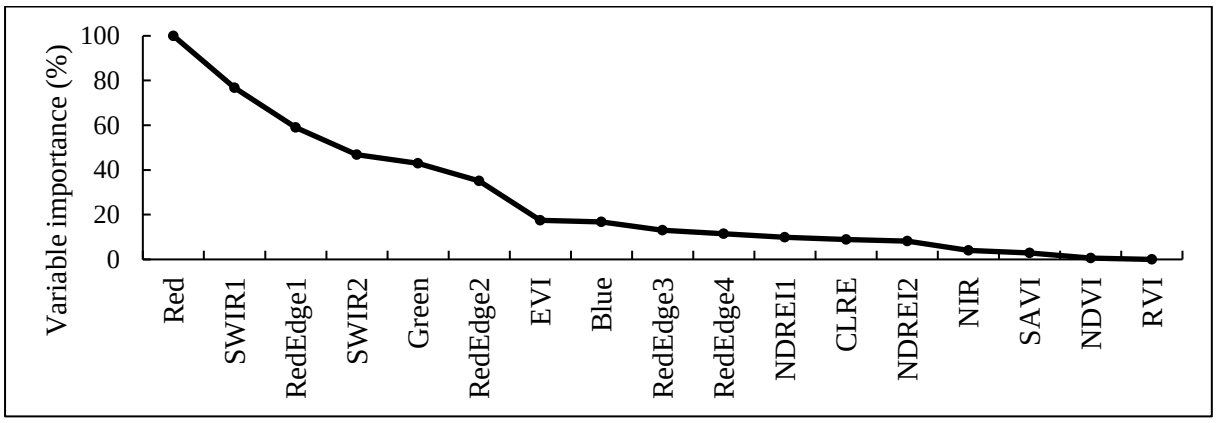

Figure 12: Variable importance (%) of the 2021 RF classification model.

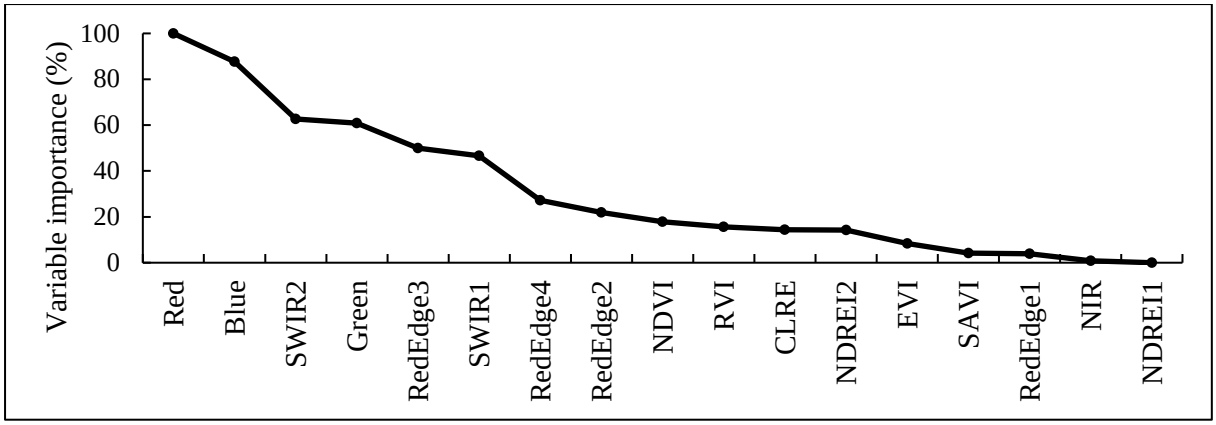

Figure 13: Variable importance (%) of the 2019 RF classification model.

Spatial distribution of land use and land cover

In the Kaaimans catchment area, the spatial distribution of the seven LULC classes exhibits distinct patterns over the years (Figure 15). Natural forest and shrubland areas predominantly occupy the northern and central regions, demonstrating a significant concentration of these classes. Conversely, built-up areas and bare land are primarily located in the southern and western regions. Agricultural land exhibits a dispersed pattern, scattered around the edges of the built-up area. The plantation forests class show a unique spatial distribution, nestled between built-up and agricultural lands on the southern side, and natural forests and shrubland on the northern side.

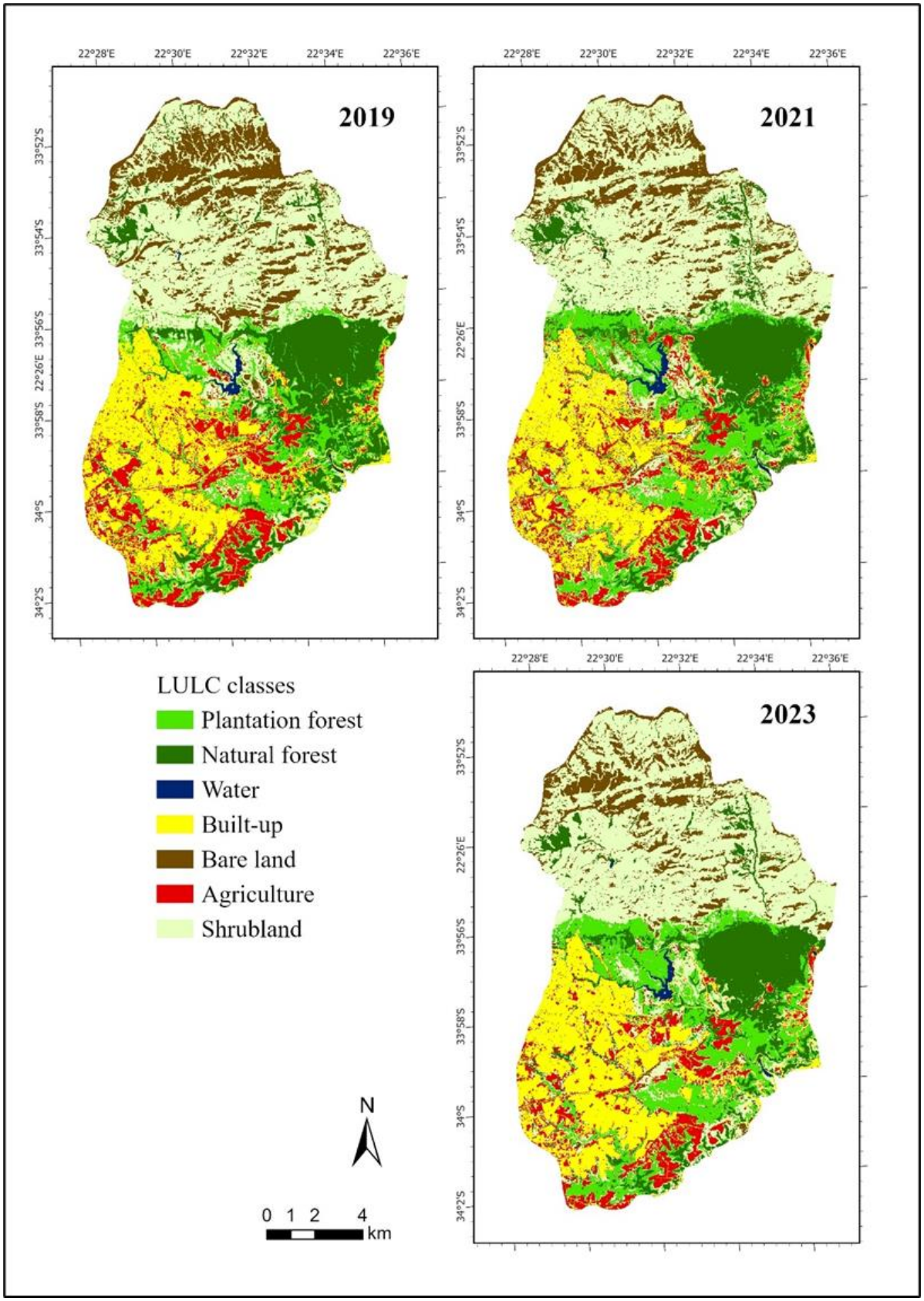

Figure 14: The 2019, 2021, and 2023 Kaaimans catchment LULC classification maps.

Land use and land cover change analysis

During the first period (2019 - 2021), there was an increase in the plantation forest (4.80 km²), and the shrubland class (3.80 km²) (Figure 16). There was a decrease in agriculture (-2.40 km²), bare land (-3.08 km²), built-up (-1.71 km²) and natural forest (-2.24 km²). From 2021 to 2023, a different pattern is observed. Along with plantation forest (0.58 km²), and shrubland (5.25 km²) seeing an increase for this second period, built-up area (1.20 km²) has also seen a net increase. Natural forest (-0.15 km²), agriculture (-2.40 km²), bare land (-3.83 km²) and the least prevalent class, water (-0.23 km²), saw a net decrease from 2021 to 2023.

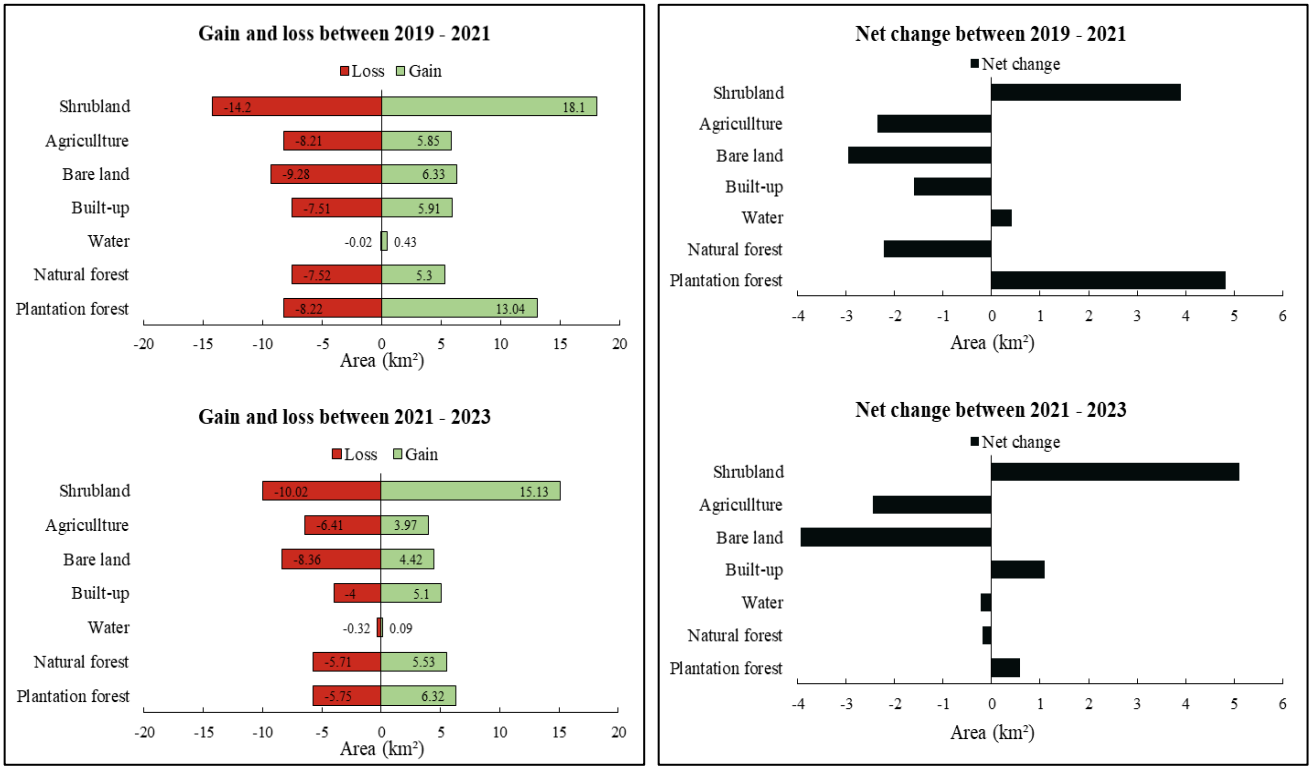

Figure 15: The gains, losses, and net change graphs for the periods 2019 to 2021 (top) and 2021 to 2023 (bottom).

The built-up area is the dominant LULC type of the catchment area which covered 18.29% in 2019, 17.44% in 2021, and 18.02% in 2023 (Table 20). The second largest LULC class in the catchment area, the natural forest, has seen a decline from 2019 (14.75%) to 2023 (13.52%). Plantation forest has seen an increase from 2019 (10.80%) to 2023 (13,52 %). From 2019 to 2021, the largest rate of increase per year is seen with the plantation forest class (2.40 km²/year). Bare land has shown the largest decrease per year for that period with 1.54 km² per year. For the second period, smaller amounts of changes are observed, but the same patterns of change take place.

Table 20: The area and change statistics for the current LULC in the Kaaimans catchment based on the three classified maps. The rate of change was calculated with equation 5 and the percentage change, with equation 4.

	Area						Change					
	2019		2021		2023		2019 - 2021			2021 - 2023		
	Area (km²)	Area (%)	Area (km²)	Area (%)	Area (km²)	Area (%)	km²	%	Rate of change (km²/year)	km²	%	Rate of change (km²/year)
Plantation forest	21.04	10.80	25.84	13.29	26.42	13.57	4.80	22.81	2.40	0.58	2.24	0.29
Natural forest	28.73	14.75	26.49	13.63	26.34	13.52	-2.24	-7.80	-1.12	-0.15	-0.57	-0.07
Water	0.79	0.41	1.20	0.62	0.97	0.50	0.41	51.90	0.21	-0.23	-19.17	-0.12
Built-up	35.61	18.28	33.90	17.44	35.10	18.02	-1.71	-4.80	-0.86	1.20	3.54	0.60
Bare land	25.61	13.15	22.53	11.59	18.70	9.60	-3.08	-12.03	-1.54	-3.83	-17.00	-1.92
Agriculture	19.38	9.95	16.98	8.74	14.58	7.49	-2.40	-12.38	-1.20	-2.40	-14.13	-1.20
Shrubland	63.62	32.66	67.42	34.69	72.67	37.31	3.80	5.97	1.90	5.25	7.79	2.63

Contributors to the net change of LULC classes

The plantation forest class shows a net gain to natural forest (2.78 km²) and shrubland (2.71 km²), indicating that areas of plantation forest are expanding into these classes, causing a loss of shrubland and natural forest from 2019 to 2023 (Figure 17; Table 21). Shrubland replaced mostly bare land (7.39 km²) and agriculture (2.37 km²) areas. Built-up areas expanded by replacing agricultural areas (2.11 km²).

