3rd Sem Mini Project Report on

Examining the Effect of Urbanization on a landmass using LULC, ML and RS techniques

Submitted in partial fulfillment of the requirement for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by:

Student Name	University Roll No.
Pragya Thapliyal	2023516

Under the Mentorship of
Dr. Hemant Singh Pokhariya

Department of Computer Science and Engineering
Graphic Era (Deemed to be University)
Dehradun, Uttarakhand

January-2025

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project report entitled **"Examining the Effect of Urbanization on a landmass using LULC, ML and RS techniques"** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the undersigned under the mentorship of **Dr. Hemant Singh Pokhariya**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Pragya Thapliyal	2023516	Pragya

The above mentioned student shall be working under the supervision of the undersigned on the **"Examining the Effect of Urbanization on a landmass using LULC, ML and RS techniques**
"

Dr. Hemant Singh Pokhariya **Head of the Department**

<u>Examination</u>

Name of the Examiners: **Signature with Date**

1.

2.

Table of Contents

Chapter 1

Abstract and Introduction

1.1 Abstract

It is vital for public awareness and decision-makers to investigate the impact caused by changes in urban land use and land cover on ecosystem service values (ESVs) in order to promote sustainable urban development, protect the environment, protect public health, and ensure communities' long-term well-being. Changes in ecosystem services brought on by changes in land use and cover (LULC) are a crucial early indicator and warning sign of ecological shifts. Using remote sensing and GIS tools, this study examines the shifting dynamics of ESVs in a part of Dehradun district of Uttrakhand state, India that covers Lachhiwala range,Ramgarh range, the river Song and major area of Doiwala, Teliwala and Bullawala . The evaluation of land use land cover (LULC) changes was carried out by analyzing satellite images of the research region of 2018 and 2024. In addition, we examined changes in ecosystem services according to the study area's LULC classifications. Using a random forest machine learning classifier, satellite images were classified into five categories of land cover: water bodies, forests, urban area, agricultural land, and barren land.

1.2 Introduction

LULC refers to changes brought to the land through continuous human use. These changes have become pivotal in contemporary environmental regulation and biodiversity oversight. In current strategies for sustainable land resource management, LULC changes play a crucial role [1].

To detect and monitor LULC patterns, RS and GIS have proven to be very useful [2]. These methods are less time-consuming, and the use of high-resolution images makes it easier to classify areas for study. GIS and remote sensing techniques employ various image classification methods, such as supervised and unsupervised classification, for digital image analysis [3]. Among these methods, supervised classification is the most commonly used for LULC classification [4].

Various types of machine learning algorithms used include feature bagging, SVM, and neural networks [5].Machine learning is one of the most dependable methodologies for categorizing nonlinear systems. These methods have demonstrated broad success in LULC studies [6]. Machine learning classifier algorithms can produce reliable results even with complex, large, high-dimensional data [7]. The classification of high-resolution satellite images becomes a computation-intensive process and poses challenges. Some popular machine learning algorithms used include classification and regression, k-nearest neighbor, support vector machine, convolutional neural network, and artificial neural network. CART builds standard decision trees from the given dataset, while random forest uses random subsets of the dataset to develop multiple decision trees [8]. K-Nearest neighbor uses data from nearby pixels to group the underlying patterns of the training dataset [9].

In recent times, the machine learning random forest classifier has been widely utilized for classifying heterogeneous landscapes [10]. It aggregates results from multiple decision trees to provide more accurate outcomes, with the number of trees being a parameter. By combining decision trees that incorporate randomly selected features from training samples, it aims to accommodate all relevant characteristics. A single decision tree may not fully capture the significance of each input attribute and could favor certain features during classification, potentially leading to biased results. Therefore, random forest helps in understanding the relative importance of different factors extracted from satellite image bands. Due to its robustness against noise and outliers, random forest has gained significant relevance and usage. Google Earth Engine, a cloud-based platform, offers multi-processing environments for conducting image processing and classification tasks.

