Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 1192752, 12 pages
https://doi.org/10.1155/2022/1192752

Research Article

Digital and Geographical Feature Detection by Machine Learning Techniques Using Google Earth Engine for CPEC Traffic Management

Sarah Mazhar,[1,2] **Guangmin Sun**,[1] **Anas Bilal**,[3] **Yu Li**,[1] **Muhammad Farhan**,[4] **and Hamid Hussain Awan**[5]

[1]*Faculty of Information Technology, Beijing University of Technology Beijing, China*
[2]*Faculty of Engineering and Computer Science, National University of Modern Languages, Islamabad, Pakistan*
[3]*College of Electronics and Information Engineering, Beibu Gulf University, Qinzhou, China*
[4]*Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Pakistan*
[5]*IBADAT International University Islamabad, Pakistan*

Correspondence should be addressed to Sarah Mazhar; sarah.mazhar@emails.bjut.edu.cn

Received 20 May 2022; Accepted 18 August 2022; Published 23 September 2022

Academic Editor: Hasan Ali Khattak

The center of human settlements is in the cities, which must have high-quality habitats for their inhabitants. Many megachallenges of urbanization, population development, global advancement, environmental destruction, traffic management, and climate change must be addressed. This study is aimed at understanding how to maintain balanced land development in rapidly urbanizing towns to solve this challenge and mobility issues. Climate and weather forecasts, land cover, environmental indices, nonoptical and optical wavelengths, water history, and air quality are only some of the datasets available on Google Earth Engine, a publicly usable data repository. Machine learning techniques, i.e., random forest (RF), support vector machine (SVM), and classification and regression tree (CART), are used to monitor spatial-temporal change regarding water, vegetation, and urbanization for Pakistan from 2013 to 2021 using Landsat 8. The detection of urban land suitability concerning multiple metrics such as ecological response variables, environmental tension, socio-economic development potential, and natural resource potential is also found. Dataset features were classified as bands in the Google Earth Engine. Moreover, for 2020 and 2021, classification results showing the change in water, vegetation, and urbanization are also represented concerning China Pakistan Economic Corridor (CPEC) highway and the railway track to monitor and control traffic and its management.

1. Introduction

Every year, urban land maps in the middle Yangtze River basin (MYRB) were facilitated by the Google Earth Engine from 1987 to 2017. After a manual topological dispute processing, random training samples for the current year, i.e., 2021, were created and submitted to GEE using modified OSM land-use info. The features of urban growth patterns, traces, and hotspots were investigated further. The resulting dataset is supposed to include explicit knowledge regarding MYRB's urban land distribution. According to the scientists, this proposed method can be extended and tested in other world areas to help better explain and measure different types of urban-related issues [1, 2].

The most recent literature on the Google Earth Engine found on PubMed is extracted, and word-cloud is drawn. The keywords are highlighted with assorted colors and sizes, as shown in Figure 1. The words with higher frequency are shown with bigger sizes and other colors.

AgKit4EE is one toolkit of Google Earth Engine designed to make using the Cropland Data Layer (CDL) product easier. The modeling (crop frequency, sequence, and trust layer) and

FIGURE 1: Word cloud of recent literature with document title and abstract.

land-use change analysis are only a few commonly utilized functions in the toolkit for use with CDL. The geospatial modelers working in agriculture and environment with CDL data benefit from the proposed program because it dramatically minimizes their workload [3].

Temporal consolidation is a technique that takes advantage of recent improvements concerning cloud processing capacity and free satellite data availability. Using Google Earth Engine, thirty-two datasets for Wales were developed by cloud mask automation and data aggregation in different timelines. Single-sensor datasets were outperformed by combined datasets, while spectral index-based datasets had the lowest accuracy. This research shows that temporal aggregation is a precise method for efficiently combining large volumes of images [4]. The inferior quality of cloud masking and automatic image selection can be compensated by temporal aggregation. It also proves that integrating data from various sensors can help increase classification accuracy [5]. For the precise comparison of processed image composites and manually chosen ones, the analysis emphasizes the need for determining appropriate satellite data combinations and parameters of aggregation [6]. The study was presented in the journal "Astronomy and Astrophysics."

In Google Earth Engine, as far as straightforward unmixing of optical Landsat imagery is concerned, TropWet allows multiple users to chart wetlands in vegetation-dominated areas. The findings indicate a substantial degree of precision in transmitting the conclusions throughout the African continent. The flood extent can map flood to the level of many formal market devices in Namibia. The impact of El Nino, i.e., Southern Oscillation occurrences, can also be calculated on the plants and swamplands in Southern Africa. It could include critical data to aid wetland management decisions at the state, provincial, and continental levels [7, 8].

Four considerations and fourteen requirements were chosen to perform the suitability study and land-use planning. The research was carried out in a specific region of Dhaka, one of the world's fastest expanding megacities. According to the research findings, extremely desirable land (13%) can be used for urban residential zones. A small amount of suitable land can be set aside for farming purposes and open fields, with 10% for conservation [9].

In the last thirty years, the cities of Srinagar and Jammu have seen alarming population development. This research is aimed at figuring out the suitability of the urban property for providing urban amenities. Land-use suitability evaluation is crucial for developing urbanization and a decision-making framework. Slope, height, land cover, and current comfort status are the variables included in the analysis. The research revealed the current trend regarding urban land use, the position of amenities, and the land suitability for potential urban amenities [10, 11].

Researchers in this literature merged land suitability modeling, remote sensing, landscape ecological research, and GIS to create a spatial analytical method for urban extension and land management. In Changsha Region, China, this tool defined constraints and expected benefits for potential land protection and growth. Environmental conservation, land use, and regional planning may all benefit from this approach [12].

Within the context of this article, the LeIGIS program enables analytical work. The model is built on the FAO land classification system for crops and data that defines an agricultural region regarding soil mechanics and environmental factors. When considering demand limitations, income maximization is part of economic research. The expert structure was created to assist in the land assessment and prepare for law changes. The established program will evaluate and display any comparable spatial dataset without specific programming skills [13].

Since the 1990s, land suitability assessment has advanced quickly in China, which has played an essential role in planning land use land cover (LULC). A modern assessment perspective was offered by moving along the implementation of

the landscape ecology philosophy. A geographic information system (GIS) to assess land suitability has become increasingly widespread. Land suitability assessment has been more scalable thanks to introducing evaluation templates into a GIS [14].

Primary purpose zoning has been hailed as a promising new spatial approach in China. This research is aimed at investigating the quantitative method using ecological economics perspectives. According to the study, identification requires thoroughly examining an interconnected regional ecosystem. The research could help to advance theory and the approach in China's cities and elsewhere [15, 16]. The process model and steps for feasibility study using GEE presented by [17] are shown in Figure 2.