Table 21: The contributors to the net change in km² of the LULC classes from 2019 to 2023.

		Net change from:						
		Plantation forest	**Natural forest**	**Water**	**Built-up**	**Bare land**	**Agriculture**	**Shrubland**
Net change contributed by:	**Plantation forest**	-	-2.78	0.06	0	-0.27	0.3	-2.71
	Natural forest	2.78	-	0	-0.21	-0.09	-0.09	0.01
	Water	-0.06	0	-	-0.03	-0.01	-0.01	-0.07
	Built-up	0	0.21	0.03	-	0.34	-2.11	2.03
	Bare land	0.27	0.09	0.01	-0.34	-	-0.52	7.39
	Agriculture	-0.3	0.09	0.01	2.11	0.52	-	2.37
	Shrubland	2.71	-0.01	0.07	-2.03	-7.39	-2.37	-

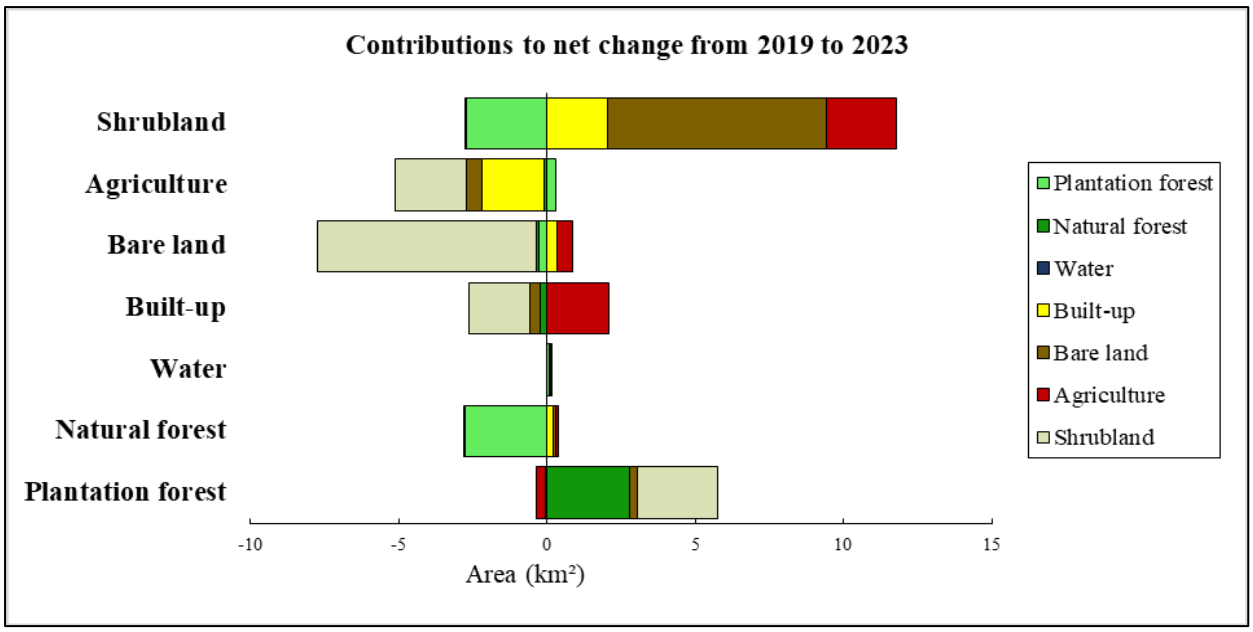

Figure 16: The negative and positive contributions by LULC classes to the net change in each class from 2019 to 2023.

LULC change maps

Between the years 2019 and 2023, the plantation forest class shows areas where it has been changed to other LULC classes (losses), in the centre of the catchment area, with areas of gains stretching across the centre (Figure 18). The natural forest losses are mostly located in the areas of high human activity, at the urban areas to the south western region where George Town is located. The large area of natural forest persistence in the eastern area, which is a PA, displayed a higher degree of loss in the southern part compared to the northern part. Shrubland has seen losses to other LULC classes at the centre of the study area at the southern borders of persistence. Similarly, more losses of shrubland are seen near the urban areas than towards the north of study area at the GRNP. The bare land class has seen large areas of losses in the northern region, and smaller areas of gains in the south, which is the opposite trend showed by the natural forest and shrubland classes. The agriculture class experienced losses within the urban area, maintaining persistence in the eastern regions outside urban areas. Built-up areas expanded near their persistent regions, with minor losses on the eastern borders.

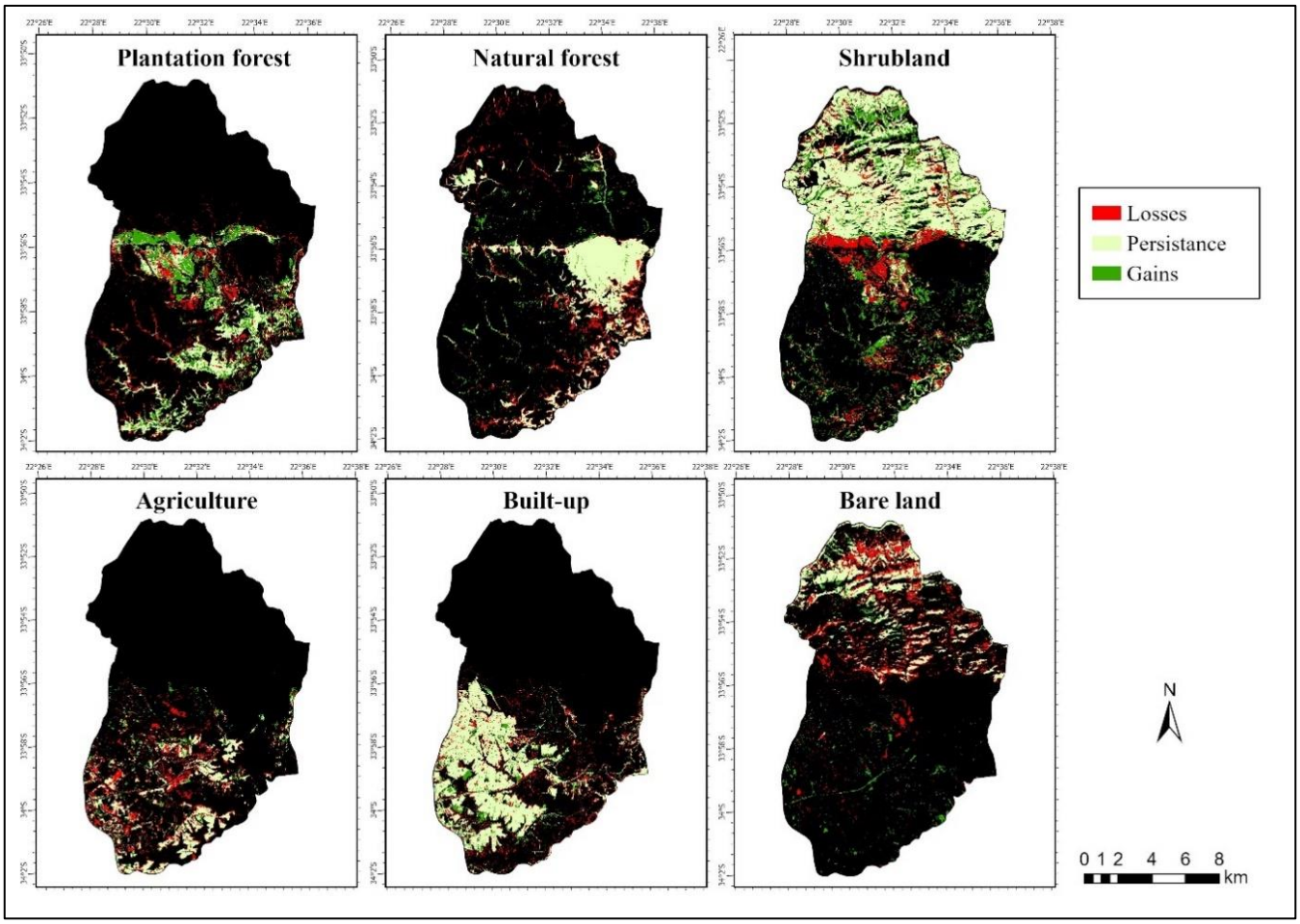

Figure 17: The maps of the six main LULC classes in the catchment area, showing each class's loss, gain, and persistence from 2019 to 2023 (the water class underwent no change, therefore was excluded from this map)

2.4 Discussion Chapter 2

The results of this Chapter indicate that all three RF models performed well in classifying LULC classes, with overall classification accuracies exceeding 85% (Table 13). The QADI values for the training set of each model suggest that the 2023 and 2019 models have high confidence in their classification accuracy, while the 2021 model exhibits moderate confidence. The accuracy of the three classified maps suggests that the resultant classifications are sufficient to be utilized for modelling future LULC trends in chapter 3.

The confusion matrix results of the 2021 and 2023 RF models reveal that the most common misclassifications occurred between the bare land class and the built-up area (Table 16 and Table 14). This could be attributed to the bare land and concrete infrastructure in built-up areas having similar reflectance and spectral characteristics which is shown in other studies to cause classification errors (Kuc & Chormanski, 2019; Osgouei, et al., 2019). Similarly, there was some misclassifications between the plantation and natural forest classes. This could be attributed to the spectral similarities between these two classes in terms of their density and spectral reflectance. Additionally, the encroachment of natural vegetation into older plantation lands might have resulted in incorrect training points between the two classes, and vice versa (Kraaij, et al., 2018). According to the variable importance results for each model (Figure 12, 13 and 14), the RedEdge and SWIR Sentinel-2A bands consistently demonstrated a significant contribution to all three models. Due to their sensitivity to vegetation features and moisture content, these two bands are known to improve model accuracy in contexts where there is a high vegetation density present in the region under study and the distinct characteristics of the three vegetation classes (Lui, et al., 2021).

In terms of the spatial characteristics of the Kaaimans catchment area, plantation forests interestingly show an in-between spatial distribution, located between built-up and agricultural lands on the northern side, and natural forests and shrubland on the southern side (Figure 15). This unique positioning indicates a transitional landscape of human utilisation. During the first period (2019 - 2021), both the plantation forest, and the shrubland class experienced the largest increase, with a percentage change of 22.81% and 5.91%, respectively (Table 20). These large increases, could be due to their recovery after the destructive 2018 George wildfires, which burned plantations, infrastructure, and natural vegetation (Quiroz, et al., 2023). Specific to the study area, the fires burned the entire northern region in the Outeniqua mountains, and some of the plantation forests in the centre of the region around the Garden Route Dam. This is where the most gains can be seen from the plantation forest and shrubland classes in Figure 18. Built-

up (-4.80%), natural forest (-7.80%), bare land (-12.03%), and agriculture (-12.38%) showed to all decrease during this period (Table 20). From 2021 to 2023, a different pattern is observed. The plantation forests increased slower with a smaller (2.24%) change, and shrublands increased more with a higher (7.79%) change than the first period. Interestingly, built-up areas, which had previously shown a decrease between 2019 and 2021, expanded by 3.54%. This could be indicative of urban expansion and development activities in the region (Tizora, 2016). The substantial 52% growth in George Town's population from the 2011 to the 2022 Census lends further support to the observed expansion of urban areas (Statistics South Africa, 2023). This suggests that the urban development is likely a strategic response to accommodate the significant increase in population. In contrast, agriculture showed an even further decrease in area with a 14.13% decrease from 2021 to 2023. Further studies could provide more in-depth insights into the ecological and socio-economic factors influencing this pattern.

From 2019 to 2023, current LULC maps reveal a trend of plantation forests gradually encroaching upon natural forest areas (Figure 15 and 18). This is particularly noticeable in regions of high human activity in the southern and central parts. However, this trend is less pronounced in the more inaccessible northern regions of the Outeniqua mountains. This trend is supported by the fact that the largest contributor to plantation forest expansion is the encroachment into natural forests, which accounted for an area of 2.78 km² during the same period (Table 21; Figure 17). The natural forest class has decreased in area from 28.73 km² in 2019 to 26.34 km² in 2023 (Table 20).

This study utilized the RF algorithm, remote sensing indices, and Sentinel-2A imagery to generate LULC classification data, without the use of in-situ validation data. A recent study demonstrated the potential of using an RF classifier with Sentinel-2A satellite imagery for LULC maps, thereby reducing the need for ground truth data (Ali, et al., 2022). The methodology proved to be cost-effective and efficient, eliminating the need for ground truth data in LULC mapping. An in-depth analysis of historical LULC data was conducted to gain a comprehensive understanding of the study area's LULC changes, as outlined in chapter 1. This, coupled with the analysis of the spectral signatures of the collected training samples for each model, and using freely available auxiliary remotely sensed imagery such as the NGI aerial imagery and Google Earth Pro imagery to generate the training samples, enabled the creation of accurate, up-to-date LULC classification maps for the Kaaimans catchment area. This approach demonstrates a time and cost-effective way to ensure accurate results without the use of in-situ data.

This chapter does have certain limitations that need to be acknowledged. Using satellite data of a higher spatial resolution will facilitate the creation of more detailed LULC maps, thereby enhancing our understanding of the study area (Basheer, et al., 2022; Feng & Li, 2020). This approach is particularly beneficial in differentiating more effectively between different types of plant species and in rectifying misclassifications, such as those between built-up areas and bare land, and between natural forests and plantation forests.

2.5 Conclusion Chapter 2

In conclusion, the results of this chapter demonstrated the effectiveness of applying RF models and Sentinel-2A satellite imagery in LULC classification in a large scale, heterogenous catchment area with densely vegetated, urban, and agricultural areas. The RF models of 2019, 2021 and 2023 performed well (with overall accuracies exceeding 85%), showing a spatial distribution closely aligned with the 2020 SANLC map. This was based on the understanding of the LULC dynamics in the area gathered from the analysis of the 30-year period of the SANLC datasets in chapter 1. During the first period (2019 - 2021), plantation forest and shrubland classes experienced the largest increase in the middle to northern region of the study area, likely due to recovery after the 2018 George wildfires. However, from 2021 to 2023, with an expansion of built-up areas in the southern region, possibly due to urban development activities. The results showing plantation forest species encroaching into natural vegetation areas, such as shrublands and natural forests, is a trend that has been corroborated by both the SANLC datasets and the LULC classification map analysis. Invasive species and non-native plantation species spreading into natural vegetation is a significant ecological concern. Once they leave plantations, they can invade natural habitats, altering local vegetation and outcompeting native species. This spread can impact global biodiversity and human welfare, affecting crucial resources like water. This highlights the importance of managing the spread of alien species from plantations within the catchment area.

