Exploring Factors Affecting Land Surface Temperature Using Machine Learning Technique: A Case Study on Meghalayan Terrain

Puja Saha[1], Teddyson Marbaniang[2], and Amitabha Nath[3*]
[1,2,3]Department of Information, Technology
North-Eastern Hill University Meghalaya, India
E-Mail: pooja12saha@gmail.com, teddysonmarbz@gmail.com, *amitabha.me@gmail.com
*Corresponding Author

Abstract—Earth's surface temperature continues to rise steadily over time. A Long-term assessment of Earth's Land Surface Temperature (LST) trend and changes in land use/land cover (LULC) can help understand its cause and effect on the environment. This study aims to use LST trend and LULC change information to analyze its impact on the state of Meghalaya. As an indicator, the trend in LST is used to examine the effects of natural and human-induced activities, supported by the decadal land use change assessment. The study utilizes a Mann-Kendall test to identify the trend in LST from the last 22 years (2000–2021). Similarly, a machine learning technique (Random Forest) is used to analyze the dynamics of decadal LULC change. Inter class differences in conjunction with LST trends are then cross-validated to draw any conclusion. The test suggests an increasing trend in LST for the Meghalayan region. Corresponding analysis of LULC change suggests that the conversion of vegetation land to buildup areas rose from 28% in 2000-2010 to 45% in 2010-2021. Similarly, 20% of forest land was converted to buildup areas in the first half, with a slight decrease to 14% in the second half. Furthermore, the importance of temperature, rainfall, and DEM in modelling LST variations was also outlined as a part of the parameter importance analysis performed in this study. By utilizing these insights and methodology, effective measures can be developed to counter the harmful effects of deforestation and urbanization, thus preserving the ecological balance of our ecosystems for future generations.

Keywords: Land surface temperature, LULC change, Mann-Kendall test, environmental impact, Meghalaya.

INTRODUCTION

Increasing Earth's surface temperature has become a global issue. It is estimated that the Earth's surface temperature is expected to increase by ~0.2°C per decade with an additional increase of 1.8°C to 4.0°C by the end of this century (Staszewski, 2023; Zheng *et al.*, 2022).This can cause significant environmental stress, driven by population growth, unplanned urbanization, industrialization, and deforestation. With the rising global population, the demand for resources will also increase. This will ultimately lead to the degradation of natural resources and put additional pressure on our ecosystem (Gondwe *et al.*, 2021; Hossain *et al.*, 2023; Sajan *et al.*, 2023).

This article presents a case study of Meghalaya, a north-eastern state of India. The region's extensive forest cover is a sanctuary for numerous endangered species, making it a vital ecological hotspot. However, the state has undergone notable changes in its LULC forms, particularly in its rich vegetation and forested areas (Singh *et al.*, 2020). Human encroachment of forest areas poses a grave threat, resulting

https://doi.org/10.22232/stj.2025.13.01.16

in the loss of invaluable forest resources and habitat degradation. These transformations have significantly impacted the region's ecological balance, exerting deep effects on the livelihoods of nearly 81% of the total population who rely heavily on forests for sustenance and cultural practices.

The literature survey reveals it's a fact that LULC and LST have a strong correlation and temperature rise has often been subjugated to transformation from natural vegetated areas into the built-up class (Moisa *et al.*, 2022). There are multiple studies specifically done to investigate LST and LULC changes for observing variations in local temperature for a given region (Ezimand *et al.*, 2024; Vohra *et al.*, 2024). Existing studies primarily used remote sensing imageries to analyze changes in surface temperature by utilizing LULC and Thermal Field Variance Index (TFVI) maps. In addition, researchers have extensively explored the use of various land cover spectral indices, such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Bareness Index (NDBaI), and the Normalized Difference Buildup Index (NDBI) in delineating different LULC classes.