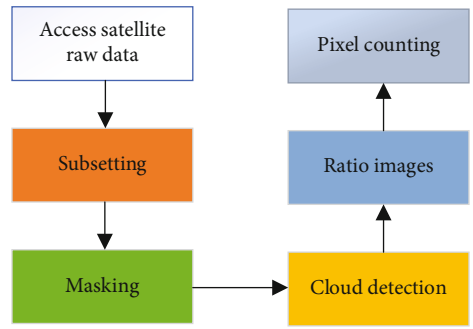

Figure 2: Feasibility method for elementary steps for data analysis using GEE.

Urban residential land demand grows yearly in most nations as the economy and technology progress. However, the world's rapidly urbanizing population has addressed significant concerns about urban residence quality, for example, air contamination, natural area divisions (e.g., wetlands, space, green, and open space), and traffic congestion [18, 19]. Urban development authorities face constraints and stress from the climate, community, and track at the social-economic stage [20]; leveraging urban residential property is complex. Thus, an accurate, fast, quantified, and fine-grained land suitability analysis for urban residents has become necessary for any planning department. It will further aid in formulating urbanization conditions and better understanding the urbanization method [21].

A land tract's suitability for certain users based on specific criteria, expectations, or predictors of these activities can be evaluated by land-use suitability analysis [22]. Residents do not like to reside in a chaotic, dirty, or dangerous setting. As a result, suitable urban residential land is figured out by various factors such as protection, comfort, and convenience. Precisely, safety demands that humanity may be protected from all-natural calamities and disasters like floods, storms, ice, snow, and other outside threats on any residential land. Comfort is the fact that people can perform day-to-day activities (eat, enjoy spare time, rest, sleep, exercise, and revive mental and physical strength completely) without restriction. Convenience refers to the ease of traveling for work, shopping, and visiting people using a facilitating transport system. As a result, the urban residential land suitability review identifies and locates the best possible land planning sites [23]. After the land use planning is completed, the aim is to find the most proper settlements for residential construction.

Concerning ecology or environment, sensitive areas are the ones that are critical for the long-time preservation of ecological diversity, water, soil, or other natural resources, both locally and regionally [24]. Examples are wildlife habitat zones, steep hills, lakes, and prime productive fields [25]. Protecting these areas is often seen as a critical parameter for assessing an ecosystem's stability and vulnerability [26].

Eurostat's guiding force-pressure-state-impact-response model motivated the idea of an environmental stress index (ESI). A European Environment Agency (EEA) metric provides a concise overview of the most significant human actions with harmful ecological consequences. The common topics for ESI research are climate change, biodiversity destruction, and air pollution [27]. The model was generalized to one indicator composed of two primary indexes based on the available emissions index and landscape deterioration index for this analysis. ESI is intended to augment stagnant ecological sensitivity research on the environmental structure and feature distribution by offering a more detailed view of the actual environmental condition as a product of historical urbanization and economic growth in dynamic contexts [28].

The World Bank and other organizations use social growth indices to assess the extent of development in countries and regions. A country's or city's current development level is a crucial determinant of competition for potential development from a growth standpoint. In today's global economy, a city's growth capacity is primarily defined by its ability to draw creative elements such as finance, knowledge, and professional knowledge, rather than conventional regional advantages such as property, energy, productivity, and other material wealth assets [29, 30].

The impact of natural capital on regional growth is no longer the same as it was in the past, thanks to rapid technological advancements and the ongoing restructuring of economic structures; this statement also extends, to some degree, to harmful elements [31]. On the other hand, inadequate natural resources limit an area's population and economic size, especially in dense urban areas.

Based on conditions and the economic value of urban land suitability, an intelligent land suitability model is developed economically by examining Google Earth Engine geo-environmental remote sensing datasets using machine learning techniques in this proposed research report. Due to the vast abundance of agricultural land and Pakistan's position as one of the world's largest nations, this research study will concentrate on urban land in Pakistan [32]. This research project is aimed at creating an intelligent machine learning-based model that figures out whether urban land is appropriate for construction based on ecological vulnerability, environmental stress, socio-economic development, and natural resource potential [33].

Checking was critical to creating the whole algorithms and fragments to see whether the requests exceeded the