The results provide valuable insights into the spatial and temporal dynamics of LULC in the Kaaimans catchment area, which can inform land management and conservation strategies. The chapter has successfully created accurate up-to-date LULC maps of the study area, which will be used for LULCC modelling in chapter 3.

Chapter 3: Objective 3

3.1 Introduction

This Chapter focuses on objective three, which is utilizing the current classification maps, produced in Chapter two using satellite imagery from the years 2019, 2021, and 2023, to predict future LULC for the Kaaimans catchment up to the year 2030. The aim is to provide a comprehensive understanding of potential future LULC changes, which is crucial for effective planning and sustainable development. This chapter will delve into the methodologies used, the LULCC analysis performed, and the insights gained from this objective.

3.2 Methods

Land Change Modeller (LCM)

To predict future LULC for the Kaaimans catchment area to the year 2030, a Decision Forest (DF) integrated Markov Chain (MC) model and Multi-layer Perceptron (MLP) integrated MC were used within the TerrSet 2020 software (Clark Labs, 2022). This study presents a comparative analysis between the less frequently used DF-MC model and the more commonly used MLP-MC model. The modelling of both models was done in the Land Change Modeller (LCM) of TerrSet v.19.0.8. TerrSet is a geospatial monitoring and modelling software that was used to do the future LULC predictions and change analysis based on the classified satellite data from chapter 2. The LCM is based on three main steps: change analysis between two LULC maps (in this study's context, 2019 and 2023), transition potential modelling by means of driver variables and transition potential maps, and a "change prediction" analysis that uses MC matrices to simulate future LULCC based on the probabilities of change (Figure 19). Many of LULC focused studies have used the LCM for analysing LULC changes and to make future projections based on the MC model with high accuracy results (Leta, et al., 2021; Shafie, et al., 2023; Shen, et al., 2020; Shrestha, et al., 2023). The study limited its LULC predictions to a 7-year horizon due to the increasing uncertainty and potential for errors in longer-term forecasts (Arfasa, et al., 2023). The two maps used for the prediction had a short 4-year interval, which may not capture sufficient variability in LULC changes for accurate long-term predictions. Furthermore, any errors present could compound over time, leading to larger inaccuracies in the predictions.

Driver variables

In the LCM, driver variables are critical components that help in predicting future land use and land cover changes. These spatial variables are factors that influence or drive the changes in LULC. For this study, the driver variables selected were elevation, slope, aspect, distance from rivers (including water bodies), distance from plantations, distance from built-up, distance from main roads (Figure 7). These driver variables were used to create the transition potential maps that are used by both the predictive models to make future LULC projections.

Transition potential modelling

The transition potential modelling precedes the change prediction step, and it involves modelling the potential of the transitions between the earlier and later LULC maps. First, the transitions between the two selected classified maps were filtered out to only consider LULC transitions larger than 4 km² between the two LULC maps. This was done to eliminate any noise that might be created by misclassifications among smaller LULC transitions and to concentrate solely on the major LULC changes in the study area. These filtered LULC transitions included only the substantial transitions between LULC classes: plantation forest to shrubland, natural forest to plantation forest, bare land to shrubland, agriculture to built-up, shrubland to plantation forest. These transitions were grouped into a single sub-model.

Predictive model 1: DF-MC

One of the two predictive models used in this study was a DF-MC model in the LCM tool of TerrSet v19.0.8. The DF algorithm, which is an ensemble of decision trees and a modification of the original RF algorithm (Breiman, 2001), is rarely used for LULCC modelling because it was only recently released by Clark Labs in the 2020 TerrSet version for the LCM, and since then it has only been used in a handful of studies (e.g., Azari, et al., 2022; Safabakhshpachehkenari & Tonooka, 2023). However, it has had high accuracy in these recent studies, and has shown to have advantages over other predictive models (Al-Sharif & Pradhan, 2014; Samardžić-Petrović, et al., 2015). The DF works by constructing a decision 'forest,' which is a collection of decision trees and each tree is formed using a bootstrapped subset of data. For every node in the tree, a random subset of the data is chosen as predictors, and the optimal split among the predictors is selected based on a criterion known as the Gini impurity. This procedure continues until the tree is fully grown, and this process is repeated across multiple trees to form the 'forest.' For LULC modelling, each decision tree in the forest symbolizes a potential sequence of LULC changes, which are based on different combinations of the input data. The randomness is introduced by the bagging and random feature selection during the modelling process, which guarantees that each tree captures distinct aspects of the LULC transition process (Safabakhshpachehkenari & Tonooka, 2023). The prediction of a cell is first predicted by each tree in the forest, which results in multiple predictions for that cell. The mode or mean of those multiple predictions will determine the final prediction. This process is what the DF model uses to predict future LULC transitions based on various driver factors. The parameters for the DF model were set to evaluate 2 variables at each split. This

was based on the split value being equivalent to the floor function of the square root of the total number of independent variables. Additionally, the number of trees were left at the default value of 100 trees.

Predictive model 2: MLP-MC

The second LULCC prediction model that was used in this study is the most used artificial neural network (ANN), the hybrid MLP-MC neural network method. The MLP-MC model is a very widely used LULCC model within TerrSet software and has shown high accuracies with simulating future LULC (Aguejdad, et al., 2017; Kumar & Agrawal, 2023; Saha, et al., 2022; Wang & Maduako, 2018). An MLP consists of three layers of nodes: an input layer that first receives the data from the earlier (for example 2019 LULC map) and later (2023) LULC map, a hidden layer where the data processing and the automatic assigning of weights according to the LULCC patterns over time, and the output layer which produces the results of the given inputs. The MLP learns by using forward propagation, which is where the inputs are processes from the input layer forward to the output layer. The computed error (difference between prediction and actual changes) from the output is then back propagated through the system and is sent in reverse through the network to locate the error connections. The iterations for MLP back-propagation were set to 10 000, which is the default. During the backpropagation, the weights within the hidden layers are updated (with gradient descent) to minimize the output error. This whole process is repeated multiple times to ensure that the output error is as small as possible. The MLP training parameters were set to use automatic training and a dynamic learning rate.

Markov Chain

After the transition potential modelling was done, the “Change Prediction” analysis was the next step and it is where the LULCC prediction process takes place. After the prediction date is specified, the quantity of change in each transition was modelled through the MC analysis. The outputs of this process are a LULC classification map of the predicted year (hard prediction map), and a continuous map that shows the probability of change across the area (soft prediction map). The MC is a stochastic mathematical model that predicts a future state based on the probability of transition and the current state, not considering the past states. This model makes use of the two LULC images from two time periods and the transitional potential maps, to create a transition probability matrix, transitional areas matrix, and conditional probability images. The MC process can be represented by the following equations (Moodley, et al., 2023):

$$S(t, t+1) = P_{ij} * S(t) \qquad \text{(Eq. 6)}$$

In this equation, *S(t)* represents the state of the system at a given time *t*, while *S(t+1)* signifies the state of the system at the next time step, *t+1*. The term *Pij* is the transition matrix that represents the probabilities of transitioning from one state to another.

The transition matrix, *Pij*, is defined as follows:

$$P_{ij} = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,n} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n,1} & P_{n,2} & \cdots & P_{n,n} \end{bmatrix}, (0 \leq P_{ij} \leq 1 \; and \; \sum_{j=1}^{n} P_{ij} = 1, i, j = 1, 2, \ldots, n) \qquad \text{(Eq. 7)}$$

Here, each element *Pi,j* in the matrix represents the probability of transitioning from the current state *i* to another state *j*. The probabilities are such that they are always between 0 and 1 inclusive. The term *Pn* refers to the probability state at any given time.

Model validation

An important step in developing any predictive change model is validation (Aburas, et al., 2019). Validation is a procedure used to assess the quality of the predicted LULC map against a reference map. The Validate module in TerrSet V19.0.8 was used to validate the predicted 2023 maps from the MC and CA-Markov models using the supervised classified LULC map of 2023 as a reference. This is a comparative analysis based on the Kappa Index of Agreement (KIA). KIA is a statistical measure that states the agreement accuracy between the predicted LULC map and the reference, adjusted for the chance of agreement (Pontius, 2002). A KIA value of 1 means that the two maps are in perfect agreement, while a value of 0 means that they are in agreement by chance alone. Negative values close to -1 indicate that there is a disagreement between the two maps. This module calculates the KlocationStrata, Kstandard, Kno, and Klocation which provides a complete analysis of the overall accuracy in terms of location and quantity. To assess the strength of the agreement, the following statistics were also considered: AgreementChance, AgreementQuantity, AgreementGridcell, DisagreeGridcell, and DisagreeQuantity. Additionally, the areas (in km²) of the projected 2023 maps of the two models were compared to the reference 2023 map to quantify the spatial accuracy of the predicted LULC map.

LULC change analysis

The analysis of LULC change between 2023 and the predicted 2030 maps was conducted using the Change Analysis section in TerrSet software. This function identifies and quantifies the changes between the initial and final LULC states. The output was a change map that visually represents areas of change that provides a detailed quantification of the changes, including gains, losses, and net changes for each LULC class. This method allows for a comprehensive understanding of the spatial patterns and magnitude of LULC changes over the specified period. The contributors to each LULC class's change between the years 2023 and 2030 was also analysed.

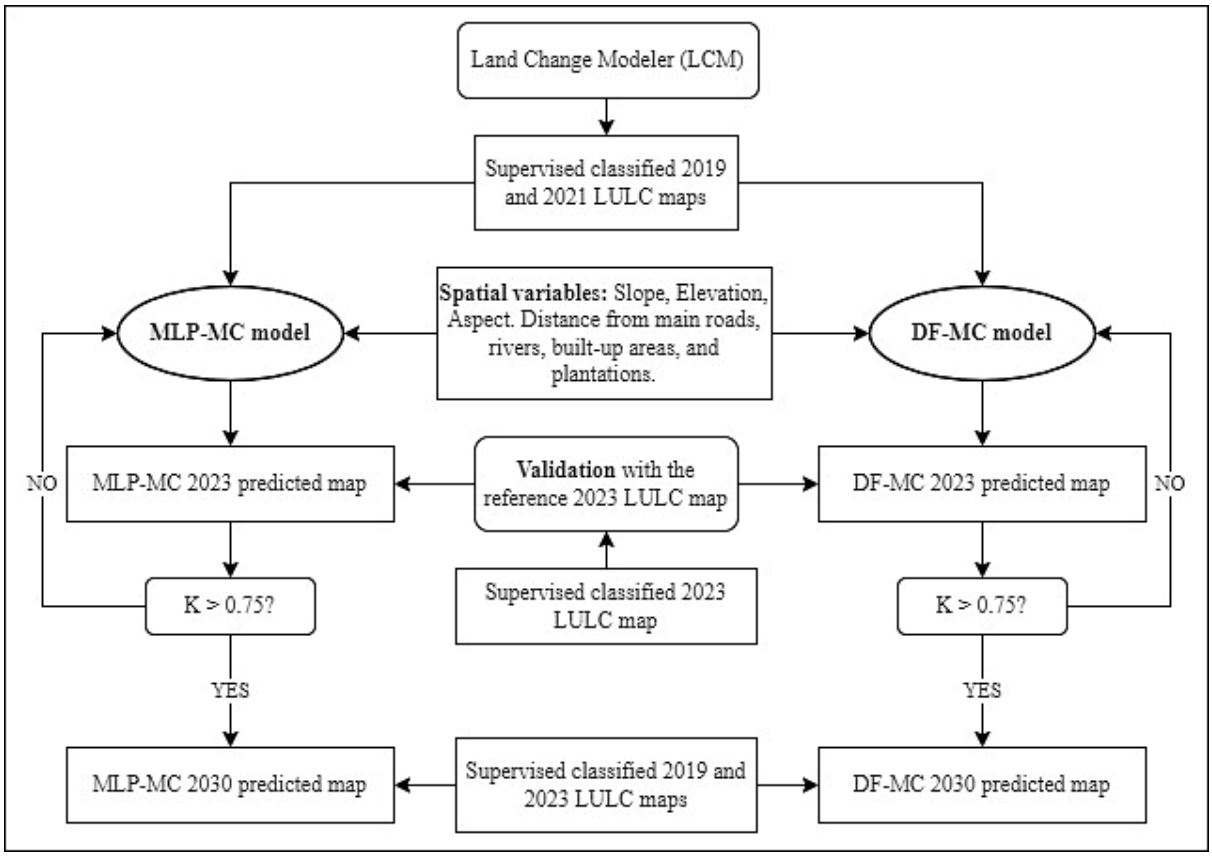

Figure 18: Methodology flowchart of objective 3.

3.3 Results

Model validation

The model performance metrics show small differences between the models, and both models performed similarly overall with all Kappa-index values larger than 0.75 (Table 22). The DF-MC model generally has slightly higher Kappa statistics, indicating slightly better performance in accuracy and location. Both models show high proportions of the total quantity that is correctly predicted (AgreementQuantity), and of the individual grid cells that are correctly predicted by the models (AgreementGridcell). This indicates that both models have a strong agreement between their predicted 2023 map and the reference 2023 map. The DF-MC model has 5.6% of individual grid cells that are incorrectly predicted, while the MLP-MC model has a slightly higher disagree grid cell of 6.3%. Both models have the same proportion of the total incorrectly predicted quantity of 1.5%, as well as the same agreement by chance value of 12.5%.

Table 22: The validation result analysis (agreement and disagreement components) and the k-index values of each model predicted 2023 map compared to the reference.

	DF-MC (%)	**MLP-MC (%)**
AgreementChance	12.5	12.5
AgreementQuantity	40.5	40.5
AgreementGridcell	40.0	39.3
DisagreeGridcell	5.6	6.3
DisagreeQuantity	1.5	1.5
	DF-MC	**MLP-MC**
Kno	0.919	0.911
Klocation	0.877	0.862
KlocationStrata	0.877	0.862
Kstandard	0.850	0.835

Due to the models having the same MC transition matrix, and the same 2019 and 2021 LULC maps for predicting the 2023 LULC map, they will predict the same area (km²) for each LULC class (Figure 20). The comparison of the predicted area of each class with the reference map's areas shows that area predictions of the classes are close to the reference areas, hence, the prediction models were found to be capable of simulating LULC for the year 2030.