Human activities, such as burning fossil fuels, have released greenhouse gases that can trap the sun's heat and cause global warming. Consequently, nearly all land areas are experiencing increasingly hotter days. For instance, 2020 ranked among the hottest years ever recorded. LST serves as an important indicator of ecosystem health. Elevated temperature creates stress on the ecosystems (e.g., droughts and heat waves) (Zhang *et al.*, 2023). It can worsen environmental conditions, reducing biodiversity, causing ecosystem degradation, and increasing vulnerability to disturbances (Rogers-Bennett and Catton, 2022). Moreover, surface temperature influences various ecosystem processes, including photosynthesis, evapotranspiration, and nutrient cycling. Temperature changes can disrupt these processes, affecting ecosystem productivity, water availability, and nutrient dynamics (Duan *et al.*, 2022). By tracking surface temperature, scientists can assess how ecosystems respond to changing environmental conditions and predict potential ecological patterns and function shifts. Also, LST monitoring enables the implementation of adaptive management strategies to mitigate potential impacts and promote ecosystem resilience. Factors that affect LST are equally important. Multiple studies have shown that the Digital elevation model (DEM) and the normalized difference vegetation index (NDVI) have a strong correlation with surface temperature (Feng *et al.*, 2019; Ham *et al.*, 2022). In addition, the Normalized Difference Built-Up Index (NDBI) is also found to be positively correlated with LST (Saha *et al.*, 2024). In this study, we have considered temperature, rainfall, DEM, u-component of wind, v-component of wind, pressure, and NDVI to predict the LST of Meghalaya.

Likewise, the Land Use Land Cover (LULC) also plays a crucial role in understanding various aspects of our environment. Studies have shown that LST variations can be determined by analyzing the LULC change (Guha, 2021; Keerthi Naidu and Chundeli, 2023; Sajan *et al.*, 2023). LULC maps, derived from satellite imagery and ground surveys are some of the credible sources to provide valuable insights into the patterns of LST and land use changes. By analyzing LULC maps, we can identify signs of deforestation, urbanization, and agricultural expansion in a region. This information is crucial for understanding the influence of human activities on the environment and develop strategies for sustainable land management. Furthermore, LULC maps offer critical insights into vegetation health, changing patterns, and urbanization trends. They help identify areas of concern, such as degraded ecosystems or expanding urban areas, allowing policymakers to prioritize conservation and land-use planning initiatives.

Despite this issue's critical importance, only a few number of studies have been carried out on Meghalaya's persistent LULC transformation processes and the factors driving these changes (Bhuyan *et al.*, 2023; Lamare *et al.*, 2019; Prokop, 2020). Moreover, these studies did not consider LST in consideration, which is a direct indicator of global warming. In this study, we wish to develop a comprehensive framework by incorporating LST and LULC change to address this problem. It will help us to understand the kind of environmental stress the region is facing and the factors responsible for bringing this change.

Recent advancements in RS/GIS technology have made it easier to analyze and monitor environmental changes (Chughtai *et al.*, 2021; Kanagasundaram *et al.*, 2022). These powerful tools provide capabilities for processing and interpreting geospatial information (Kaur *et al.*, 2023). Additionally, the availability of geospatial repositories further enhances accessibility to a vast array of satellite imagery and other geospatial datasets, facilitating research, decision-making, and innovation across various domains. Many recent studies have used Google Earth Engine (GEE) to perform complex geospatial analyses without the need for extensive computational resources or specialized software (Sajan *et al.*, 2023). By leveraging Google's infrastructure, users can process large-scale environmental data efficiently and extract valuable insights into land cover, land use, vegetation dynamics, climate change, and more.

Based on our literature survey, we hypothesize that the increase in land surface temperature is driven by changes in Land Use and Land Cover (LULC), subsequently affecting biodiversity. To validate this hypothesis, we have outlined the following objectives:

- **Objective 1:** To analyze the changes in LULC practices for the state of Meghalaya using high-satellite imagery from the last22 years (2000–2021).
- **Objective 2:** To analyze the trend in LST from 2000 to 2021 and develop a model to capture the relationship between LST and its affecting factors. Subsequently, determine the relative importance of each constituent attribute.