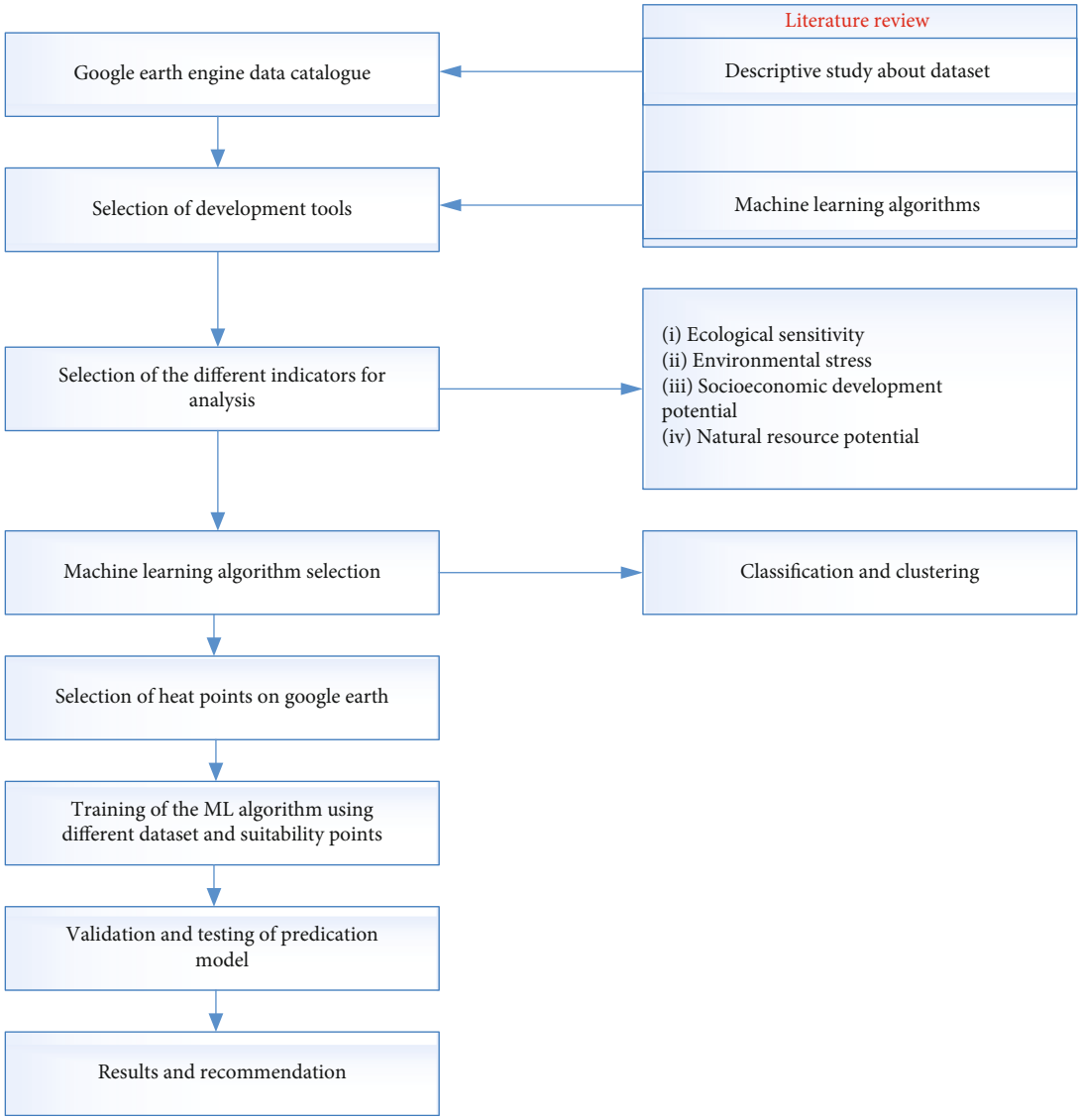

Figure 3: Methodology for development of geographical feature detection model for Pakistan.

usage cap on Google's servers [34]. As a first step, Google has set a limit of three queries per second, ensuring that resource-intensive algorithms do not degrade GEE's overall availability. Furthermore, client-side computations run after three minutes, while exports only receive the same error message after a more extended period: they reach the pre-compute limits after two hours of computation time or the on-the-fly limits after ten minutes per function.

This threshold is met depending on the scale of the original image collection and the imported feature collection. One can get around these limitations by adopting one's data and parameters. In particular, the scale parameter of regional counts of pixels decreases computing time almost immediately, albeit at the cost of precision and description. However, suppose one's algorithms produce exceptional and positive outcomes. In that case, one can request additional computing power, but this has only been granted for a few programs, such as "Global Forest Watch." It might be worthwhile to qualify for such a data use boost depending on the future implementations.

The suitability investigation of land use has undergone a shift from qualitative to quantitative with the help of GIS technologies [35], and one can see the "land-use suitability analysis" with a GIS base has become one of the most practical implementations of GIS [36]. GIS is a high-quality analysis platform that integrates various data forms to improve decision-making. A multicriteria assessment can rank each element in terms of significance and apply weights to each. As a result, the most popular approach for producing a final suitability map is to incorporate different methods into a GIS [37]. The contributions of the paper are as follows:

(i) To investigate using Google Earth Engine's global geo-environmental datasets to grow urban land suitability in Pakistan

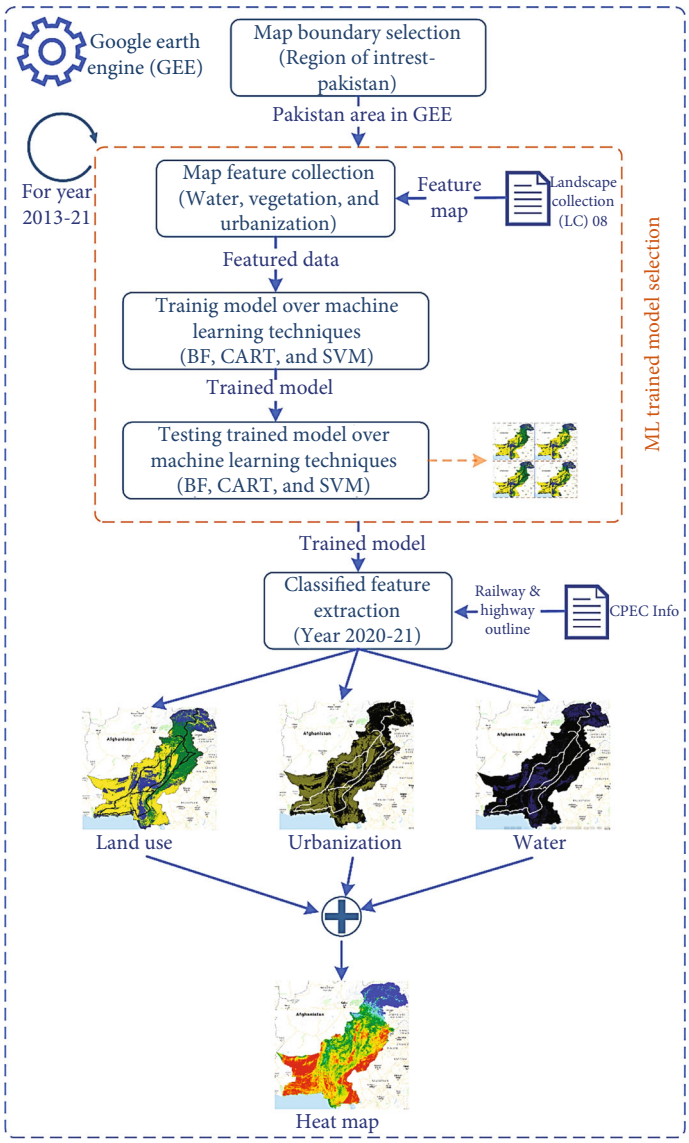

Figure 4: Pakistan geographical area coverage detection mechanism.

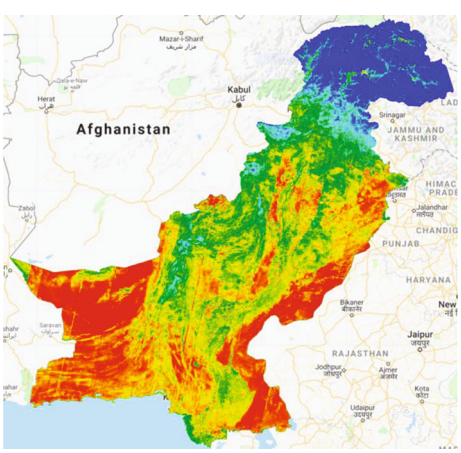

Figure 5: Pakistan as region of interest heat map.