Chapter 2

Material and Method

2.1 Study Area

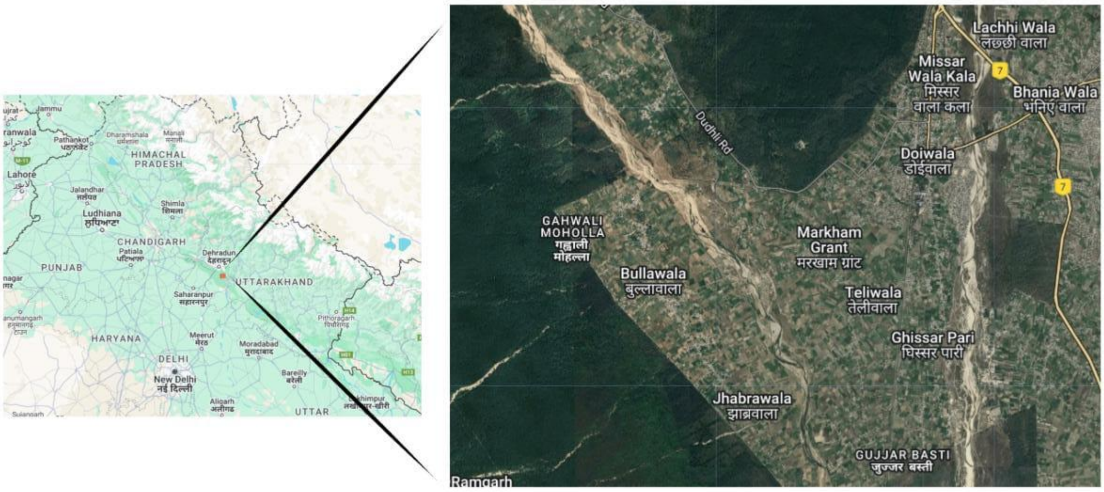


Fig. 2.1 Region of Interest

The chosen Area of Interest (AOI) is situated in Indias Uttarakhand state, in the city of Dehradun known for its landscapes and serene valleys dotted with forests and rivers aplenty.This research hones in on the vicinity surrounding Rishikesh including the Lachhiwala and Ramgarj forest areas along with the Song River as notable zones, like Doiwala,Teliwala and Bullawala.

This area is an option, for classifying Land Use and Land Cover (LULC) because of its scenery encompassing water bodies, dense forests, urban zones, barren land and farmlands. The diverse range of land types offers a backdrop for researching and understanding land use trends. How they interact with the environment, around them.

2.2 Data Used

In remote sensing and classifying different land forms over a piece of land, every land form that was taken into consideration and was made as a separate class .These classes had to be studied by different spectral bands provided by the Sentinel-2 images. Sentinel-2 is a Europian satellite designed for Earth Observation. It carries 13 spectral bands ranging from 10 to 60-meter pixel size. (B2) band for blue, (B3) band for green, (B4) band for red and (B8) band for near-infrared have a 10-meter resolution. Other than that, band (B5) for its red edge, bands (B6, B7, and B8A) for ear infrared (NIR) and bands (B11 and B12) for short-wave infrared, have a ground sampling distance of 20 meters. And at last band (B1) for coastal aerosol, band (B9) for water vapor and the band (B10) for circus band, have 60-meter pixel size.

For achieving our goal , here we use bands like B2, B3, B4, B8 and B11 and their various combinations to classify the study area.