The manuscript is organized into five sections to enhance readability. Section 2 provides an overview of the study area and the methodologies used in the data acquisition, classification, prediction, evaluation, and trend analysis. In Section 3, the results obtained from these various methods are presented. Section 4 delves into a detailed discussion of the findings and their broader significance. Finally, Section 5 encapsulates the key insights and implications of the study as the concluding remarks.

MATERIALS AND METHOD

Study Area

We have selected the state of Meghalaya as a case study because of recent developments in the name of industrialization and other expansion activities (Chakraborty and Saikia, 2022). Moreover, as per the latest Sustainable Development Goal (SDG) India Index 2023-2024 report, the state finds itself among the least performing states and union territories. The study area extends from latitude 24°58’ to 26°07’N and longitude 89°48’ to 92°51’E, as depicted in Fig. 1. Meghalaya ranks among the wettest regions globally, with high annual rainfall attributed to its steep terrain and proximity to the Bay of Bengal. Cherrapunji and Mawsynram in Meghalaya hold the record for the highest yearly rainfall in the world (Bhuyan *et al.*, 2023; Chakraborty and Saikia, 2022). The temperature fluctuates across the region, with the central highlands experiencing 2 to 24 °C and the western half ranging from 12 to 33 °C (Suri *et al.*, 2023).

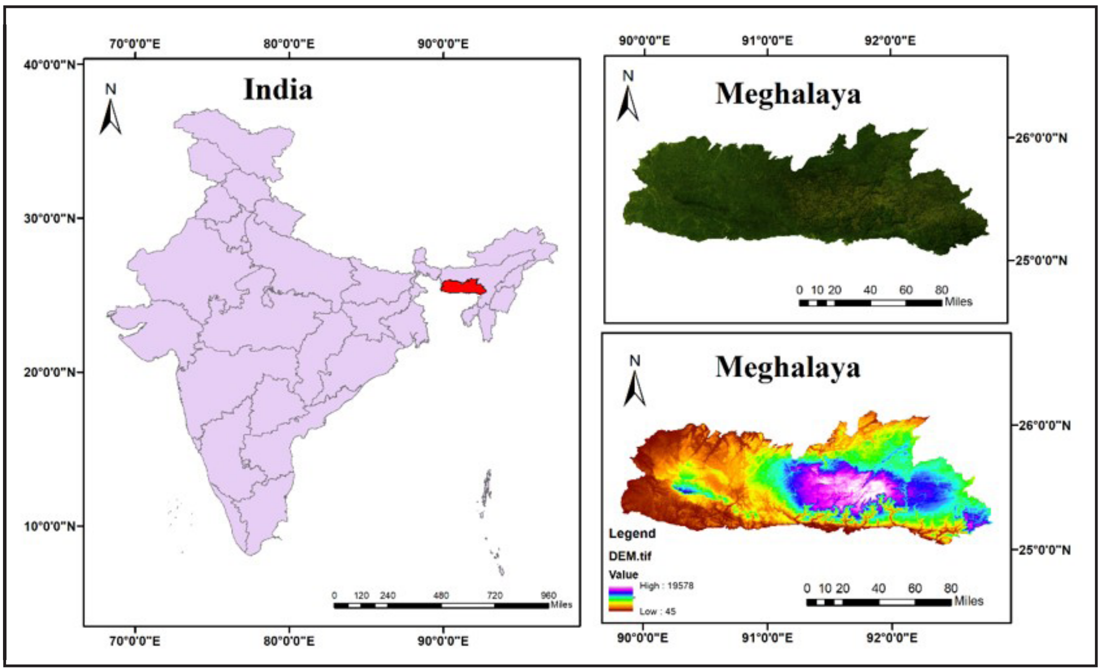


Fig. 1: Study Area Map

In recent decades, rapid urbanization in Meghalaya has exerted considerable pressure on the natural environment and initiatedvarious infrastructure development projects. This has resulted in reduced vegetation cover, loss of water bodies, and alterations in drainage systems. Over this period, Meghalaya's population has surged by 27.95%, rising from 2.32 million

(23.19 lakhs) to 2.97 million (29.67 lakhs) (Census of India data 2011) (Keerthi Naidu and Chundeli, 2023) .