(ii) Appropriate steps can be taken sooner to provide a facility for dealing with environmental stress, resulting in safer and more sustainable community growth

(iii) It is detecting the suitability of urban property early for management to make an informed judgment to improve socio-economic growth ability

(iv) To combine machine learning algorithms for detecting the suitability of urban land with a case analysis of Pakistani regions

2. Machine Learning Algorithms

2.1. Support Vector Machine. A support vector machine (SVM) is an algorithm that may be used for classification and regression. Consequently, it is often used in the

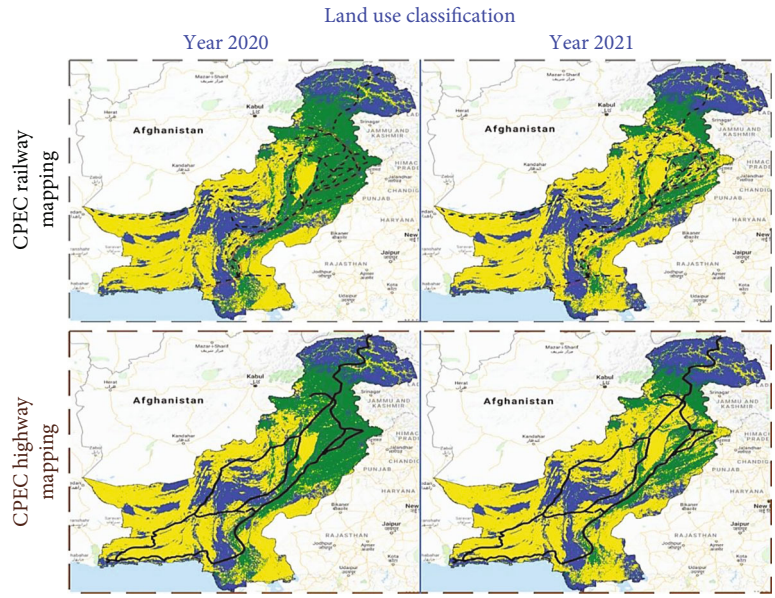

Figure 6: Land use classification for CPEC highway and railway mapping.

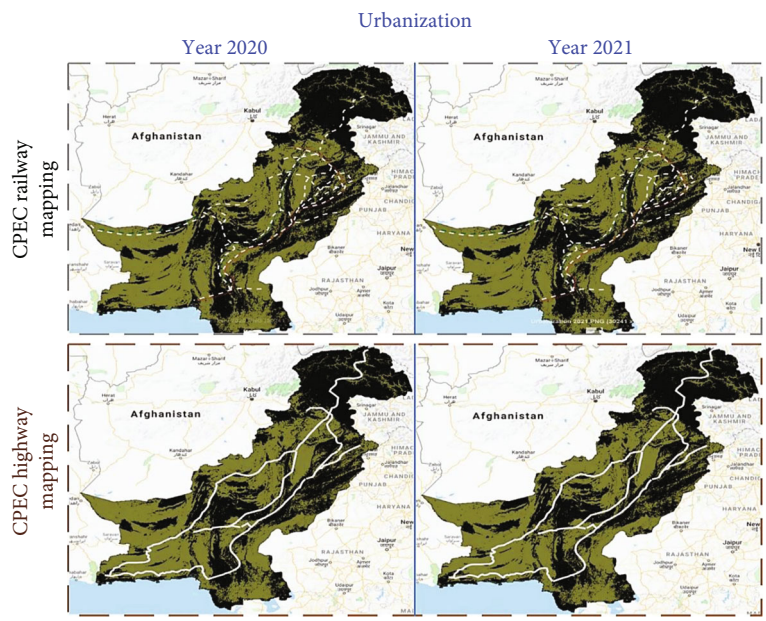

Figure 7: Urbanization classification for CPEC highway and railway mapping.

classification of objects. This technique is aimed at finding a hyperactive plane in N-dimensional space (N = the number of parameters) that divides the data points into distinct groups. Ninety percent of the entries are used for training, while ten percent are used for testing. It calculates the values by extracting the parameters from the test data. It obtains a favorable outcome, analyzes it using rigorous values, and provides the model's accuracy rate. Let the training samples have a dataset $\text{Data} = \{yi, xi\}\,; i = 1, 2, \cdots, n$ where $xi \in R\,n$ represents i the vector and $yi \in R\,n$ represents the target item. The optimal hyperplane of the form $f(x) = w^t x + b$ is found by the linear SVM, where w represents a dimensional coefficient vector and b represents an offset. It is carried out by solving subsequent optimization problems.

$$\text{Min}w, b\xi i \frac{1}{2} w^2 + C \sum_{i=1}^{n} \xi i, \tag{1}$$

$$\text{St.}yi\left(w^t xi + b\right) \geq 1 - \xi i, \xi i \geq 0, \forall i \in \{1, 2, \cdots, m\}. \tag{2}$$

2.2. Random Forest Algorithm. Random forest is an ensemble learning-based supervised machine learning technique. Multiple versions of the same algorithm are grouped in

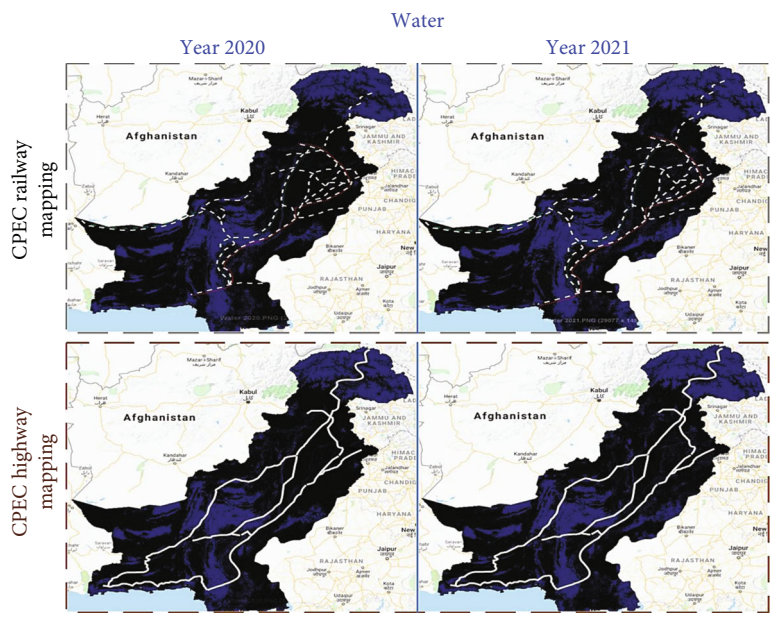

Figure 8: Water classification for CPEC highway and railway mapping.