Chapter 3

Methodology

Workflow used in this study is shown in the Fig. 3.1

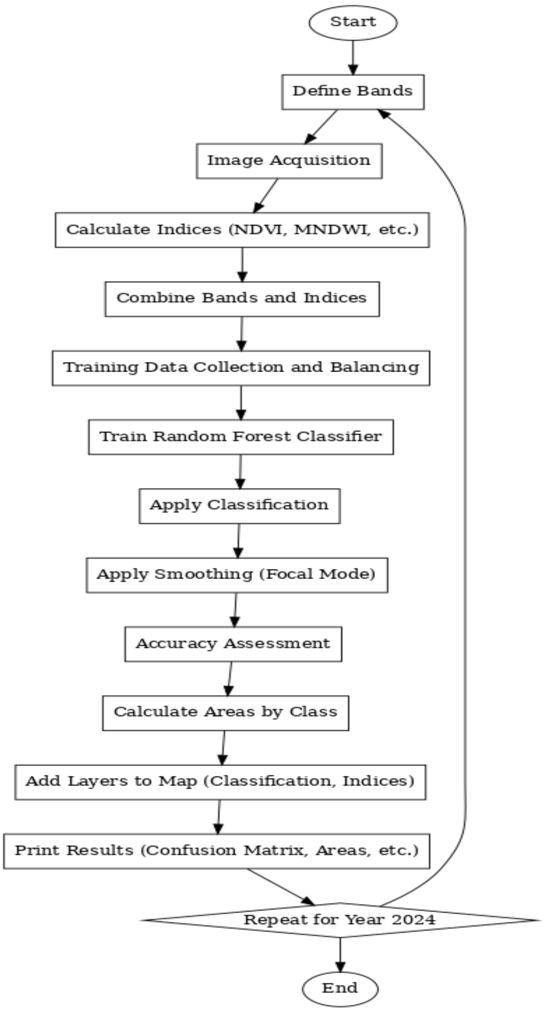


Fig. 3.1 Workflow chart used in this study

3.1. Data acquisition and preprocessing

The first step is to select a study area and gather the required Landsat images. The preprocessing of data involved filtering the images. Preprocessing refers to the transformation of unprocessed data into a format that can be understood and analyzed by computers and machine learning algorithms. The following stage involves filtering the data because different topography and atmospheric conditions can lead to changes in the data. Therefore, in this investigation, radiometric corrections, including atmospheric correction, were used.

3.2. Classification

The training/testing samples are created by adding markers for each class using Google Earth Engine. The Google Earth Engine provides the provision to generate the training samples. The study area was divided into 5 distinct, non-overlapping sub regions based on the land cover classes. 80 elements for 'Water' class, 138 for 'Agriculture', 87 for 'Forest', 146 for 'Barren_land' and 170 for 'Urban_area' were selected very carefully. These markers were marked by experimenting with different band combinations and using one that was suitable for that particular class.

3.2.1 NDVI

NDVI highlights vegetation by comparing the Near-Infrared (NIR) band and the Red band. Healthy vegetation reflects more NIR and less Red light.

$$NDVI = \frac{NIR - RED}{NIR + RED} \qquad \ldots(3.1)$$

3.2.2 NDBI

NDBI is an index used to identify built-up or urban areas in satellite imagery. NDBI values closer to 1 indicate urban areas, while negative values indicate water or non-urban regions.

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \qquad \ldots(3.2)$$

3.2.3 MNDWI

MNDWI emphasizes water bodies by contrasting the Green band with the Shortwave Infrared (SWIR1) band.

$$MNDWI = \frac{GREEN - SWIR1}{GREEN + SWIR1} \qquad \ldots(3.3)$$

3.3. Accuracy assessment

Accuracy evaluation is a crucial stage in the processing of data in remote sensing. Accuracy assessment is used to validate the results obtained from the LULC analysis. It determines whether the image pixels from remotely sensed data are correctly classified or not, hence determining the quality of the information obtained. The result could be graphs, diagrams, tables, matrices, etc.Here to measure the accuracy we have found the kappa coefficient and the overall accuracy(%).