Data Collection

Landsat 7 satellite images with a spatial resolution of 30 meters by 30 meters from the years 2000, 2010, and 2021 were acquired using Google Earth Engine. These images were selected to examine Land Use and Land Cover (LULC) changes for the specified timeframe. To ensure data accuracy, cloud-free images from November to December timeline were specifically chosen. The deliberate exclusion of cloud cover improves visibility and maintains the precision of the study.

The study was conducted in multiple steps. These steps include data preprocessing, feature collection, LULC classification, accuracy assessment, calculation of LST trend, and the analysis of the factors affecting LST. Fig 2 presents the workflow diagram of the study.

To minimize data preprocessing time and to ensure a likeness, the GEE cloud-based platform was used. With ample satellite imaginary and geospatial datasets available in this platform, it merges geospatial datasets using JavaScript and makes computational processing on satellite images easier (Kaur *et al.*, 2023). Several preprocessing steps were taken in this study including copy raster, mosaic to a new raster with all bands, and extract by mask.

For cloud and shadow masking the “pixelqa” band of Landsat-7and the Band Quality Assessment (BQA) function were used. The cloud confidence bit of pixelqa band less than 3 was considered as good pixels and those having confidence bit equal to 3 or more were considered as bad pixels. These identified “bad” pixels are then masked from the image, ensuring that only clear-sky pixels are retained. Finally, edge pixels were removed from all bands to refine the quality of the resulting image.