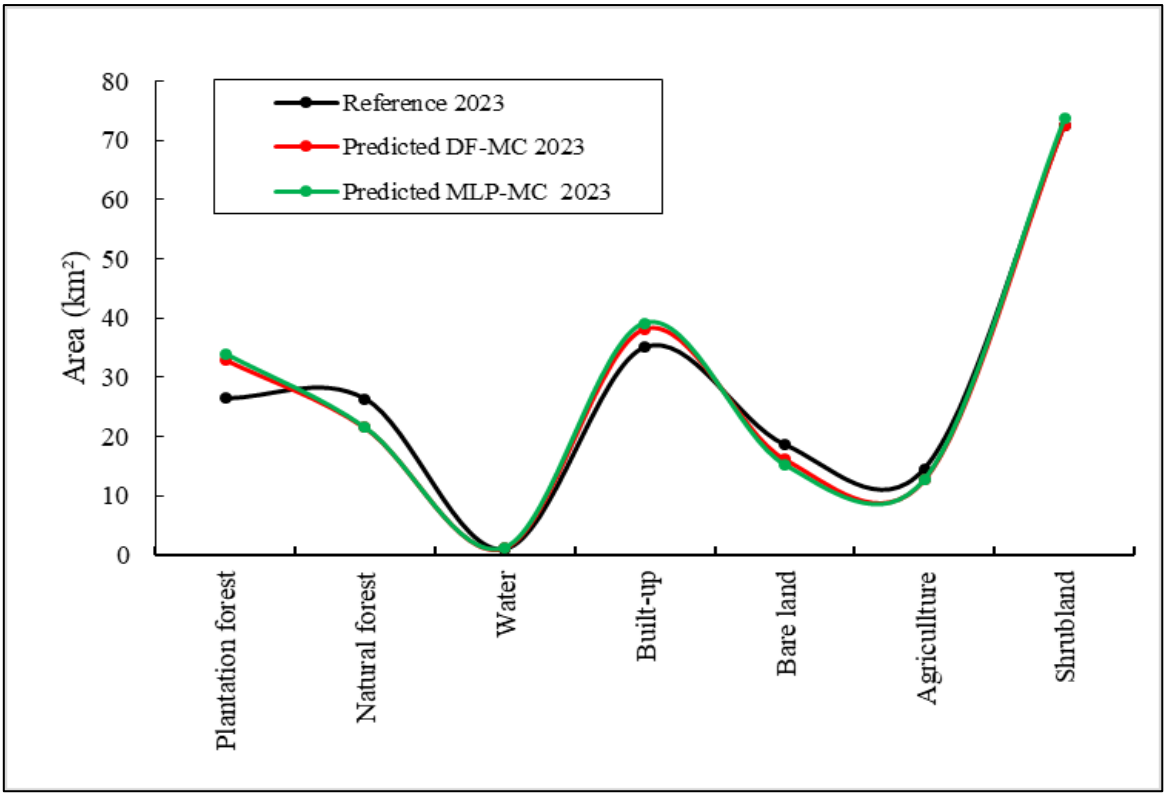

Figure 19: The predicted area of each LULC class by the MLP-MC and DF-MC models of their predicted 2023 maps compared to the reference 2023 area per class.

Transition potential modelling for 2030

The provided MC transition matrix is a crucial component in predicting the LULCC for the year 2030 (Table 23). The matrix indicates a cell currently classified as plantation forest has a 42.82% chance of remaining the same, while it has a 21.49% chance of transitioning to natural forest, which is the highest transition probability for this class. Similarly, a cell classified as natural forest has a 63.29% chance of remaining the same, with the highest transition probability being to plantation forest at 27.90%. In the case of the water class, in the absence of any large infrastructure developments, the probability of remaining the same is extremely high at 99.87%, indicating a very low likelihood of transition to any other class. Built-up areas, on the other hand, have a 58.16% chance of remaining the same, with the highest transition probability being to agriculture at 21.95%. Bare land has a 64.85% chance of remaining the same, with the highest transition probability being to shrubland at 22.47%. Agriculture has a 62.18% chance of remaining the same, with the highest transition probability being to built-up areas at 20.11%. Lastly, shrubland has a 33.91% chance of remaining the same, with the highest transition probability being to bare land at 32.67%. These probabilities, along with the current state of each cell, are used by the models to predict the state of each cell in 2030.

Table 23: The MC transition probability matrix (%) of LULCC for predicting to the year 2030, used by both models.

	Probability of changing to:						
Given:	**Plantation forest**	**Natural forest**	**Water**	**Built-up**	**Bare land**	**Agriculture**	**Shrubland**
Plantation forest	42.82	21.49	9.78	3.22	2.79	10.15	9.75
Natural forest	27.9	63.29	1.41	0.46	0.82	1.58	4.54
Water	0.06	0	99.87	0.02	0.04	0	0.01
Built-up	2.46	1.6	5.36	58.16	6.4	21.95	4.07
Bare land	4.95	1.82	2.67	2.18	64.85	1.06	22.47
Agriculture	3.89	1.28	2.32	20.11	4.86	62.18	5.36
Shrubland	10.29	9.09	9.29	1.37	32.67	3.38	33.91

The transitional potential maps are equal to the number of major change transitions between LULC classes (Figure 21). The change potential for each transition for the two models differ, thus, the spatial distribution of LULC for each model will be different. However, the trends of change potential are relatively the same, with similar areas having higher change potential between the two models. For example, the transition from natural forest to plantation forest for both models show high change potential at areas close to the urban area in George Town.

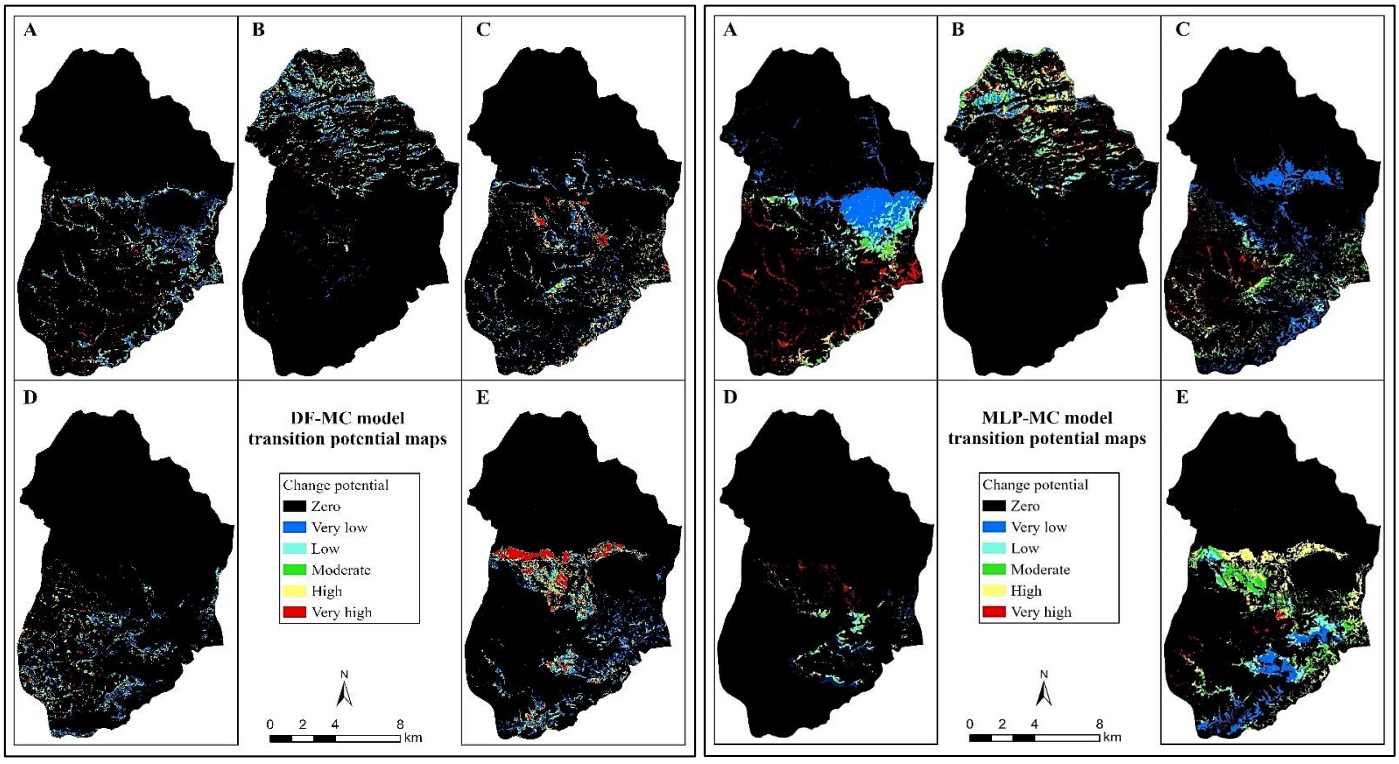

Figure 20: Transition probability maps of major transition categories; A is natural forest to plantation forest, B is bare land to shrubland, C is shrubland to plantation forest, D is built-up to agriculture, E is plantation forest to shrubland.

Both models show a high change potential at the foot slopes of the Outeniqua mountains, in the centre of the study area, for the plantation forest to shrubland transition. However, for the built-up to agriculture transition, the DF-MC model show high change potentials around George Town, where the MLP-MC model show high change potential at the centre of the study area, around the Garden Route Dam. Overall, the very high change potentials of the MLP-MC model are located around the built-up area of George Town where most human activity take place, and the same trend is seen with the DF-MC models, but at a lesser extent.

2030 LULC prediction

From 2019 to 2023, the Kaaimans catchment area experienced various LULC changes (Figure 22 and Chapter 2). Plantation forest and built-up areas saw an overall increase, while natural forest, bare land, and agriculture experienced a decrease. Water areas remained relatively stable. Shrubland saw a decrease initially but increased again by 2023. The DF-MC and MLP-MC models predict further LULC area changes by 2030. Both models forecast an increase in plantation forest and built-up areas, a decrease in natural forest, bare land, and agriculture, and stability in water areas and shrubland.

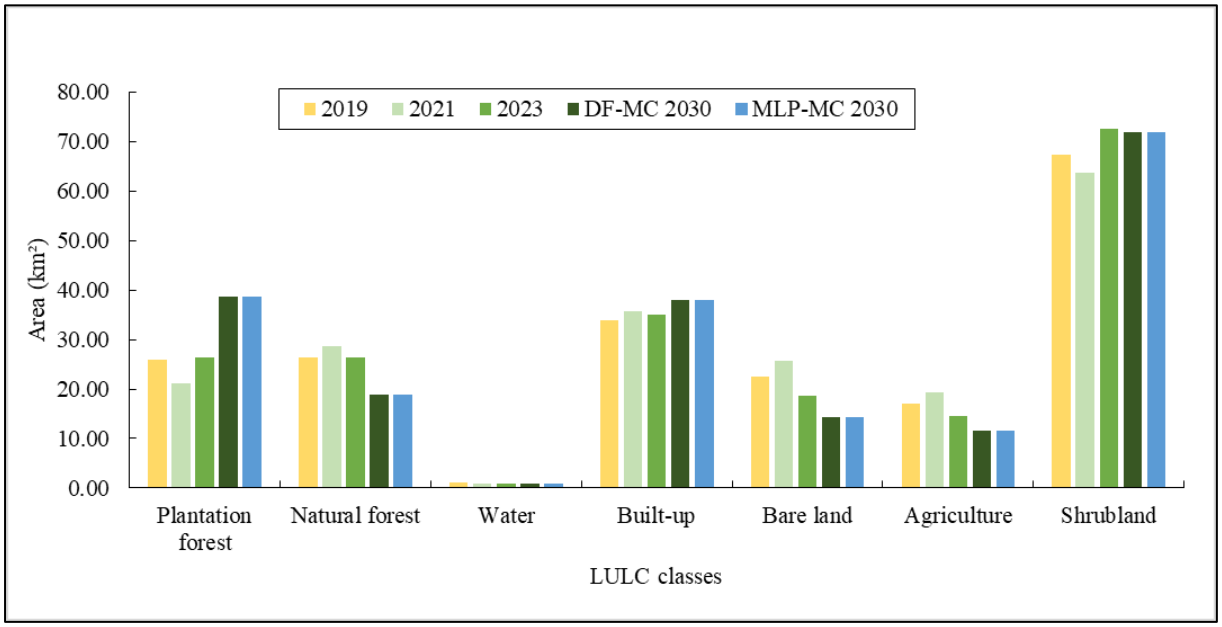

Figure 21: The LULC change in area from the three supervised classification maps 2019, 2021, and 2023 from Chapter 2; and the two predictive models for the year 2030 from Chapter 3.

The area of plantation forest is predicted to increase from 26.42 km² in 2023 to 38.64 km² in 2030 according to both models. The water and shrubland classes are predicted to remain relatively stable from 2023 to 2030. The natural forest area is expected to decrease from 26.34 km² in 2023 to 18.97 km² in 2030 as per both models. The built-up areas saw a small increase

from 35.10 km² in 2023 to 37.92 km² in 2030. Bare land and agriculture both underwent a big decrease in area, with 18.70 km² to 14.41 km² and 14.58 km² to 11.62 km², respectively.

Table 24: LULC area (km²) coverage of classified and predicted images.

	2019	**2021**	**2023**	**DF-MC 2030**	**MLP-MC 2030**
Plantation forest	25.84	21.04	26.42	38.64	38.64
Natural forest	26.49	28.73	26.34	18.97	18.97
Water	1.20	0.79	0.97	0.97	0.97
Built-up	33.90	35.61	35.10	37.92	37.92
Bare land	22.53	25.61	18.70	14.41	14.41
Agriculture	16.98	19.38	14.58	11.62	11.62
Shrubland	67.42	63.62	72.67	71.83	71.83

The DF-MC and MLP-MC models forecast relatively similar spatial distributions of LULC in the Kaaimans catchment area for the year 2030 (Figure 23 and 24). The encroachment of plantation forest into the surrounding areas, especially along the south-eastern region of the study area, is predicted by both models. Adjacent to the plantation forest areas, the natural forest is shown to be receding and replaced by plantation forest. The water class serves as stable landmarks in the predicted maps. A larger expansion of built-up areas into agricultural areas is shown by the MLP-MC model compared to the DF-MC which shows a smaller transition. The northern half of the study area, that borders with the plantation and natural forests, has remained largely undisturbed by other LULC classes in both maps, with the exception of shrubland replacing bare land. The change potential maps of both models show the areas where each model predicted to be the most potential for change between classes. When comparing the two models, it is evident that the DF-MC model's change potential map indicates fewer areas with very high (red) change potentials. This suggests that, according to the DF-MC model, fewer regions in the catchment are likely to experience significant LULC transitions. On the other hand, the MLP-MC model's change potential map shows more areas with very high change potentials, indicating a broader extent of predicted LULC transitions across the catchment. For both models, most areas with moderate to extremely high change potentials are concentrated in the central region of the study area, extending towards the southeast where George Town is located. This region, where most LULC transitions occur, is situated between the urban area in the south-west and the northern half of the study area which is part of the GRNP.