3.3.1 Kappa Coefficient

Kappa analysis is a discrete multivariate technique to determine accuracy and serves as a gauge of correctness or agreement [11].

$$KC = \frac{(TSxTCS) - \sum(Column\ Total \times Row\ Total)}{TS^2 - \sum(Column\ Total - Row\ Total)} \qquad \ldots(3.4)$$

Where, TCS = Total Number of Correctly Classified Pixel

TS = Total Number of Reference Pixel

3.3.2 Overall Accuracy

It measures the proportion of correctly classified pixels in relation to the total pixels, indicating the classifier's performance. Higher accuracy signifies better classification reliability.

$$OA = \frac{sum\ of\ corerct\ predictions \quad X\ 100}{total\ no.of\ predictions} \qquad \ldots(3.5)$$

Chapter 4

Result and Discussion

4.1 LULC Classification

Classified maps of year 2018 and 2024, shown in Fig. 4.1(a) and 4.1(b) resp., are obtained using supervised classification. Random forest classifier is used to classify LANDSAT images into 5 different land cover classes. These five LULC classes are water, agriculture, forest, barren land and urban area which are represented by blue, yellow, green, grey and red color respectively.

From these images this can be clearly observed that the water of river Song have increased , this is the result of Global Warming that the glaciers are melting and hence increasing the Water .

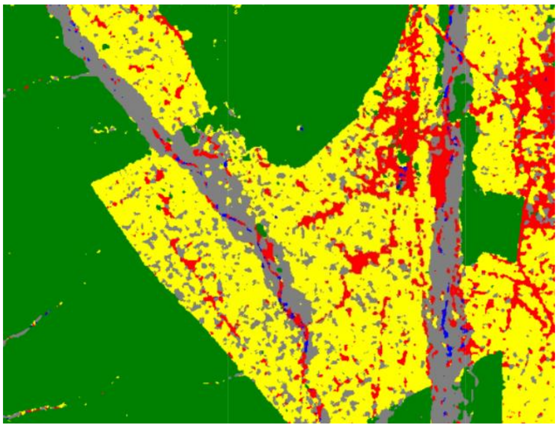


Fig. 4.1(a)
LULC image of the year 2018

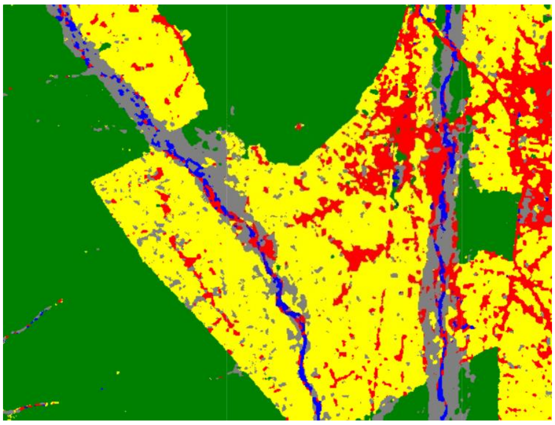


Fig. 4.2(b)
LULC image of the year 2024

Fig. 4.2(a)
Sentinel-2 image of ROI (2018)

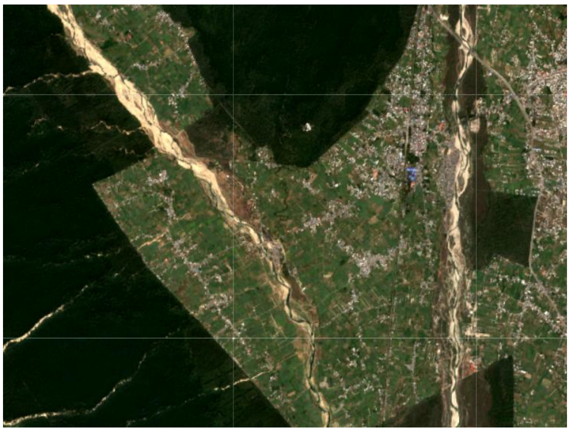


Fig. 4.2(b)
Sentinel-2 image of ROI (2014)

From Fig 4.2 we can also see the increase in agricultural land and decrease in barren land from 2018 to 2024.