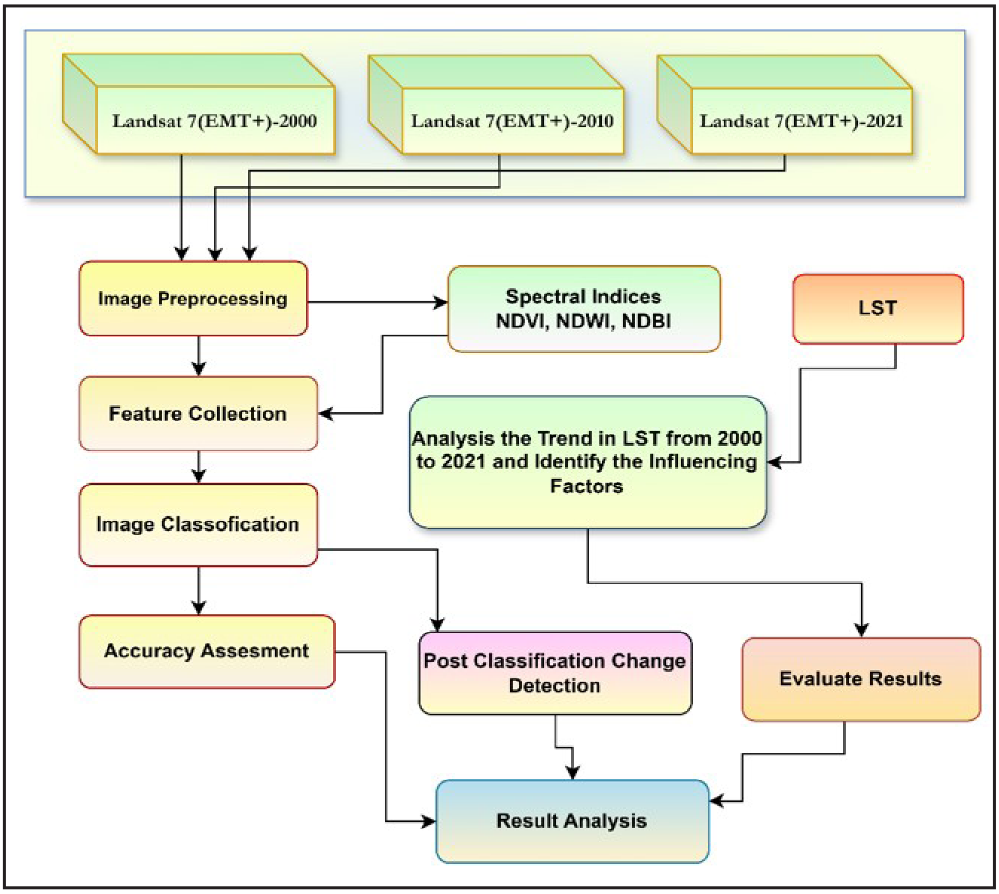


Fig. 2: Schematic Framework of the Methodology

Feature Selection

After the pre-processing stage, feature selection was carried out. Spectral indices such as NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), and NDBI (Normalized Difference Built-up Index) were calculated. These indices offer valuable insights into vegetation cover, water bodies, and built-up areas, respectively (Bhuyan *et al.*, 2023; Keerthi Naidu and Chundeli, 2023). By incorporating these spatial indices alongside standard band values, it is possible to capture the spectral features and enhance the model's accuracy.

Normalized Difference Vegetation Index (NDVI)

The NDVI is a useful tool for analyzing vegetation growth, monitoring changes in land cover, and researching the effects of environmental conditions on plant health.

$$NDVI = \frac{(NIR - \mathrm{Red})}{(NIR + \mathrm{Red})}$$

Here, NIR stands for near-infrared light reflectance and red stands for red light reflectance. Typically, reflectance measurements are obtained from satellite or aerial images. The resultant NDVI values vary from -1 to +1. Higher positive numbers imply thick and healthy vegetation, whereas lower or negative values indicate a lack of vegetation or a low vegetation density (Huang *et al.*, 2021).

Normalized Difference Water Index (NDWI)

NDWI stands for Normalized Difference Water Index. It is a spectral index used in remote sensing to identify and observe the presence of water bodies (Gao, 1995).

$$NDWI = \frac{(NIR - \mathrm{SWIR})}{(NIR + \mathrm{SWIR})}$$

Here, NIR represents the reflectance of near-infrared light, and SWIR represents the reflectance of shortwave infrared light. NDWI values vary from -1 to +1, with greater positive values indicating higher water concentrations and lower or negative values indicating non-water features (Gao, 1996). NDWI is frequently used to map and monitor water bodies such as lakes, rivers, and wetlands and to measure changes in water content over time.

Normalized Difference Built-Up Index (NDBI)

The Normalized Difference Built-Up Index (NDBI) is a spectral index used in remote sensing to detect and map built-up areas or urbanized regions (Zha *et al.*, 2003). NDBI compares the reflecting intensity of NIR and SWIR light wavelengths to determine the relative abundance of built-up or man-made materials within a given area.

$$NDBI = \frac{(SWIR - NIR)}{(SWIR + \mathrm{NIR})}$$

NDBI values span from -1 to +1, where positive values close to 1 indicate built-up areas, such as buildings and roads, while low values close to -1 represent non-built-up areas, such as vegetation or bare soil. NDBI is mainly used for urbanization monitoring, land use planning, and urban growth analysis.

LULC Classification

After collecting the features, the next phase in the LULC change study is the classification of Landsat images. Classification is a procedure that uses spectral properties to assign pixels in a view to distinct land cover classes (Amini *et al.*, 2022). In this particular study, a machine learning classifier, Random Forest, has been employed for the classification task. It is works on the decision tree principle, with numerous decision trees integrated to produce an ensemble model (Ho, 1995). Each decision tree in the random forest is trained on an arbitrary portion of the training data and makes judgments using a random subset of features (Fratello and Tagliaferri, 2018). This randomization adds variation to the trees, lowering the danger of over-fitting and increasing the model’s overall performance.