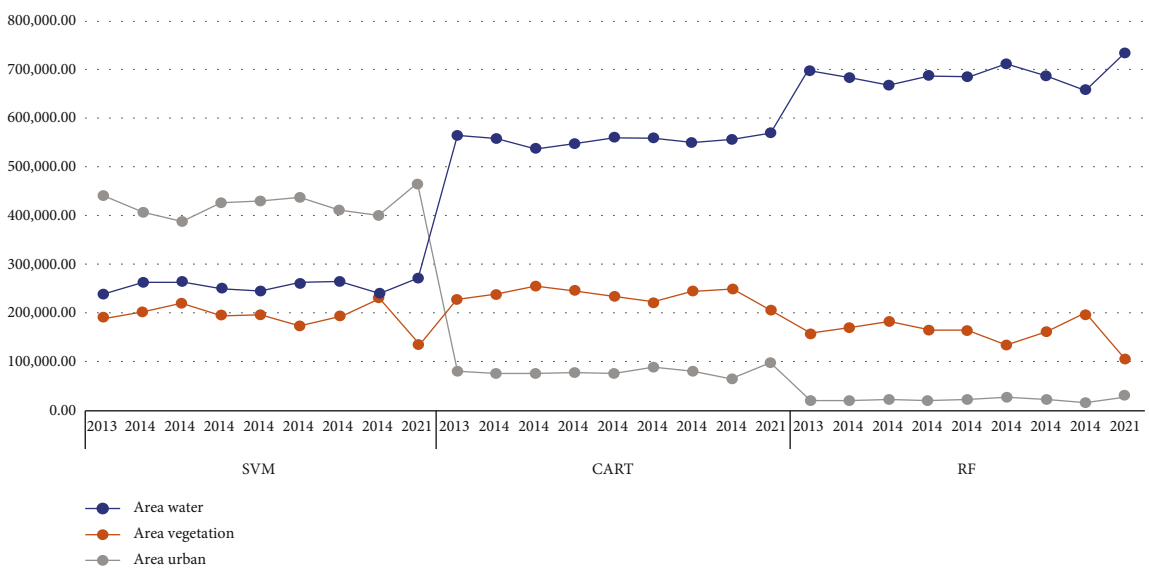

Figure 9: Area-wise coverage for ML techniques SVM, CART, and RF for the years 2013-21.

ensemble learning to develop a more effective prediction model. Many algorithms of the same sort as random forest, such as various decision trees, are combined to create a forest of trees. Regression and classification jobs can benefit from using the random forest approach. Random forest works faster than decision tree algorithms since it selects random values to predict the value. Several decision trees are built and incorporated by RF to get the best result. The bootstrap aggregating or bagging is applied for tree learning. For a given data, $X = \{x1, x2, x3, \cdots, xn\}$ time with responses $Y = \{x1, x2, x3, \cdots, xn\}$ repeats the bagging from $b = 1$ to B. The unseen samples $[x]'$ are made by averaging the predictions $\sum_{b=1}^{B} fb(x')$ from every individual tree on x':

$$j = \frac{1}{B}\sum_{b=1}^{B} fb\left(x'\right). \tag{3}$$

The standard deviation is used to calculate the uncertainty of a forecast on a tree.

$$\partial = \sqrt{\frac{\sum_{b=1}^{b}\left(fb\left(x'\right) - f\right)^2}{B-1}}. \tag{4}$$

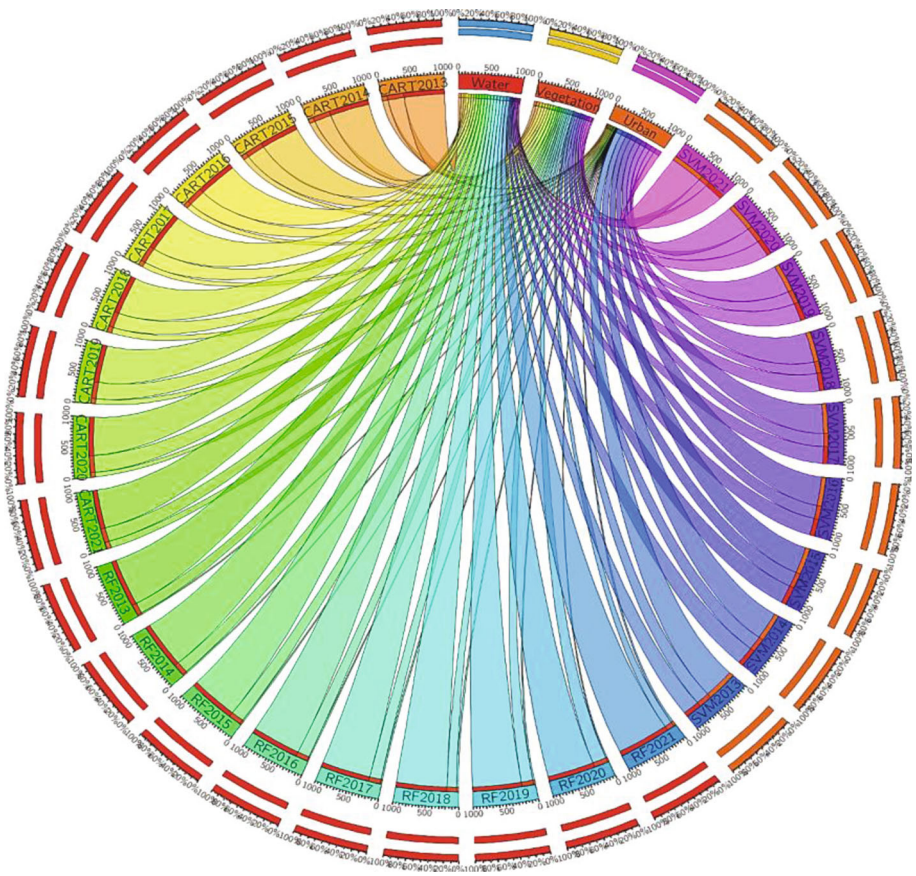

Figure 10: Values predictive distribution for SVM, CART, and RF for the years 2013 up to 2021.

2.3. Classification and Regression Trees. The Classification and Regression Tree (CART) algorithm is one of the supervised learning algorithms family. The primary purpose of employing a CART is to regressively predict some class or value of the target variable, i.e., training data. CART starts at the tree's root to forecast a class label for a record. In comparison to earlier approaches, this algorithm produced accurate and overprotective results. The advantage of this algorithm is that it does not reveal difficulties with overfitting. The trees are built by giving the high entropy input to sample data. Divide and conquer (DAC) approach is used to construct the fast and simple trees. Irrelevant samples are deleted on sample data D, called tree pruning.

$$\text{Entropy} = \sum_{j=1}^{m} pi_j \log_2 pi_j. \tag{5}$$

3. Materials and Methods

3.1. Datasets. The Landsat satellite series provide the longest continuous record of satellite-based observations. Landsat is an essential resource used in policymaking to track world change and provide medium spatial resolution Earth observations. Landsat-8 receives up to 740 images per day, compared to before in 2013, i.e., 550 images per day [38].

With sun elevations over five degrees, images are defined as day-lit images. There is a criterion set for the inclusion of images. Scheduling of all candidate scenes is done if less than 740 images per day exist. The images are excluded as a cloud cover prediction function and long-term cloud cover statistics dealing with more than 740 candidate scenes. The priority is increased if the cloud cover prediction is better than the long-term average. In case of rejection of an acquisition, a missed opportunity priority increases the probability of future acquisitions [39].

Figure 3 illustrates the candidate scene distribution during 2019. The green region is the area containing acquired scenes. The orange region depicts the area that did not meet the day-lit criteria. The regions with yellow color are particular request images that represent the rejected ones because of cloud cover thresholds. Due to resource reservations, images in the areas under the blue regions cannot be acquired, for example, maneuvers or calibration activities. The black horizontal line represents the 740-image daily limit. Due to the daily limit, scenes in the red region were rejected. Over Antarctica, the most abandoned locations were observed where revisit times of up to once every two days were allowed by the side lap between paths [39].