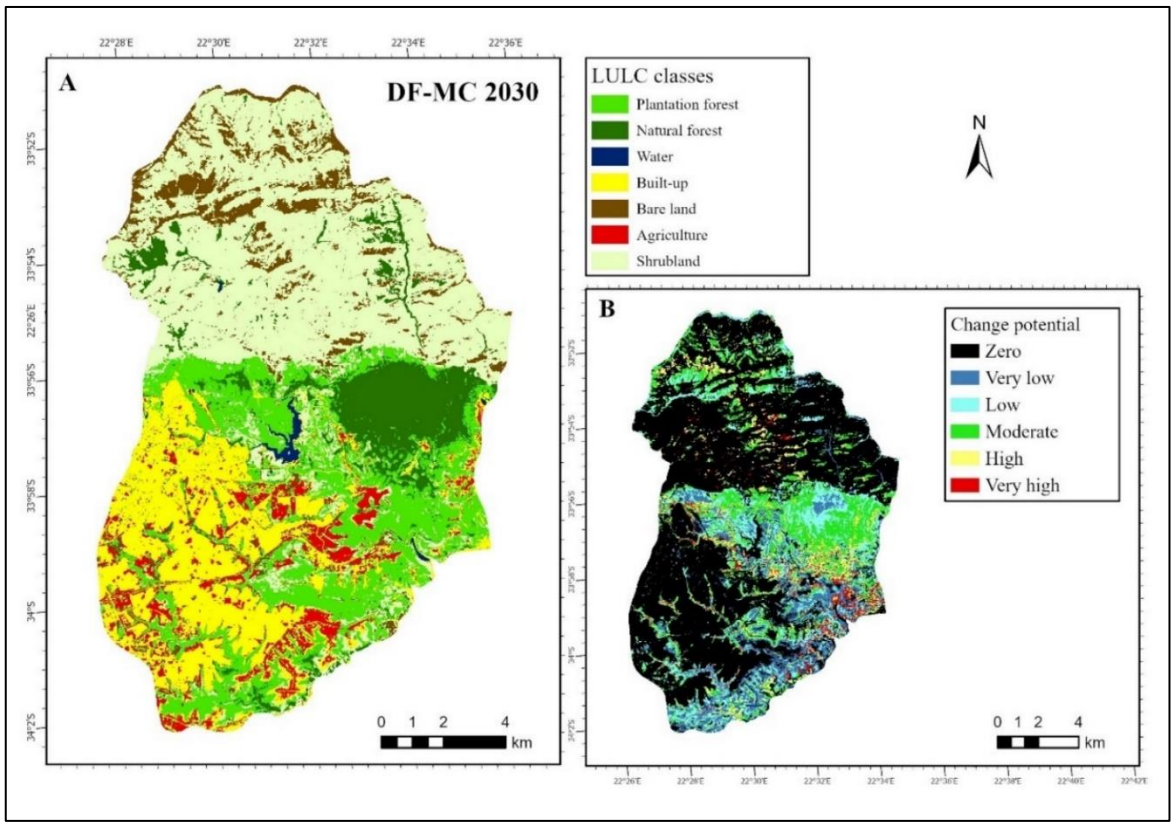

Figure 23: The DF-MC model prediction of LULC spatial distribution in the Kaaimans catchment area for the year 2030. A is the hard prediction map, B is the soft prediction map that shows the change potential.

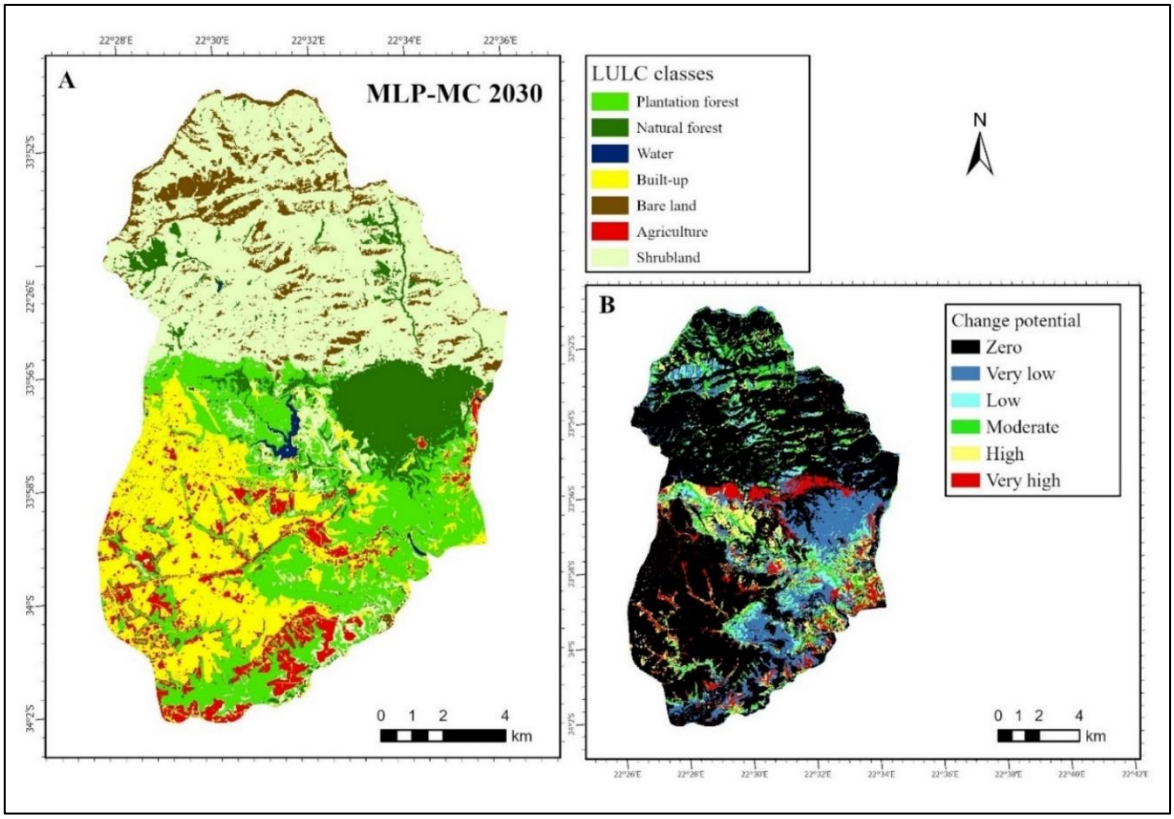

Figure 23: The DF-MC model prediction of LULC spatial distribution in the Kaaimans catchment area for the year 2030. A is the hard prediction map, B is the soft prediction map that shows the change potential.

LULC change analysis

The most significant gain is projected to occur in the plantation forest, with an increase of approximately 14.81 km² (Figure 25). Conversely, the natural forest is predicted to experience the most substantial loss, with a decrease of approximately 7.34 km². This data suggests a shift in LULC patterns between 2023 and 2030, with a notable expansion in plantation forests and a corresponding reduction in natural forests. Further losses are predicted for agriculture, and bare land, with gains in built-up areas for the year 2030.

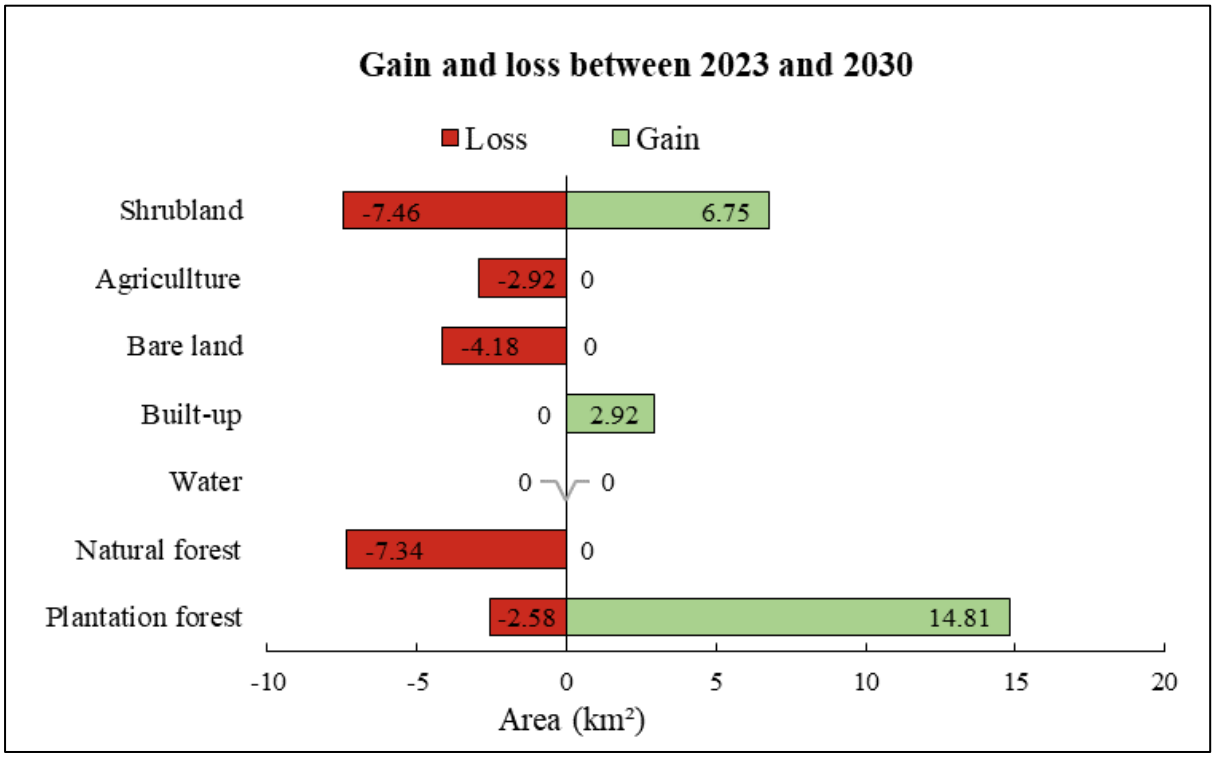

Figure 24: The gains and losses experienced by each LULC during the period of 2023 to the predicted year 2030.

Between 2023 and 2030, it is evident that most categories are projected to experience a net decrease (Figure 26). Specifically, shrubland is expected to decrease by 0.71 km², agriculture by 2.92 km², bare land by 4.18 km², and natural forest, the largest net decrease in area, with 7.34 km². On the other hand, two categories are anticipated to demonstrate an increase from 2023 to 2030. Built-up area is projected to increase by approximately 2.92 km². The water class appears to remain unchanged during this period, because the only water bodies in the study area are the Garden Route Dam and other smaller agricultural dams. Plantation forest has the largest net increase during this period, with 12.23 km².

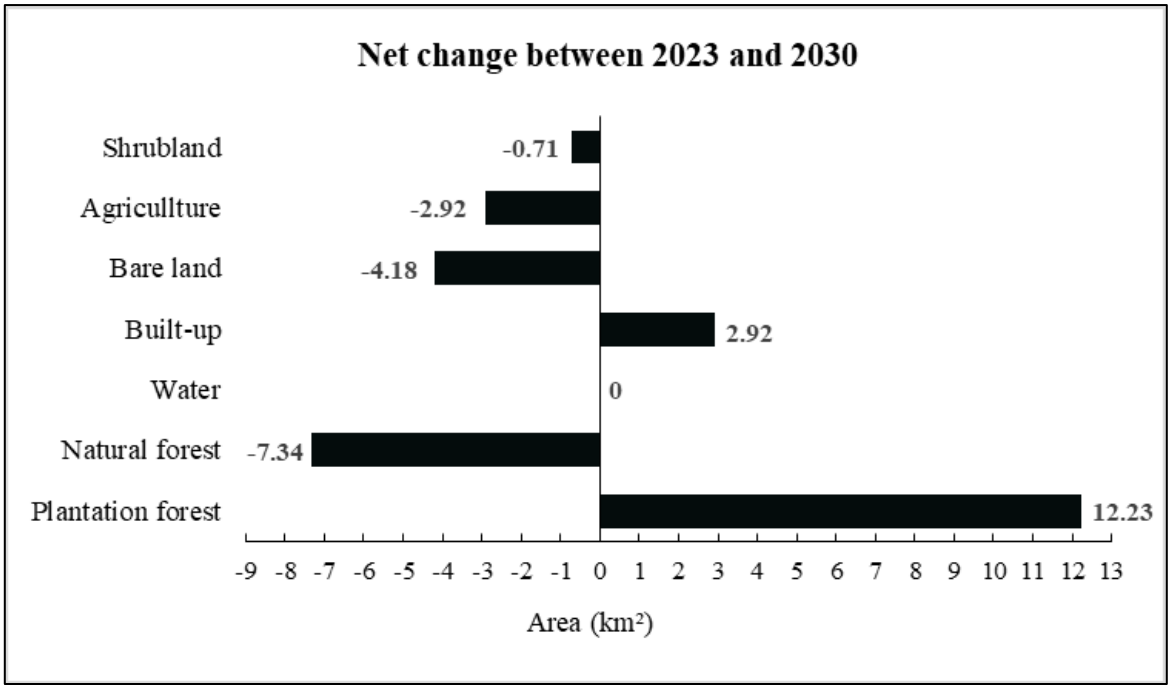

Figure 25: The net change in square kilometres for each LULC class from 2023 to 2030.

Based on assessing the contributors to the net change in each LULC class (Figure 27), the plantation forest class shows a net increase by replacing natural forest (7.34 km²) and shrubland (4.89 km²). This indicates that areas of plantation forest are expanding into these classes, causing a loss of shrubland and natural forest. Shrubland has replaced mostly bare land (4.18 km²). Built-up areas have expanded by replacing agricultural areas (2.92 km²), indicating urban expansion into agricultural lands.