Random Forest can handle categorical and continuous variables, making it well-suited to the wide variety of data types commonly encountered in LULC datasets. Second, it can handle high-dimensional data successfully, supporting many input variables or features commonly seen in LULC datasets, such as spectral bands, indices, and texture measures. Furthermore, Random Forest resists noisy and correlated input variables, which is useful when working with remotely sensed imagery data susceptible to various noises and artifacts. Different Land cover classes identified in this study are presented in Table 1.

Table 1 : LULC Class Description

LULC class	*Description*
Water	This class comprises lakes, rivers, ponds, and reservoirs, which serve as important habitats for aquatic organisms and as a significant water supply.
Forest	This class includes open forests, dense forests, and big plantations, each indicating a different level of tree cover density and type of wooded region.
Vegetation	This category includes non-forest vegetation such as grasslands, shrubs, cultivated crops, and other types of plants.
Barren Land	Areas with little or no vegetation, such as deserts or stony terrains, have low ecological production.
Built-Up area	Comprises urban or developed areas with infrastructure, buildings, roads, and human settlements.

Data accuracy is crucial for processing and analyzing classified remote sensing data. The accuracy assessment used a confusion matrix to evaluate errors across five natural and man-made feature classes. Confusion matrix, the Overall Accuracy (OA) and the Kappa Statistic were calculated to validate the LULC classification results (Cohen, 1960; Pearson, 1904). Equation 4 is used to calculate the overall accuracy. Here, TP is the true positive, TN is the true negative, FP is the false positive, and FN represents the false negative value. Equation 5 is used to calculate the kappa coefficient. Here, Observed is $\frac{TP + TN}{N}$, and the calculation of random value shown in Equation 3. p1, p2 are calculated with Equation 1 and Equation 2, respectively.

$$p1 = \frac{TP + FP}{N} * \frac{TP + FN}{N} \tag{1}$$

$$p1 = \frac{FN + TN}{N} * \frac{FP + TN}{N} \tag{2}$$

$$Random = p1 + p2 \tag{3}$$

$$Overall\ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$kappa = \frac{\text{Observed} - \text{Random}}{1 - random} \tag{5}$$

Post-classification Change Detection

Change detection is a method that identifies a feature or phenomenon by examining it at its characteristic time (Afaq and Manocha, 2021), also defined as a technique of identifying geographical changes from Geographic Information System (GIS) data as a result of natural or artificial events. Change detection is vital for observing ecosystem changes, changes in land use and land cover, and land mapping changes (Shah *et al.*, 2023). Examining environmental conditions, such as consequences of natural disasters and urban growth, identifying changes in agriculture, assessing deforestation, and identifying specific urban or natural variances in the ecosystem, is usually done through LULC detection. Following the study of LULC using remote sensing data from the years 2000, 2010, and 2021 of Meghalaya, the next stage is to discover changes from 2000 to 2010 and 2010 to 2021. In this study, the QGIS SCP plugin was used to generate the change detection map for these periods. The objectives of change detection are to identify and measure changes in land cover classes over different timescales. Analyzing variations between 2000 and 2010 exposes previous LULC dynamics while evaluating changes between 2010 and 2021 indicates more recent trends. Change detection helps to monitor the environment, inform land management policies, and facilitate decision-making for sustainable land use practices (Jiang *et al.*, 2022).

LST trend and Factors Affecting the Trend

LST has emerged as a critical indicator for analyzing and interpreting the Earth's surface's thermal dynamics with vegetation changes and associated climatic events (Hossain *et al.*, 2023). From 2000 to 2021, the lowest and maximum temperatures in Meghalaya grew considerably, rising from 15.33°C to 20.25°C and 31.61°C to 35.81°C, respectively, as observed from LST data. Mann-Kendall (MK) test was performed to analyze the LST trend of Meghalaya. The MK test is a non-parametric statistical test that uses time series data to detect trends (Hamed and Ramachandra Rao, 1998). The output of the test contains h, p, tau, and slope. '*h*' represents the hypothesis test result; it indicates any significant trend present in data. '*p*' represents the probability of observing if the null hypothesis is true. '*Tau*' also known as Kendall's Tau represents the strength and direction of data. '*Slope*' represents magnitude; it represents the rate of change of a variable over time.Surface reflectance Landsat 7 ETM+ imaginary dataset is used to collect the LST data of Meghalaya.