The Landsat 8 completes the Earth's orbit in a sun-synchronous near-polar orbit at an altitude of 705 km. It is inclined at 98.2 degrees and completes one Earth orbit every 99 minutes. Data collected by the spacecraft and sensor

Table 1: Biyearly differences in water, vegetation, and urban areas with its consecutive year.

Classifier	Biyearly difference (year 1–year 2)	Area (KM)^2 Water	Vegetation	Urban
SVM	2013-2014	-23,760.09	-10,935.41	34,695.50
	2014-2015	-2,056.47	-17,522.78	19,579.24
	2015-2016	15,504.99	24,847.01	-40,351.99
	2016-2017	3,960.95	-1,297.32	-2,663.64
	2017-2018	-16,245.27	24,239.77	-7,994.50
	2018-2019	-2,412.10	-23,344.04	25,756.13
	2019-2020	22,644.64	-35,654.68	13,010.04
	2020-2021	-31,118.53	97,317.19	-66,198.65
CART	2013-2014	6,307.90	-8,836.35	2,528.45
	2014-2015	20,306.58	-18,234.71	-2,071.88
	2015-2016	-9,347.49	10,103.24	-755.74
	2016-2017	-13,328.13	11,438.08	1,890.05
	2017-2018	1,872.48	11,165.00	-13,037.48
	2018-2019	10,055.88	-20,503.38	10,447.51
	2019-2020	-7,094.91	-7,064.69	14,159.59
	2020-2021	-14,266.01	46,175.12	-31,909.11
RF	2013-2014	13,088.99	-14,039.76	950.77
	2014-2015	14,528.61	-12,383.93	-2,144.68
	2015-2016	-20,342.96	18,263.19	2,079.77
	2016-2017	1,965.17	1,868.30	-3,833.47
	2017-2018	-26,336.40	29,401.03	-3,064.64
	2018-2019	24,048.19	-27,710.51	3,662.32
	2019-2020	29,732.99	-35,058.66	5,325.68
	2020-2021	-76,089.03	89,523.10	-13,434.08

Table 2: Biyearly differences starting from 2013 to 2021 for the classifiers and their prediction of water, vegetation, and urban areas in KM^2.

Classifier	Yearly difference from base year to each year	Area (KM)^2 Water	Vegetation	Urban
SVM	2013-2014	-23,760.09	-10,935.41	34,695.50
	2013-2015	-25,816.56	-28,458.18	54,274.74
	2013-2016	-10,311.57	-3,611.17	13,922.75
	2013-2017	-6,350.62	-4,908.50	11,259.11
	2013-2018	-22,595.89	19,331.27	3,264.62
	2013-2019	-25,007.99	-4,012.76	29,020.75
	2013-2020	-2,363.35	-39,667.44	42,030.79
	2013-2021	-33,481.88	57,649.75	-24,167.86
CART	2013-2014	6,307.90	-8,836.35	2,528.45
	2013-2015	26,614.48	-27,071.06	456.57
	2013-2016	17,266.99	-16,967.82	-299.17
	2013-2017	3,938.86	-5,529.74	1,590.88
	2013-2018	5,811.34	5,635.25	-11,446.60
	2013-2019	15,867.22	-14,868.13	-999.09
	2013-2020	8,772.31	-21,932.82	13,160.50
	2013-2021	-5,493.70	24,242.30	-18,748.61
RF	2013-2014	13,088.99	-14,039.76	950.77
	2013-2015	27,617.60	-26,423.69	-1,193.91
	2013-2016	7,274.64	-8,160.50	885.86
	2013-2017	9,239.81	-6,292.20	-2,947.61
	2013-2018	-17,096.59	23,108.83	-6,012.25
	2013-2019	6,951.60	-4,601.68	-2,349.93
	2013-2020	36,684.59	-39,660.34	2,975.75
	2013-2021	-39,404.44	49,862.76	-10,458.33

payload correction data, elevation data provided by a digital elevation model, and the processing level used are determined by ground control points [40]. It has the following parameters, i.e., north-up MAP image orientation, world geodetic system 84 as the datum, universal transverse Mercator as map projection, 30-meter and 60-meter pixel size as reflective bands, cubic convolution as resampling method, and GeoTIFF as output format of images.

4. Methodological Framework

A detailed literature review and descriptive analysis are undertaken on Google Earth Engine geo-environmental datasets in the proposed research phase, as seen in Figure 3. Later in the research phase, the informative study of the dataset's characteristics or attributes is defined for interpretation purposes. The next step was a literature review and descriptive analysis of various machine learning algorithms to obtain their efficiency and applicability to multiple data types. The selection of machine learning algorithms and techniques is based on the dataset characteristics and a detailed analysis performed earlier.

After projecting the respective dataset layer on the selected land boundary, this analytical study obtains random training samples. Training samples can be obtained in dark and clear conditions, and gathered pieces have come from different years. Variation in training samples allows algorithms to learn more about features and training data, thus improving the efficiency or accuracy of machine learning algorithms. Besides, Google Earth Engine offers a cloud-based computing facility for analyzing petabytes of data. After a detailed examination of geo-environmental datasets and machine learning algorithms, selected algorithms are applied to training datasets, yielding land suitability detection models.

Based on conditions and the economic value of urban land suitability, an intelligent land suitability model is developed economically by examining Google Earth Engine geo-environmental remote sensing datasets using machine learning techniques. Due to the vast abundance of agricultural land and the fact that China is one of the largest countries, this research study will concentrate on urban land in China. This research is aimed at creating an intelligent machine learning-based model that detects whether a decision can be made based on ecological sensitivity, environmental stress, socio-economic growth potential, and natural resource potential while considering urban land for development activities.

Implementing an automated thresholding function enables the classification of heat points that provide good results even under the changing conditions of the resulting image composite. Although it displays a tool that should be further refined by using correct reflectance values for topographic effects, it offers the possibility of computing a fair estimate of values over large regions. Furthermore, the objective to complement already available heat sources with a mask from debris-covered areas was achieved by creating an image composite for the entire regions in Pakistan. Despite the dependency on the quality of the prior scene selection, the results provide good insights for a better understanding of ablation processes and their implications within a specific region.

4.1. Pakistan Geographical Area Coverage Detection Mechanism. In the method, Google Earth Engine is used for training and feature extraction for region of interest (ROI) from the Earth map, as shown in Figure 4. Firstly, it is necessary to define the boundary for the region to collect and apply different image processing features while extracting values like heat, water, urbanization, vegetation, and many more, whereas library LSIB published in 2017 provides country-level boundary selection. The following code snippet helps in the sample for Pakistan as the region of interest in our particular case study: var roi = ee.FeatureCollection('USDOS/LSIB_SIMPLE/2017').filterMetadata('country_co', 'equals', 'PK');

Now, data points for training machine using Google Map feature selection with pinning mechanism while putting them in different classes like water, vegetation, and urbanization. Landscape Collection 08 (LC 08) is used for Earth map feature collection, which provides "dplease ata" values for 2013 to 2021. So, the developed mechanism trains and tests three different machine learning techniques like BF, CART, and SVM to observe which model best fits accuracy and precision while detecting map features like water, vegetation, urbanization, and land use.