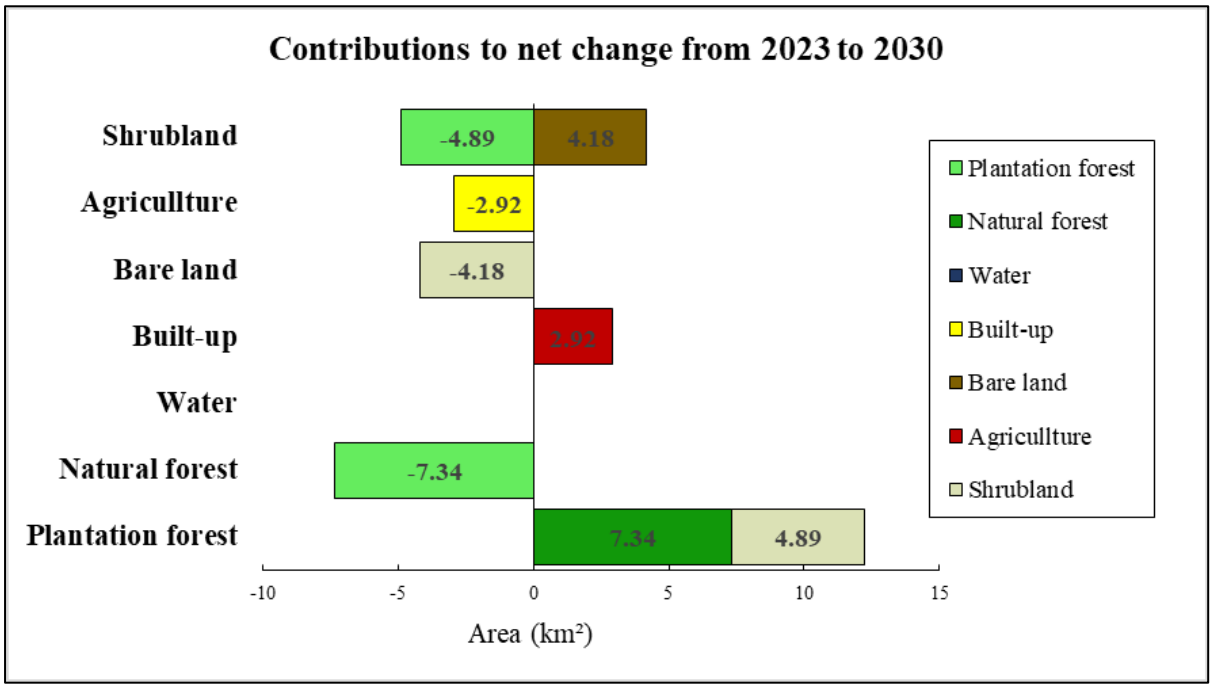

Figure 26: The positive and negative contributions to the net change of each LULC class during the period 2023 and 2030.

From the year 2023 to 2030, the individual LULC classes experienced various changes across the study area, including increases, decreases, and areas of no change (Figure 28). Both the MLP-MC and DF-MC models present similar trends of these changes for each class, albeit with minor variations. For instance, the MLP-MC model forecasts a decline in plantation forest in the northern part of the study area, whereas the DF-MC model anticipates a greater decline in the central region. The maps of natural forest from both models reveal that the losses of this class to other LULC classes are widespread across the lower half of the study area. Shrubland is also showing losses in this same region. The DF-MC model shows more agricultural land losses to the north, while the MLP-MC model shows more agricultural losses to the south.

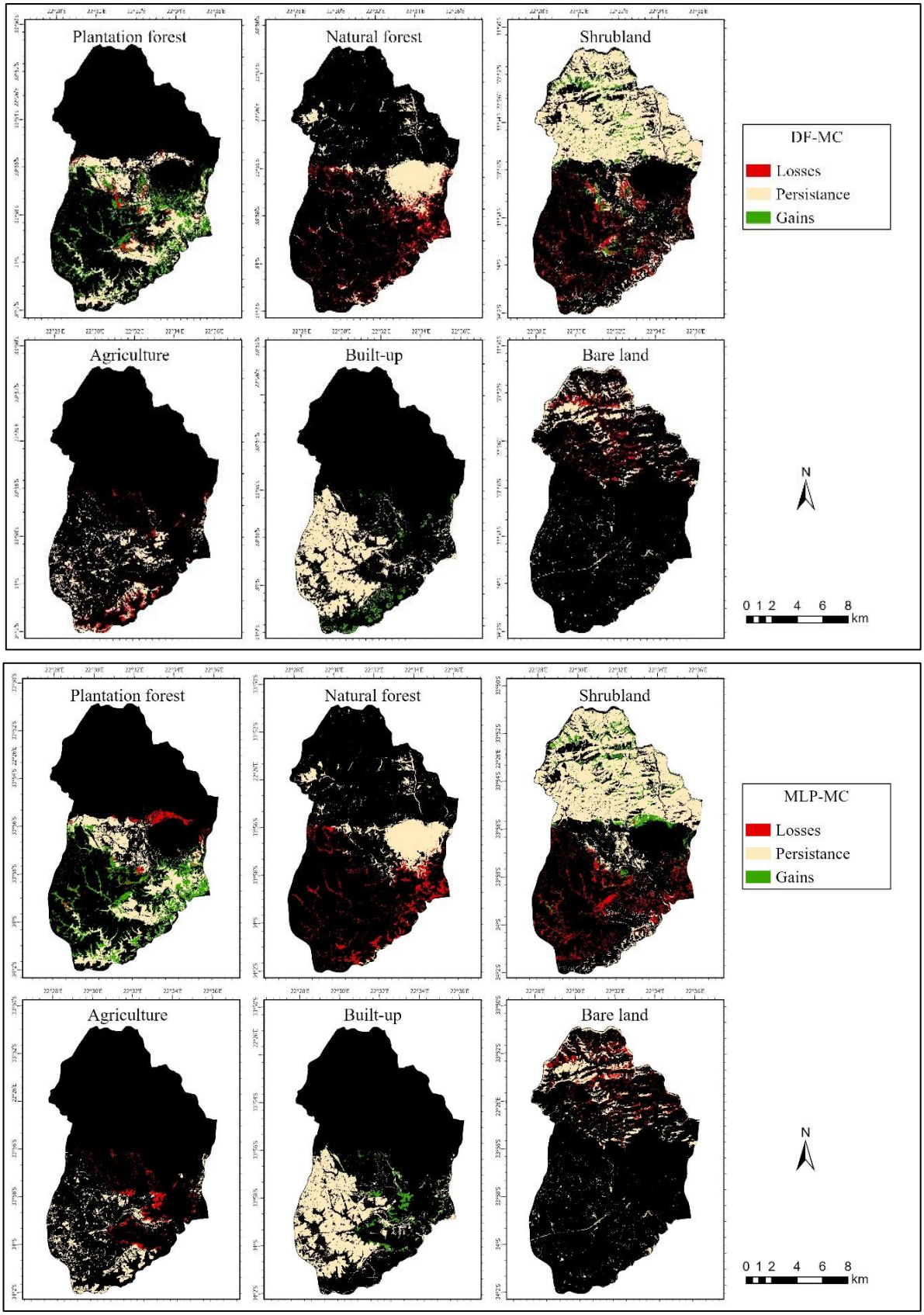

Figure 27: The losses, gains, and persistence of each LULC class based on the DF-MC (top), and MLP-MC (bottom) models, between the years 2023 and 2030.

3.4 Discussion Chapter 3

This Chapter successfully modelled the future LULC of the Kaaimans catchment area to the year 2030, using the current LULC classified maps of the years 2019, 2021, and 2023. The models incorporated static and dynamic spatial variables, providing a comprehensive approach to predicting future LULC changes. Two different models were developed using the same input maps, spatial variables, and MC matrix and therefore, thus both models predicted the same quantity of LULC change for each class, but they differed in the location and spread of the LULC classes within the study area. The purpose of creating these two predictive models was to evaluate the spatial distribution of the predicted LULCC each model produced in the context of this study area where they have been applied in the contextual geographical factors of aspects, elevation, slope, and distances from urban areas, rivers, roads, and plantation forests. This dual-model approach allowed us to cross-verify our predictions and ensure a more robust understanding of potential future changes.

The results indicate that both the DF-MC and MLP-MC models are highly effective in predicting LULCC for the year 2030, with a strong agreement between their predicted 2023 map and the reference 2023 map during the validation process (Table 22 and Figure 20). Notably, the DF-MC model exhibited a slight advantage in performance, with fewer individual grid cells incorrectly predicted and higher accuracy, as indicated by its marginally higher Kappa values (Table 22). It is worth noting that, based on a Google Scholar search of “DF-MC” and “LULC modelling,” only one other study, to date, has employed the DF-MC model for LULC modelling: Azari, et al. (2022). This study, which was carried out in the highly urbanised state of Selangor, Malaysia, successfully modelled a future scenario of land use for the year 2030 using a DF-MC model in the LCM tool of TerrSet software (Azari, et al., 2022). Despite its less frequent use compared to the MLP-MC model, the results of this Chapter suggest that the DF-MC model's predictive capabilities should not be overlooked in future LULC modelling studies.

The results of the LULCC analysis between 2023 and the modelled 2030 year by both models, showed the expansion of plantation forests into natural forested areas and built-up expansion at the edges of existing urban boundaries, replacing agricultural land (Figure 23, 24 and 27). The increase of built-up areas is often related to urban development. While the increase of plantation forests or the encroachment of exotic plantation species along the slopes of the Outeniqua Pass could be due to the forestry industry becoming a more prominent economic

activity in the region (Department of Agriculture, Forestry and Fisheries, 2023). Alternatively, it could suggest that efforts to prevent invasive species from plantations encroaching into natural areas have been less effective than desired (van Wilgen & Richardson, 2014). Long-standing plantation forests in the Garden Route area surrounding the Kaaimans catchment has been linked to the encroachment of alien plant species into the GRNP's indigenous vegetation, a trend substantiated by numerous studies (Baard & Kraaij, 2014; Kraaij, et al., 2011). This observed trend lends confidence to the predictions made by both models, which forecast an expansion of plantation forests into areas of natural forest and shrubland, especially in regions of high human disturbance such as the central and southern parts of the Kaaimans catchment.

Hybrid models like the DF-MC and MLP-MC better represent reality, and by incorporating static and dynamic variables in this study's models improved the accuracy. However, social-economic driver variables were not included in the study's models. The exclusion of socio-economic driver variables means that the models may not fully capture all the factors that influence the system (van Niekerk, 2020). This could potentially limit the accuracy and applicability of the models. For instance, urban development, which often leads to an increase in built-up areas, is typically driven by economic growth and population increase.

According to the most recent 2022 Census, George Town has experienced a significant population growth. Since the 2011 Census, the population has surged by 52%, resulting in an increase of over 100,000 people (Statistics South Africa, 2023). This substantial growth highlights the rapid urbanization specific to the study area. The expansion of plantation forests in the catchment are, as shown by chapter 2's LULC classification maps of 2019, 2021, and 2023 and this chapter's 2030 LULC prediction maps, could be influenced by changes in local or regional economic activities. The Forestry Sector Masterplan of 2020 to 2025, for instance, outlines a strategic plan for the expansion of plantation forests in the Western Cape province (Department of Trade, Industry and Competition, 2019). This plan includes the replanting and recovery of fire-damaged plantation areas. This is after the 2017 and 2018 Knysna and George fires, which destroyed large areas of plantations, infrastructure, and natural vegetation in the regions (Quiroz, et al., 2023). Furthermore, the sector plans to have 1510 km² of new plantation forests established in Limpopo, Mpumalanga, KwaZulu-Natal and closer to the study area - in the Eastern Cape as well, by 2026. Future studies could consider integrating socio-economic data, such as population growth rates, economic development indicators, and policy changes, into the modelling process.

3.4 Conclusion Chapter 3

In conclusion, the Chapter successfully modelled the future LULC of the Kaaimans catchment area up to the year 2030. The models incorporated both static and dynamic spatial variables, providing a comprehensive approach to predicting future LULC changes. Two different models, DF-MC, and MLP-MC, were developed and showed high effectiveness in predicting LULCC for the year 2030. Notably, the DF-MC model exhibited slightly superior performance, suggesting its potential for future LULC modelling studies.

The models predicted an expansion of plantation forests into natural forested areas and built-up expansion at the edges of existing urban boundaries, replacing agricultural land. These changes could be attributed to urban development and the growing prominence of the forestry industry in the region. The rapid urbanization of George Town, as evidenced by a 52% population surge since 2011, and the strategic expansion of plantation forests as outlined in the Forestry Sector Masterplan of 2020 to 2025, are key factors influencing these changes. Future studies could benefit from incorporating these variables to capture all factors influencing the system and improve the accuracy of the predictions. This study's findings contribute to a more robust understanding of potential future changes in the Kaaimans catchment area.

Overall Discussion and Conclusion

Summary

The aim of this research was to provide a comprehensive understanding of the historical, current, and future LULCC, in an important water resource management area, the Kaaimans catchment area, within the GRBR, South Africa. The study was conducted with three main objectives:

- Objective 1: The study determined the historical LULC for the Kaaimans catchment area using the SANLC data of 1990, 2014, 2018, and 2020. The analysis revealed significant changes in the LULC from 1990 to 2020. The northern half of the area, dominated by natural vegetation, reflects effective conservation efforts. In contrast, the southern half shows the expansion of human activity, including commercial forestry, urban areas, and agricultural fields. The study suggests that natural processes, possibly veldfires, drive the major land cover transitions in the northern half, while human activities primarily drive changes in the middle and southern regions. The priority LULC classes of the catchment were identified as plantation forest, natural forest, shrubland, built-up, agriculture, bare land, and water.
- Objective 2: The study produced current LULC classification maps of the Kaaimans catchment's priority classes using RF classification models and Sentinel-2 satellite data from 2019, 2021, and 2023. The results of this objective demonstrate the effectiveness of RF models in LULC classification, and all three models performed well with an overall accuracy above 85%, which makes them suitable for use as an input for many other applications. The plantation forest, built-up, and shrubland LULC classes showed a notable increasing trend between the period 2019 to 2023. Consequently, the natural forest, bare land, and agriculture area decreased. The analysis of the contributors of each LULC class have shown plantation forest areas expanding into shrubland and natural forest areas, the built-up area replacing agricultural areas, and shrubland replacing the bare land. The spatial distribution of these LULC changes and transitions corroborates the findings of Objective 1, indicating a dominance of natural LULC class interactions in the northern half of the study area and anthropogenic LULC classes in the middle to southern half of the catchment.
- Objective 3: The study predicted the future LULC of the Kaaimans catchment area up to 2030. The current classification maps, as developed in Objective 2, were used in conjunction with hybrid MC models and driver variables. Two models, DF-MC, and MLP-

MC, were developed and demonstrated high accuracy in predicting in LULCC for 2030, with the DF-MC model showing slightly superior performance. The models forecasted an expansion of plantation forests into natural forest areas and the growth of built-up areas at the edges of existing urban boundaries, replacing agricultural land. These changes are likely due to urban development and the increasing prominence of the forestry industry in the region.