4.2 Spectral Indices

The index layers for NDVI, NDBI, and MNDWI illustrate specific characteristics of the landscape:

- **NDVI** (Normalized Difference Vegetation Index) highlights vegetation health and density, where higher values indicate lush vegetation (shown in green). Fig 4.3
- NDBI (Normalized Difference Built-up Index) emphasizes urban or built-up areas, with higher values indicating dense infrastructure (shown in shades of red). Fig 4.4
- **MNDWI** (Modified Normalized Difference Water Index) focuses on water bodies, with higher values representing areas of open water (shown in blue). Fig 4.5

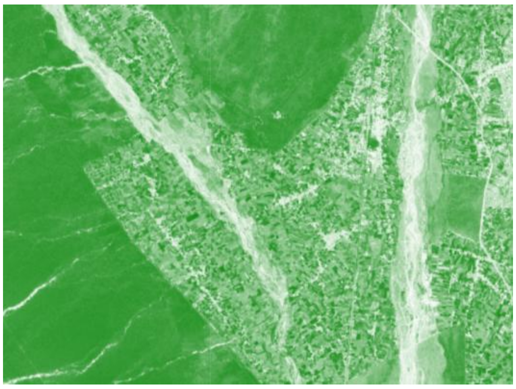


Fig. 4.3(a)
NDVI 2018

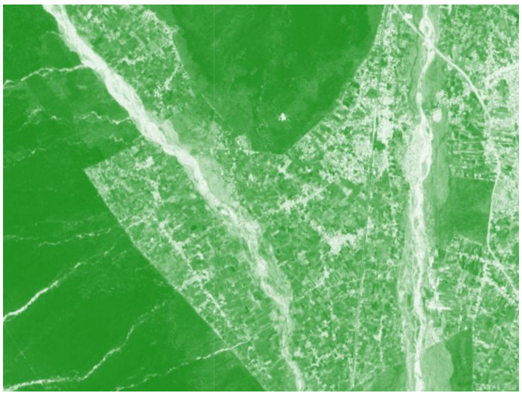


Fig. 4.3(b)
NDVI 2024

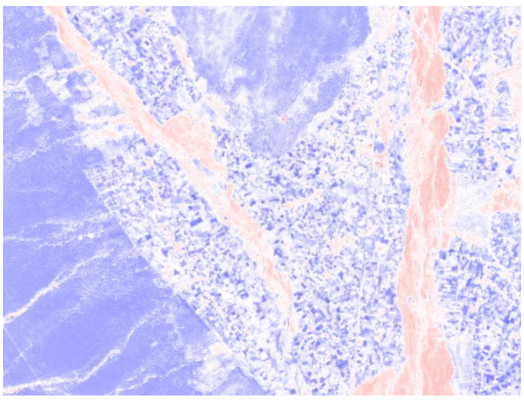


Fig. 4.4(a)
NDBI 2018

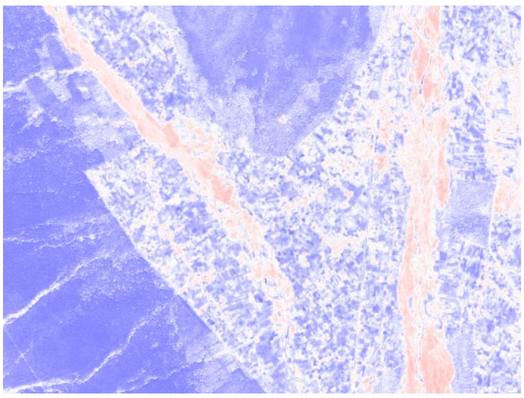
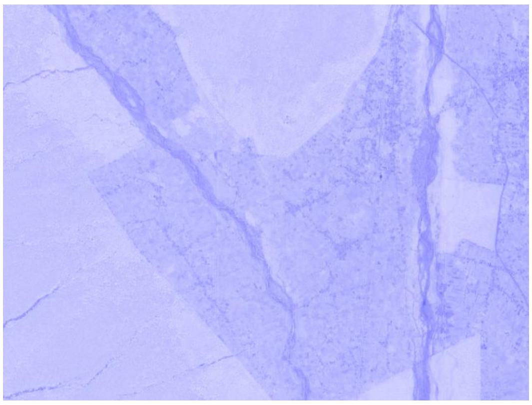