4.2. Importance of Urbanization, Water, and Vegetation Detection for Heat Analysis. After the final selection of the ML technique as a trained model used to find out values of the map for the Pakistan region to depict change in water, land use, and urbanization, all predicted features contribute to displaying the heat map for a specific area. The above mechanism is represented in Figure 5.

5. Results and Discussion

In the first phase, a scene classification algorithm was created to choose the best images for glacier mapping with the least amount of cloud cover and seasonal snow. Different approaches have been implemented and compared to a manually performed scene selection in terms of scene selection precision and estimation period. Even though GEE has certain limitations in terms of usability, such as spatial or temporal resolution and height, the automated scene selection algorithm performs well enough to present a compilation of the most suitable images for further processing and is scalable. The second algorithm computes glacier outlines and measures regions based on an image composite of an earlier scene selection. Glacier outlines have been beneficial results that may be utilized on a large scale. On the other hand, future applications may fix shadows inside individual scenes, detect debris-covered areas, highlight sources of error, and lower impact quality.

5.1. Analyzing Map Area Detection Using Different Machine Learning Techniques. For 2020 and 2021, the Pakistan region's land use, urbanization, and water are depicted using the SVM machine learning technique, as shown in Figures 6, 7, and 8, respectively. Furthermore, these maps are masked for standing for CPEC (a collective project between Pakistan and China) railway and highway plan over feature extracted. These maps show the impact over the entire year, especially separately for each year, 2020 and 2021.

Figure 6 shows the overall impact as the water and vegetation region reduced for one year between 2020 and 2021. Figure 6 shows that the urbanization area increased significantly, showing the future challenges and their impact while completing the project.

5.2. Facts Over Figures. The accuracy level of RF and CART is higher than SVM due to the overfitting factor of water feature detection over the map, as clearly seen in Figure 9.

This overfitting reflects the wrong prediction of low data values for urbanization and vegetation from 2013 to 2021. For example, for the year 2021, RF predicted a water area of 736,065.81 square kilometers out of the total area of Pakistan, which is 881,913 square kilometers coming up to unrealistic by making a false prediction of 83% area filled with water. Prediction results assure that SVM brings much better results than ML techniques like CART and RF. In Figure 10, the distribution of data values shows that water dominates over other features like vegetation and urbanization when using ML techniques like RF and CART.

The calculated differences in water, vegetation, and urban areas with their consecutive years are shown in Table 1. The starting year for the difference calculation is 2013, and the ending year is 2021. The overall vegetation trend increases while water and urban areas decrease. The calculated trend is more or less the same across all the classifiers. Some years' differences are significant, e.g., the metropolitan areas' SVM in 2020-2021.

Three machine learning (ML) models are used to predict the values of water, vegetation, and urban. These ML models are SVM, CART, and RF. All these models predict the years 2013 to 2021. Hence, the last year of experiments is 2021, and these are performed the year before the completion of the year. It is expected that the values for the year 2021 may change if these are performed after the completion of the year 2021. The total number of experimental years is 9, so each model predicts these. Each model has a separate entry for each year, i.e., RF2021, which means the prediction of the RF model for the year 2021. Each fragment of the diagram is divided into 1000 subsections, and the out part of the diagram shows percentage-wise division, where each

central axis stands for the 20% interval as shown in Figure 10.

The calculated differences in water, vegetation, and urban areas predicted by different classifiers are shown in Table 2. The starting year is 2013, and the ending year is 2021. The overall vegetation trend increases while water and urban areas decrease. The trend is more or less the same across all the classifiers.

6. Conclusion and Future Work

The independent identification of temperature variations requires the study of glaciers and their changes overtime. In addition to their hydrologic importance on a regional to the global scale, they can trigger or incite natural disasters, needing a monitoring system that allows for routine observations. Since satellite imagery encompasses such a large globe region, it is possible to track glacier change with high spatial and temporal precision. These resolutions have been pushed to new heights, resulting in an ever-growing publicly accessible satellite data archive—the advantages of the volume of data used for glacier-monitoring decrease manual workload and device loading times. The Google Earth Engine (GEE) can map, measure, and imagine glacier distribution, a realistic way to cope with the increasing workload.

Data Availability

The data supporting this study's findings are available from the corresponding author or Sarah Mazhar upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 11527801 and grant 41706201.

References

[1] D. Liu, N. Chen, X. Zhang, C. Wang, and W. Du, "Annual large-scale urban land mapping based on Landsat time series in Google Earth Engine and OpenStreetMap data: a case study in the middle Yangtze River basin," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 337–351, 2020.

[2] N. A. Wahap and H. Z. M. Shafri, "Utilization of Google Earth Engine (GEE) for land cover monitoring over Klang Valley, Malaysia," *IOP Conference Series: Earth and Environmental Science*, vol. 540, no. 1, 2020.

[3] C. Zhang, L. Di, Z. Yang, L. Lin, and P. Hao, "AgKit4EE: a toolkit for agricultural land use modeling of the conterminous United States based on Google Earth Engine," *Environmental Modelling and Software*, vol. 129, article 104694, 2020.

[4] S. Das, P. K. Shit, and P. P. Patel, "Ecosystem services value assessment and forecasting using integrated machine learning algorithm and CA-Markov model: an empirical investigation of an Asian Megacity," *Geocarto International*, pp. 1–23, 2021.

[5] F. Ullah, J. Wang, M. Farhan, S. Jabbar, M. K. Naseer, and M. Asif, "LSA based smart assessment methodology for SDN infrastructure in IoT environment," *International Journal of Parallel Programming*, vol. 48, no. 2, pp. 162–177, 2020.

[6] K. Razzaq Malik, T. Ahmad, M. Farhan, F. Ullah, K. Amjad, and S. Khalid, "Multiagent semantical annotation enhancement model for IoT-based energy-aware data," *International Journal of Distributed Sensor Networks*, vol. 12, no. 6, Article ID 9103265, 2016.

[7] A. Hardy, G. Oakes, and G. Ettritch, "Tropical wetland (TropWet) mapping tool: the automatic detection of open and vegetated waterbodies in Google Earth engine for tropical wetlands," *Remote Sensing*, vol. 12, no. 7, 2020.

[8] H. Ibrahim, Z. Khattab, T. Khattab, and R. Abraham, "Expatriates' housing dispersal outlook in a rapidly developing metropolis based on urban growth predicted using a machine learning algorithm," *Housing Policy Debate*, pp. 1–21, 2021.

[9] K. M. Ullah and A. Mansourian, "Evaluation of land suitability for urban land-use planning: case study Dhaka city," *Transactions in GIS*, vol. 20, no. 1, pp. 20–37, 2016.