Conclusion

This research provided a robust understanding of the LULCC in the Kaaimans catchment area, highlighting the significant role of both natural processes and human activities in shaping the landscape. The structure of this study based on first analysing the historical LULC data, then creating and assessing current LULC data for the area, led to a well-rounded understanding of the interactions among the dominant LULC transitions. This thorough understanding has introduced confidence in predicting future LULC based on these observed trends. The results of this study demonstrates that the use of Sentinel-2 satellite imagery for LULC classifications, when combined with remote sources for training sample creation and a meticulous study structure, is an effective and precise method for generating accurate LULC maps over time.

This research has addressed the critical challenges faced by South Africa's Pas and BRs in managing LULCC sustainably. The lack of up-to-date LULCC assessments and forecasts in areas like the Kaaimans catchment area in the GRBR has been a significant difficulty in effective land management. This study has filled this gap by providing a comprehensive analysis of the historical and current LULC data, and predicting future LULC trends in the catchment area. The study's findings have bridged the science-policy gap by providing timely, appropriately-scaled data for reserve managers, policy makers, and land use planners. The use of Sentinel-2 satellite imagery for LULC classifications and the successful application of predictive models have proven to be effective tools for generating accurate LULC maps. These maps can now serve as current datasets of LULC, thereby aiding in the sustainable management of LULCC in the region. Moreover, the study's insights into the potential impacts of urban development and the growth of the forestry industry on the Kaaimans catchment area can guide policies and strategies to manage these changes sustainably. This co-designed LULC modelling research has thus provided a robust baseline for LULCC in a strategic water management area within the GRBR, addressing the area's critical need for up-to-date data to guide land use planning and related policies.

Limitations

This study acknowledges certain limitations encountered at each chapter. The SANLC datasets used in chapter 1, may have inconsistencies due to varying methodologies over the years. The introduction of the fallow land class in later datasets and the lack of significant validation data points in the study area further complicate the interpretation of these datasets. In chapter 2, the potential for misclassifications between LULC classes and the need for more detailed LULC maps suggest that higher resolution satellite data could be beneficial. However, this study was limited to using freely available data sources. Lastly, chapter 3 highlights the potential enhancement of predictive LULC models through the integration of socio-economic data, where only topographical variables were used.

Recommendations

For future work in this catchment area, it is recommended to create more detailed LULC maps based on higher spatial resolution imagery and ground truth points that do not only rely on historical aerial imagery. Given the unique LULC composition of the study area, particularly the large prevalence of plantation forests, and the risk of invasion into natural forest areas, obtaining more detailed data on the dispersion of plantation tree species across the landscape would be of interest. The necessity for such detailed information has been underscored by this study. In the light of using higher spatial resolution satellite imagery, drones equipped with high-resolution sensors can capture detailed images of the landscape, allowing for a more precise mapping of the spatial distribution and degree of spread of invasive plant species into natural vegetation (e.g. Sladonja, et al., 2022). This could greatly aid projects focused on the removal of invasive species in the area. Furthermore, spectral samples collected using a spectrometer could provide valuable data on the major invasive plant species in the catchment (e.g. Underwood, et al., 2007). Spectrometers can measure the unique spectral signatures of different plant species, which can be used to improve the accuracy of LULC classifications. This could be particularly useful for distinguishing between different types of vegetation and identifying areas of invasion.

Considering the misclassification between natural and plantation species observed in this study, it might be beneficial to explore the use of hyperspectral imagery for future LULC classifications. Hyperspectral imagery, with its ability to capture a wide range of the electromagnetic spectrum, could potentially provide more detailed and accurate information

about different vegetation types, thereby improving the accuracy of LULC maps (e.g. Papp, et al., 2021).

Furthermore, expanding this study's low cost and effective approach that created accurate up-to-date LULC data and future LULC predictions, to other priority areas within the GRBR could significantly aid management and planning decisions. This study's methodology can provide current data in a short timeframe which makes it a valuable tool for creating current data for the BR. Each catchment area within the reserve has its own unique LULC composition and dynamics, and understanding these dynamics is crucial for effective land management. Therefore, conducting similar studies in other catchment areas and integrating the findings could provide a more comprehensive understanding of the LULC dynamics within the entire BR.

References

Abu Hatab, A., Cavinato, M. E. R., Lindemer, A. & Lagerkvist, C.-J., 2019. Urban sprawl, food security and agricultural systems in developing countries: A systematic review of the literature. *Cities,* Volume 94, pp. 129-142.

Aburas, M. M., Ahamad, M. S. S. & Omar, N. Q., 2019. Spatio-temporal simulation and prediction of land-use change using conventional and machine learning models: a review. *Environmental Monitoring and Assessment,* Volume 191.

Aguejdad, R., Houet, T. & Hubert-Moy, L., 2017. Spatial Validation of Land Use Change Models Using Multiple Assessment Techniques: A Case Study of Transition Potential Models. *Environmental Modelling and Asessment,* Volume 22, p. 591–606.

Ali, U. et al., 2022. Limiting the Collection of Ground Truth Data for Land Use and Land Cover Maps with Machine Learning Algorithms. *ISPRS Int. J. Geo-Inf.,* Issue 11, p. 333.

Al-Sharif, A. & Pradhan, B., 2014. A novel approach for predicting the spatial patterns of urban expansion by combining the chi-squared automatic integration detection decision tree, Markov chain and cellular automata models in GIS. *Geocarto International,* Volume 30, pp. 858-881.

Arfasa, G. F., Owusu-Sekyere, E. & Doke, D. A., 2023. Predictions of land use/land cover change, drivers, and their implications on water availability for irrigation in the Vea catchment, Ghana. *Geocarto International,* 38(1).

Asamoah, E. F., Beaumont, L. J. & Maina, J. M., 2021. Climate and land-use changes reduce the benefits of terrestrial protected areas. *Nature Climate Change,* Volume 11, pp. 1105-1110.

Azari, M., Billa, L. & Chan, A., 2022. Multi-temporal analysis of past and future land cover change in the highly urbanized state of Selangor, Malaysia. *Ecological Processes,* 11(2).

Baard, J. A. & Kraaij, T., 2014. Alien flora of the Garden Route National Park, South Africa. *South African Journal of Botany,* Volume 94, pp. 51-63.

Baard, J. A. & Kraaji, T., 2014. Challenges to sustainability in the Garden Route: Water, Land, and Economy. *Alien flora of the Garden Route National Park, South Africa,* Volume 94, pp. 51-63.

Barnes, E. et al., 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. *Proceedings of the 5th International Conference on Precision Agriculture,* 16-19 July.

Basheer, S. et al., 2022. Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques. *Remote Sensing,* 14(19).

Behera, M. D. et al., 2012. Modelling and analyzing the watershed dynamics using Cellular Automata (CA)–Markov model – A geo-information based approach. *Earth System Science,* Volume 121, pp. 1011-1024.

Belgui, M. & Dragut, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing,* Volume 114, pp. 24-31.

Boerner, R. E. et al., 1996. Markov models of inertia and dynamism on two contiguous Ohio landscapes. *Geographical Analysis,* 28(1), pp. 56-66.

Borrelli, P. et al., 2017. An assessment of the global impact of the 21st century land use change on soil erosion. *Nature Communications,* Volume 8.

Bourne, A. et al., 2016. A Socio-Ecological Approach for Identifying and Contextualising Spatial Ecosystem-Based Adaptation Priorities at the Sub-National Level. *PLoS ONE,* 11(5).

Breiman, L., 2001. Random forests. *Machine learning,* 45(1), pp. 5-32.

Bridgewater, P. B., 2002. Biosphere reserves: special places for people and nature. *Environmental Science Policy,* Volume 5, pp. 9-12.

Burguillo, J. C., 2017. Cellular Automata. In: *Self-organizing Coalitions for Managing Complexity.* Springer, pp. 57-67.

Cheng, K.-S.et al., 2021. Quantifying Uncertainty in Land-Use/Land-Cover Classification Accuracy: A Stochastic Simulation Approach. *Frontiers Environmental Science,* Volume 9.

Clark Labs, 2022. *TerrSet Geospatial Monitoring and Modeling Software.* [Online]
Available at: https://clarklabs.org/terrset/

ClimateData.org, 2023. *George Climate (South Africa).* [Online]
Available at: https://en.climate-data.org/africa/south-africa/western-cape/george-7176/
[Accessed 6 November 2023].

Cockburn, J. et al., 2016. Evaluating the outcomes and processes of a research-action partnership : the need for continuous reflective evaluation. *African Biodiversity and Conservation,* 46(2).

Cockburn, J. et al., 2016. How to build science-action partnerships for local land-use planning and management: lessons from Durban, South Africa.. *Ecology and Society,* 21(2).

Coetzer, K. L., Witkowski, E. T. F. & Erasmus, B. F. N., 2013. Reviewing Biosphere Reserves globally: effective conservation action or bureaucratic label?. *Biological Reviews,* 89(1), pp. 82-104.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Phychological Measurement,* 20(1).

De Visser, J., 2016. Devolution by court injunction: The case of land use planning and management in South Africa. *The Ugandan Journal of Management and Public Policy Studies,* 10(1), pp. 84-99.

Department of Agriculture, Forestry and Fisheries, 2023. *Forestry.* [Online]
Available at: https://www.gov.za/about-sa/forestry
[Accessed 1 November 2023].

Department of Cooperative Governance and Traditional Affairs, 2020. *District Profile: Garden Route.* [Online]
Available at: https://www.cogta.gov.za/ddm/wp-content/uploads/2020/08/DistrictProfile_GARDENROUTE19072020-1.pdf
[Accessed 6 June 2023].

Department of Trade, Industry and Competition, 2019. *Master Plan for the Commercial Forestry Sector in South Africa 2020-2025.* [Online]
Available at: https://www.thedtic.gov.za/wp-content/uploads/Masterplan-Forestry_Sector.pdf
[Accessed 6 November 2023].

Edge, S. & McAllister, M. L., 2008. Place-based local governance and sustainable communities: lessons from Canadian Biosphere Reserves. *Journal of Environmental planning and management,* 52(3), pp. 279-295.

ESRI, 2023. *ArcGIS Pro: Release 3.1*

Farr, T. G. et al., 2007. The Shuttle Radar Topography Mission. *Review of Geophysics,* 45(2).

Feizizadeh, B., Darabi, S., Blaschke, T. & Lakes, T., 2022. QADI as a New Method and Alternative to Kappa for Accuracy Assessment of Remote Sensing-Based Image Classification. *Sensors,* 22(12).

Feng, M. & Li, X., 2020. Land cover mapping toward finer scales. *Science Bulletin,* 65(19).

Foody, G. M., 1990. The significance of border training patterns in classification by a feedforward neural network using back propagation learning. *International Journal of Remote Sensing,* 20(18), pp. 3549-3562.

Foody, G. M., 2020. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment,* Volume 239.

Foody, G. et al., 2016. The Sensitivity of Mapping Methods to Reference Data Quality: Training Supervised Image Classifications with Imperfect Reference Data. *International Society for Photogrammetry on geo-information,* 5(11), p. 199.

GEOTERRA IMAGE, 2015. *2013-2014 South African National land-cover dataset: data User Report and MetaData..* [Online]
Available at: https://egis.environment.gov.za/data_egis/data_download/current
[Accessed June 2023].

Ghimire, B. et al., 2013. An evaluation of bagging, boosting, and random forests for land-cover classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing,* 49(5), pp. 623-643.

Ghosh, P. et al., 2017. Application of Cellular automata and Markov-chain model in geospatial environmental modeling- A review. *Remote Sensing Applications: Society and Environment,* Volume 5, pp. 64-77.

Ghosh, S. & Saini, R., 2019. Analyzing the impact of red-edge band on land use land cover classification using multispectral RapidEye imagery and machine learning techniques. *Journal of Applied Remote Sensing,* 13(4).

Gidday, B., Baard, J., Vhengani, L. & Kraaij, T., 2021. The effect of adjacent vegetation on fire severity in Afrotemperate forest along the southern Cape coast of South Africa. *Southern Forests: a Journal of Forest Science,* 83(3), pp. 225-230.

Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R., 2006. Random Forests for land cover classification. *Pattern Recognition Letters,* 27(4), pp. 294-300.

Gitelson, A. A., Gritz, Y. & Merzlyak, M. N., 2003. Relationships between Leaf Chlorophyll Content and Spectral Reflectance and Algorithms for Non-Destructive Chlorophyll Assessment in Higher Plant Leaves. *Journal of Plant Physiology,* Volume 160, pp. 271-282.

Gitelson, A. & Merzlyak, M. N., 1994. Quantitative Estimation of Chlorophyll-a Using Reflectance Spectra: Experiments with Autumn Chestnut and Maple Leaves. *Journal of Photochemistry and Photobiology,* Volume 13, pp. 247-252.

Godfrey, L., Funke, N. & Mbizvo, C., 2010. Bridging the science-policy interface : a new era for South African research and the role of knowledge brokering : research article. *South African Journal of Science,* 106(5).

Google, n.d. *Google Earth Engine.* [Online]
Available at: https://code.earthengine.google.com/
[Accessed July 2023].

Guerbois, C., Brady, U., de Swardt, A. G. & Fabricius, C., 2019. Nurturing ecosystem-based adaptations in South Africa's Garden Route: a common pool resource governance perspective. *Regional Environmental Change,* pp. 1849-1863.

Guerra, C. A., Rosa, I. M. D. & Pereira, H. M., 2019. Change versus stability: are protected areas particularly pressured by global land cover change?. *Landscape Ecology,* Volume 34, pp. 2779-2790.

Huete, A. et al., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment,* 83(1-2), pp. 195-213.

Huete, A. R., 1998. A Soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment,* Volume 25, pp. 295-309.

IPBES, 2018. The IPBES assessment report on land degradation and restoration. *Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services,* p. 744.

Kavzoglu, T., 2009. Increasing the accuracy of neural network classification using refined training data. *Environmental Modelling and Software,* 24(7), pp. 850-858.

Kepe, T., 2008. Land Claims and Comanagement of Protected Areas in South Africa: Exploring the Challenges. *Environmental Management,* Volume 41, pp. 311-321.