Fig. 4.4(b)
NDBI 2024



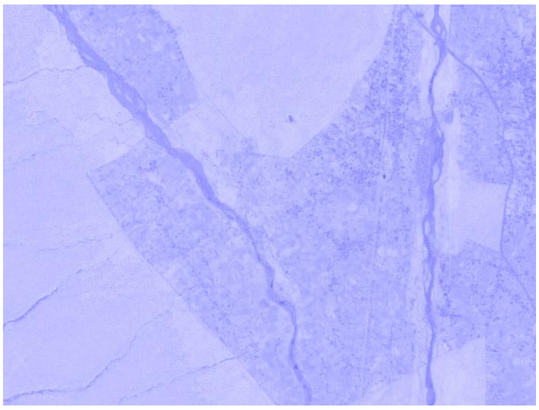


Fig. 4.5(a)

MNDWI 2018

Fig. 4.5(b)

MNDWI 2024

4.2 Change Detection from 2018 to 2024

Analyses of change detection pinpoint and quantify differences between pictures taken of the same area at various points in time (2018 to 2024). This approach greatly aids in identifying the distinct modifications that occurred in the various LULC classes. From table 5.1 and Fig. 5.1 we can infer that various changes have been brought in different land classes.

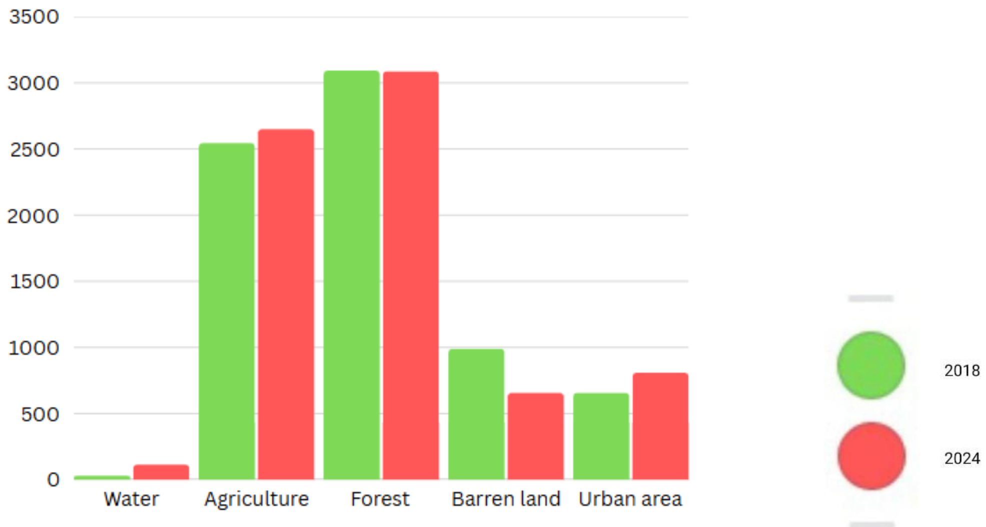


Fig. 5.1

Bar Graph of Area change of different Land Classes from 2018 to 2024

LULC Classes	Area in 2018(hec.)	Area in 2024(hec.)
WATER	**29.84**	**115.09**

AGRICULTURE	**2546.01**	**2649.66**
FOREST	**3093.61**	**3088.16**
BARREN LAND	**990.50**	**655.70**
URBAN AREA	**657.69**	**809.04**

Table 5.1 Table showing the area in hectares of different land cover classes in the year 2018 and 2024

The Kappa calculated using the code in google earth engine is 0.8412 with an OA of 87.60% in the year 2018 and 0.8689 with OA of 89.66% in the year 2024.