[10] J. A. Parry, S. A. Ganaie, and M. Sultan Bhat, "GIS based land suitability analysis using AHP model for urban services planning in Srinagar and Jammu urban centers of J&K, India," *Journal of Urban Management*, vol. 7, no. 2, 2018.

[11] P. Kulithalai Shiyam Sundar and P. C. Deka, "Spatio-temporal classification and prediction of land use and land cover change for the Vembanad Lake system, Kerala: a machine learning approach," *Environmental Science and Pollution Research*, pp. 1–7, 2021.

[12] F. Yang, G. Zeng, C. Du, L. Tang, J. Zhou, and Z. Li, "Spatial analyzing system for urban land-use management based on GIS and multi-criteria assessment modeling," *Progress in Natural Science*, vol. 18, no. 10, 2008.

[13] S. Kalogirou, "Expert systems and GIS: an application of land suitability evaluation," *Computers, Environment and Urban Systems*, vol. 26, no. 2–3, 2002.

[14] T. G. Shi, G. Q. Zheng, Z. Y. Wang, and L. L. Wang, "Progress in research on land suitability evaluation in China," *Progress in Geography*, vol. 2, no. 2, pp. 106–115, 2007.

[15] J.-L. He, Y.-G. Zong, and H. Gebhardt, "New spatial strategy in China: major function zoning of Beijing-Tianjin metropolitan area with perspective from ecological economics," *Journal of Urban Planning and Development*, vol. 137, no. 4, 2011.

[16] D. Hou, F. Meng, and A. V. Prishchepov, "How is urbanization shaping agricultural land-use? Unraveling the nexus between farmland abandonment and urbanization in China," *Landscape and Urban Planning*, vol. 214, article 104170, 2021.

[17] N. Zeltner, *Using the Google Earth Engine for Global Glacier Change Assessment, [Ph.D. Thesis]*, Geographisches Institut der Universität Zürich, 2016.

[18] M. L. Zellner, T. L. Theis, A. T. Karunanithi, A. S. Garmestani, and H. Cabezas, "A new framework for urban sustainability assessments: linking complexity, information and policy," *Computers, Environment and Urban Systems*, vol. 32, no. 6, pp. 474–488, 2008.

[19] B. Mondal and D. N. Das, "How residential compactness and attractiveness can be shaped by environmental amenities in an industrial city?," *Sustainable Cities and Society*, vol. 41, pp. 363–377, 2018.

[20] Z. Xu and Q. Li, "Integrating the empirical models of benchmark land price and GIS technology for sustainability analysis of urban residential development," *Habitat International*, vol. 44, pp. 79–92, 2014.

[21] J. Luo, P. Du, A. Samat, and L. Feng, "Evaluation on the natural suitability of urban human settlement environment using multisource data," in *2015 Joint Urban Remote Sensing Event (JURSE)*, Lausanne, Switzerland, 2015.

[22] M. G. Collins, F. R. Steiner, and M. J. Rushman, "Land-use suitability analysis in the United States: historical development and promising technological achievements," *Environmental Management*, vol. 28, no. 5, 2001.

[23] M. Zhang, M. Zhang, H. Yang, Y. Jin, X. Zhang, and H. Liu, "Mapping regional soil organic matter based on sentinel-2a and modis imagery using machine learning algorithms and google earth engine," *Remote Sensing*, vol. 13, no. 15, 2021.

[24] S. K. Tayyaba, H. A. Khattak, A. Almogren et al., "5G vehicular network resource management for improving radio access through machine learning," *IEEE Access*, vol. 8, pp. 6792–6800, 2020.

[25] F. Ndubisi, T. DeMeo, and N. D. Ditto, "Environmentally sensitive areas: a template for developing greenway corridors," *Landscape and Urban Planning*, vol. 33, no. 1–3, 1995.

[26] M. A. Al-Garadi, M. R. Hussain, N. Khan et al., "Predicting cyberbullying on social media in the big data era using machine learning algorithms: review of literature and open challenges," *IEEE Access*, vol. 7, pp. 70701–70718, 2019.

[27] A. Shankar Kshirsagar, M. A. El-Gafy, and T. Sami Abdelhamid, "Suitability of life cycle cost analysis (LCCA) as asset management tools for institutional buildings," *Journal of Facilities Management*, vol. 8, no. 3, 2010.

[28] H. Shahid, M. A. Shah, A. Almogren et al., "Machine learning-based mist computing enabled Internet of Battlefield Things," *ACM Transactions on Internet Technology (TOIT)*, vol. 21, no. 4, pp. 1–26, 2021.

[29] Y. Zhang, R. Song, K. Zhang, and T. Wang, "The Characteristics and Modes of Urban Network Evolution in the Yangtze River Delta in China from 1990 to 2017," *IEEE Access*, vol. 9, pp. 5531–5544, 2021.

[30] N. S. Sumari, P. B. Cobbinah, F. Ujoh, and G. Xu, "On the absurdity of rapid urbanization: spatio-temporal analysis of land-use changes in Morogoro, Tanzania," *Cities*, vol. 107, article 102876, 2020.

[31] R. M. Auty, "The political economy of resource-driven growth," *European Economic Review*, vol. 45, no. 4–6, 2001.

[32] M. F. Baqa, F. Chen, L. Lu et al., "Monitoring and modeling the patterns and trends of urban growth using urban sprawl matrix and CA-Markov model: a case study of Karachi, Pakistan," *Land*, vol. 10, no. 7, 2021.

[33] F. S. Ahmad, L. Ali, Raza-Ul-Mustafa et al., "A hybrid machine learning framework to predict mortality in paralytic ileus patients using electronic health records (EHRs)," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3283–3293, 2021.

[34] F. Ullah, J. Wang, M. Farhan, S. Jabbar, Z. Wu, and S. Khalid, "Plagiarism detection in students' programming assignments based on semantics: multimedia e-learning based smart assessment methodology," *Multimedia Tools and Applications*, vol. 79, no. 13-14, pp. 8581–8598, 2020.

[35] J. Yang, Y. Liu, and S. Wang, "An overview of the methods of GIS-based land-use suitability analysis," *Geoinformatics 2007: Geospatial Information Technology and Applications*, vol. 6754, pp. 1110–1120, 2007.

[36] J. Malczewski, "GIS-based land-use suitability analysis: a critical overview," *Progress in Planning*, vol. 62, no. 1, 2004.

[37] H. Huang, Q. Li, and Y. Zhang, "Urban residential land suitability analysis combining remote sensing and social sensing data: a case study in Beijing, China," *Sustainability (Switzerland)*, vol. 11, no. 8, 2019.

[38] "Landsat acquisition, Landsat missions," https://landsat.usgs.gov/landsat_acq.

[39] "Landsat Landsat Data Access," https://www.usgs.gov/landsat-missions/landsat-data-access.

[40] "Science for a changing world," https://www.usgs.gov/.