Kibena, J., Nhapi, I. & Gumindoga, W., 2014. Assessing the relationship between water quality paramenters and changes in landuse patterns in the Upper Manyame River, Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C,* Volume 67-69, pp. 153-163.

Kim, Y., Newman, G. & Guneralp, B., 2020. A Review of Driving Factors, Scenarios, and Topics in Urban Land Change Models. *Land,* 9(8).

Kleijn, D. et al., 2020. Chapter Four - Integrating biodiversity conservation in wider landscape management: Necessity, implementation and evaluation. *Advances in Ecological Research,* Volume 63, pp. 127-159.

Koomen, E. & Borsboom-van Beurden, J., 2011. *Land-use modelling in planning practice..* XVI-214 ed. s.l.:Springer.

Kraaij, T. et al., 2018. An assessment of climate, weather, and fuel factors influencing a large, destructive wildfire in the Knysna region, South Africa. *Fire Ecology,* Volume 14, pp. 1-12.

Kraaij, T., Cowling, R. M. & Van Wilgen, B. W., 2011. Past approaches and future challenges to the management of fire and invasive alien plants in the new Garden Route National Park : review article. *South African Journal of Science,* 107(9).

Kuc, G. & Chormanski, J., 2019. Sentinel-2 imagery for mapping and monitoring imperviousness in urban areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XLII-1/W2.

Kumar, V. & Agrawal, S., 2023. A multi-layer perceptron–Markov chain based LULC change analysis and prediction using remote sensing data in Prayagraj district, India. *Environmental Monitoring and Assessment,* 195(619).

Lakhraj-Govender, R. & Grab, S. W., 2019. Rainfall and river flow trends for the Western Cape Province, South Africa. *South African Journal of Science,* 115(9-10), pp. 1-6.

Le Maitre, D. C., Kotzee, I. M. & O'Farrell, P. J., 2014. Impacts of land-cover change on the water flow regulation ecosystem service: Invasive alien plants, fire and their policy implications. *Land Use Policy,* Volume 36, pp. 171-181.

Le Maitre, D., Forsyth, G. G., Dzikiti, S. & Gush, M. B., 2016. Estimates of the impacts of invasive alien plants on water flows in South Africa. *Water SA,* 42(4).

Leta, M. K., Demissie, T. A. & Tranckner, J., 2021. Modeling and Prediction of Land Use Land Cover Change Dynamics Based on Land Change Modeler (LCM) in Nashe Watershed, Upper Blue Nile Basin, Ethiopia. *Sustainability,* 13(7), pp. 37-40.

Leutner, B., 2023. *RStoolbox: Tools for Remote Sensing Data Analysis.,* s.l.: Rpackage version 0.3.0.

Lui, Y., Qian, J. & Yue, H., 2021. Comprehensive Evaluation of Sentinel-2 Red Edge and Shortwave-Infrared Bands to Estimate Soil Moisture. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* Volume 14, pp. 7448-7465.

Mani, S., Osborne, C. P. & Cleaver, F., 2021. Land degradation in South Africa: Justice and climate change in tension. *People and Nature,* 3(5), pp. 978-989.

Marques, A. et al., 2019. Increasing impacts of land use on biodiversity and carbon sequestration driven by population and economic growth. *Nature Ecology and Evolution,* Volume 3, pp. 628-637.

Matlhodi, B., Kenabatho, P. K., Parida, B. P. & Maphanyane, J. G., 2021. Analysis of the Future Land Use Land Cover Changes in the Gaborone Dam Catchment Using CA-Markov Model: Implications on Water Resources. *Remote Sensing,* 13(13), pp. 24-27.

Meganck, R. A., 1990. Biosphere reserve management concept. In: *Encyclopedia of Earth Science.* s.l.:Springer, Dordrecht, pp. 58-59.

Misra, G., Cawkwell, F. & Wingler, A., 2020. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sensing,* 12(17).

Moodley, K., Toucher, M. L. & Lottering, R. T., 2023. Simulating future land-use within the uThukela and uMngeni catchments in KwaZulu-Natal. *Scientific Africa,* Volume 20.

Moser, C. S., 1996. A partial instruction module on global and regional land use/land cover change: assessing the data and searching for general relationships. *GeoJournal,* Volume 39, pp. 241-283.

Msweli, S., Potts, A., Fritz, H. & Kraaij, T., 2020. Fire weather effects on flammability of indigenous and invasive alien plants in coastal fynbos and thicket shrublands (Cape Floristic Region). *Ecology.*

Mugari, E. & Masundire, H., 2022. Consistant changes in Land-Use/Land-Cover in Semi-Arid Areas: Implications on Ecosystem Service Delivery and Adaptation in the Limpopo Basin, Botswana. *Land,* 11(11), p. 2057.

Mukwada, G. & Manatsa, D., 2017. Acacia mearnsii management in a South African National Parks: SWOT analysis using hot topics in biological invasion as a guide. *Journal of Mountain Science,* Volume 14, p. 205–218.

Musetsho, K. D., Chitakira, M. & Nel, W., 2021. Mapping land-use/land-cover change in a critical biodiversity area of South Africa. *International Journal of Environmental Research and Public Health,* 18(16), p. 10164.

Nasiri, V. et al., 2022. Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. *Remote Sensing,* 14(9).

Neidertscheider, M., Gingrich, S. & Erb, K.-H., 2012. Changes in land use in South Africa between 1961 and 2006: an integrated socio-ecological analysis based on the human appropriation of net primary production framework. *Regional Climate Change,* Volume 12, pp. 715-727.

Nkosi, M., Mathivha, F. I. & Odiyo, J., 2021. Impact of land management on water resources, a South African context. *Sustainability,* 13(2), p. 701.

Osgouei, P. E., Kaya, S., Sertel, E. & Alganci, U., 2019. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote sensing,* 11(3).

Palliwoda, J., Buermann, A., Kreamer, R. & Schoter, M., 2021. Zoning of UNESCO Biosphere Reserves: A comprehensive Set of Geodata for Europe. *Frontiers in Ecology and Evolution,* Volume 9.

Pandy, W. R. & Rogerson, C. M., 2021. Coastal tourism and climate change: risk perceptions of tourism stakeholders in South Africa's garden route. *GeoJournal pf Tourism and Geosites,* 37(3), pp. 730-739.

Papp, L., Van Leeuwen, B., Szilassi, P., Tobak, Z., Szatmári, J., Árvai, M., Mészáros, J. and Pásztor, L., 2021. Monitoring invasive plant species using hyperspectral remote sensing data. *Land*, *10*(1), p.29.

Phiri, D. et al., 2020. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing,* 12(14).

Pontius, R. G. J., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing,* 68(10), pp. 1041-1049.

Pool-Stanvliet, R. & Coetzer, K., 2020. The scientific value of UNESCO Biopshere Reserves. *South African Journal of Science,* 116(2), p. 4.

QGIS Development Team, 2002. *QGIS Geographical Information System.* [Online]
Available at: https://qgis.osgeo.org

Quiroz, N. F. et al., 2023. Analysis of the 2017 Knysna fires disaster with emphasis on fire spread, home losses and the influence of vegetation and weather conditions: A South African case study. *International Journal of Disaster Risk Reduction,* Volume 88.

Qui, S., He, B., Yin, C. & Liao, Z., 2017. Assessments of Sentinel-2 vegetation red-edge spectral bands for improving land cover classification. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XLII-2/W7, p. 871–874.

R Core Team, 2021. *R: A language and environment for statistical computing.* [Online]
Available at: https://www.R-project.org/

Reed, M. G., 2019. The contributions of UNESCO Man and Biosphere Programme and biosphere reserves to the practice of sustainability science.. *Sustainability Science,* Volume 14, pp. 809-821.

Richards, J. & Jia, X., 2006. *Remote Sensing Digital Image Analysis..* Fourth ed. New York: Springer.

Rodriguez-Galiano, V. F. et al., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* Volume 67, pp. 93-104.

Rogers, G., 2019. *Herald LIVE.* [Online]
Available at: https://www.heraldlive.co.za/news/2019-10-22-sanparks-completes-controlled-burn-in-garden-route-park/
[Accessed 2 November 2023].

Rose, D. C. et al., 2019. Calling for a new agenda for conservation science to create evidence-informed policy. *Biologucal Conservation,* Volume 238, pp. 108-222.

Rouse, J. W. et al., 1973. Monitoring the Vernal Advancement and Retrogradation (Green WaveEffect) of Natural Vegetation. *NASA,* p. 112.

Safabakhshpachehkenari, M. & Tonooka, H., 2023. Assessing and Enhancing Predictive Efficacy of Machine Learning Models in Urban Land Dynamics: A Comparative Study Using Multi-Resolution Satellite Data. *Remote Sensing,* 15(18), pp. 44-95.

Saha, P., Mitra, R., Chakraborty, K. & Roy, M., 2022. Application of multi layer perceptron neural network Markov Chain model for LULC change detection in the Sub-Himalayan North Bengal. *Remote Sensing Applications: Society and Environment,* Volume 26.

Samardžić-Petrović, M., Dragićević, S., Bjat, B. & Kovačević, M., 2015. Exploring the Decision Tree Method for Modelling Urban Land Use Change. *Geomatica,* 69(3), pp. 313-325.

Sauka, S., 2016. Flood risk assessment of the Crocodile River, Mpumalanga. *University of the Witwatersrand, Johannesburg, South Africa.*

Shafie, B. et al., 2023. Modeling land use/cover change based on LCM model for a semi-arid area in the Latian Dam Watershed (Iran). *Environmental Monitoring and Assessment,* Volume 363.

Shen, L. et al., 2020. Multi-Layer Perceptron Neural Network and Markov Chain Based Geospatial Analysis. *Journal of Environmental Informatics Letters,* 3(1), pp. 29-39.

Shrestha, M., Mitra, C., Rahman, M. & Marzen, L., 2023. Mapping and Predicting Land Cover Changes of Small and Medium Size Cities in Alabama Using Machine Learning Techniques. *Remote Sensing,* 15(1).

Shumba, T. et al., 2021. The influence of biophysical and socio-economic factors on the effectiveness of private land conservation areas in preventing natural land cover loss across South Africa. *Global Ecology and Conservation,* Volume 28.

Simmons, C. et al., 2021. Projecting effects of land use change on human well-being through changes in ecosystem services. *Ecological Modelling,* Volume 440, p. 109358.

Sinthumule, N. I., Ratshivhadelo, T. & Nelwamondo, T., 2020. Stakeholder perspectives on land-use conflics in the South African section of the Greater Mapungubwe Transfrontier Conservation Area. *Journal of Land Use Science,* 15(1), pp. 11-24.

Skowno, A. L. et al., 2018. National Biodiversity Assessment 2018: The status of South Africa's ecosystems and biodiversity. South African National Biodiversity Institute, an entity of the Department of Environment, Forestry and Fisheries, Pretoria. pp. 1–214.

Sladonja, B., Damijanić, D., Krapac, M., Uzelac, M., Linić, I. and Poljuha, D., 2022. Development of drone-based methodology for inventory and monitoring invasive plants along river banks in Croatia. *Management of Biological Invasions*, *13*(4), pp.679-689.

Statistics South Africa, 2023. *Census 2022.* [Online]
Available at: https://www.statssa.gov.za/?p=16716 [Accessed 6 November 2023].

Steyn, C. J., 2013. Exploring ethical challenges, climate change and implications on land and water use within the agricultural sector of the Garden Route, Western Cape, South Africa. *Department of Philosophy, Stellenbosch University, Thesis*

Tanner, J. et al., 2022. Impacts of climate change on rivers and biodiversity in a water-scarce semi-arid region of the Western Cape, South Africa. *Frontiers Water,* Volume 4.

Tizora, P. L. R. A. M. G. a. C. A., 2016. Land use and land cover change in the Western Cape Province: quantification of changes & understanding of driving factors. *Planning Africa 2016; Making Sense of the Future – Disruption and Reinvention.*

Turner, B. L., Lambin, E. & Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences,* 104(52), pp. 20666-20671.

Turner, B. L. & Meyer, W. B., 1996. Global land-use/land-cover change: challenges for geographers. *GeoJounal,* Volume 39, pp. 237-240.

Tuswa, N. et al., 2019. The impacts of commercial plantation forests on groundwater recharge: A case study from George (Western Cape, South Africa). *Physics and Chemistry of the Earth, Parts A/B/C,* Volume 112, pp. 187-199.

Underwood, E.C., Ustin, S.L. and Ramirez, C.M., 2007. A comparison of spatial and spectral image resolution for mapping invasive plants in coastal California. *Environmental management*, *39*, pp.63-83.

UNESCO, 2021. *UNESCO Biosphere Reserve.* [Online]
Available at: https://en.unesco.org/node/314143
[Accessed 20 April 2023].

van Niekerk, A. J., 2020. Inclusive Economic Sustainability: SDGs and Global Inequality. *Sustainability,* 12(13).

van Wilgen, B. et al., 2016. Historical costs and projected future scenarios for the management of invasive alien plants in protected areas in the Cape Floristic Region. *Biological Conservation,* Volume 200, pp. 168-177.

van Wilgen, B. & Richardson, D. M., 2014. Challenges and trade-offs in the management of invasive alien trees. *Biological Invasions,* Volume 16, pp. 721-734.

Wang, J. & Maduako, I. N., 2018. Spatio-temporal urban growth dynamics of Lagos Metropolitan Region of Nigeria based on Hybrid methods for LULC modeling and prediction. *European Journal of Remote Sensing,* 51(1).

Weyer, V. D., Granger, J. E., Hill, T. R. & O'Connor, T. G., 2015. Land transformation and its implication for biodiversity integrity and hydrological functioning from 1944 to 1990, Karkloof catchment, South Africa. *Bothalia - African Biodiversity & Conservation,* 45(1), pp. 1-13.

Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M., 2021. Global land use changes are four times greater than previously estimated. *Nature Communications,* 11 May.Volume 12.

Wong, B. et al., 2018. Assessment of Sentinel-2 MSI Spectral Band Reflectances for Estimating Fractional Vegetation Cover. *Remote Sensing,* 10(12), pp. 19-27.

Woyessa, Y. E., Hensley, M. & van Rensburg, L. D., 2006. Catchment management in semi-arid area of central South Africa: Strategy for improving water productivity. *Water SA,* 32(5).

Wynberg, R., 2002. A decade of biodiversity conservation and use in South Africa: Tracking progress from the Rio Earth Summit to the Johannesburg World Summit on Sustainable Development. *South African Journal of Science,* Volume 68, pp. 233-243.