The spectral reflectance of barren land was difficult to distinguish from urban area as both of them reflected bright light .Factors like this resulted in lowering the kappa value and overall accuracy.

Chapter 5

Conclusion

The present study is an attempt to examine the effect of urbanization on LULC and environmental quality using RS and GIS methods. The study was conducted to understand LULC change and its relation with ecosystem service value. This study explores three SENTINEL-2 images over the period 5 years to produce accurate trends of changes in the LULC of the region of interest. The changes in the year 2018 and 2024 have been classified into five classes: water bodies, forest area, urban area, agricultural land, and barren land. Next, the random forest machine learning classifier model—which had an accuracy of 87.60% in 2018 and 89.66% in 2024 is used to analyze the classification findings.

We can also say that, the only way to implement an appropriate management plan for rapid urbanization is to examine and control the modifications made to the various LULC classes for sustainable development of land resources effectively and improved understanding on the LULC change and its impact on ESVs is required.

Guide interaction form

Guide Interaction Form

Name of the Student : Pragya Thapliyal

University Id of the Student: 23022903

Section : E

Name of the Guide : Dr. Hemant Singh Pokhariya

S. No.	Date	Task Assigned	Task Status	Guide's Sign.
01	03/09/24	Project Synopsis	Done	
02	10/09/24	Research Paper	Done	
03	17/09/24	Determining ROI	Done	
04	24/09/24	How to classify	Done	
05	8/10/24	LULC for 2018 image	Done	
06	18/11/24	Making Dataset	Done	
07	17/12/24	Refining accuracy	Done	
08	24/12/24	Project Report.	Done.	

References

[1]*Rawat J S, Biswas V and Kumar M 2013 Changes in land use/cover using geospatial techniques: a case study of Ramnagar town area, district Nainital, Uttarakhand, India India India Egypt J. Remote Sens. Space Sci.* ***16*** *111–7*

[2]Obando J A, Luwesi C N and Akombo R A 2024 The impact of land use and cover changes on river flows in Wundanyi catchment of Taita Hills, Kenya (1970–2030) Sustainable Social Development 2 2507 2024

[3]Thakur R and Manekar V L 2022 Artificial intelligence-based image classification techniques for hydrologic applications Appl. Artif. Intell. 36 1–19

[4]Rawat J and Kumar M 2015 Monitoring LULC changes using remote sensing and GIS techniques: a case study of Hawalbagh block, district Almora, Uttarakhand, India Egyptian J. Remote Sens. Space Sci. 18 77–84

[5]Zhang S Q, Zhang S K and Zhang J Y 2000 A study on wetland classification model of remote sensing in the Sangjiang plain Chinese Geographical Science 10 68–73

[6]Mehra N and Swain J B 2024 Assessment of land use land cover change and its effects using artificial neural network-based cellular automation Journal of Engineering and Applied Science 71 70

[7]Chen H et al. 2021 Changes of the spatial and temporal characteristics of land-use landscape patterns using multi-temporal landsat satellite data: a case study of Zhoushan Island, China China. Ocean & Coastal Management 213 105842

[8]Wang Z and Zhang M 2022 Evaluation and comparison of different machine learning models for NSAT retrieval from various multispectral satellite images Atmosphere (Basel) 13 1429

[9]Chirici G 2016 A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data Remote Sensing of Environment 176 282–94

[10]Ghosh R et al. 2014 Origin of visible and near-infrared photoluminescence from chemically etched Si nanowires decorated with arbitrarily shaped Si nanocrystals Nanotechnology 25 045703

[11]Rwanga S S and Ndambuki J M 2017 Accuracy assessment of land use/land cover classification using remote sensing and GIS International Journal of Geosciences 8 611–22