Urban areas, known for their substantial solar radiation absorption, have higher LST. Vegetation or water regions, on the other hand, have lower LST values due to solar radiation scattering, plant heat absorption, and transpiration (Sajan *et al.*, 2023). The LST is influenced by various factors, including the Digital Elevation Model (DEM), NDVI, eastward and westward components of 10m wind, pressure, temperature, and rainfall. The DEM, which represents the land's topography, can impact LST due to its influence on local air circulation patterns and the exposure of different surfaces to solar radiation. NDVI is a measure of vegetation density. Wind components, particularly the eastward and westward components at 10m height, can influence LST by affecting air flow and heat transfer across the surface. Variations in pressure contribute to changes in atmospheric conditions, which in turn impact the LST. Finally, rainfall can cool down the land surface by reducing solar heating and changing moisture content. A total of 2,772 data points were generated using the Fishnet feature of ArcGIS for data collection purposes (ESRI, 2016). These components work together to determine the complex dynamics of LST, providing insights into the thermal behavior of the Earth's surface across different landscapes. To understand the relationship a random forest-based predictive model was fit and a variable of importance matrix was computed. The fitness of the model and the prediction accuracy were evaluated with R-square and RMSE values.

RESULTS

We have used Landsat-7 images to analyze the changes in LULC for the state of Meghalaya. The LULC map of Meghalaya is classified into five classes, viz. forest, water, vegetation, barren land, and build for the years 2000, 2010, and 2021. LULC map is a good reporter of different land cover types, monitoring changes over time and assessing the impact of human activities or natural phenomena on the landscape. We hypothesized that the depletion of vegetation or forest land would result in higher land surface temperatures. In this regard, our first step is to create a LULC map of Meghalaya for the years 2000, 2010, and 2021 (Fig. 3).

Table 2: Presents the Total area of Each Land Cover Class in Meghalaya, Represented in Square Kilometers (km^2) and as a Percentage (%), for the Years 2000, 2010, and 2021:

Class	Land cover 2000		Land cover 2010		Land cover 2021	
	Area (Km^2)	(%)	Area (Km^2)	(%)	Area (Km^2)	(%)
Water	358.90	1.60	225.91	1.01	111.20	0.50
Forest	19126.91	85.34	18195.67	81.18	18880.62	84.26
Vegetation	1968.11	7.57	2958.23	13.18	2635.29	11.76
Barren Land	1197.84	5.34	960.56	4.29	56.80	0.25
Build up	18.43	0.08	59.63	0.27	716.09	3.19
Total	22400	100	22400	100	22400	100

Table 3: Represent user Accuracy, Producer Accuracy, Kappa, and Overall Accuracy of the Random Forest Algorithm for the Years 2000, 2010, and 2021.

	Class	2000				2010				2021			
		User Accuracy	Producer Accuracy	Kappa	Overall Accuracy	User Accuracy	Producer Accuracy	Kappa	Overall Accuracy	User Accuracy	Producer Accuracy	Kappa	Overall Accuracy
RF	Water	0.94	0.92			0.98	0.96			0.79	0.90		
	Forest	0.99	0.99			0.98	0.99			0.96	0.91		
	Vegetation	0.98	0.95	0.95	0.98	0.95	0.96	0.97	0.98	0.63	0.68	0.81	0.87
	Barren	0.94	0.92			0.97	0.96			0.79	0.94		
	Build up	0.98	0.88			0.96	0.99			0.93	0.77		

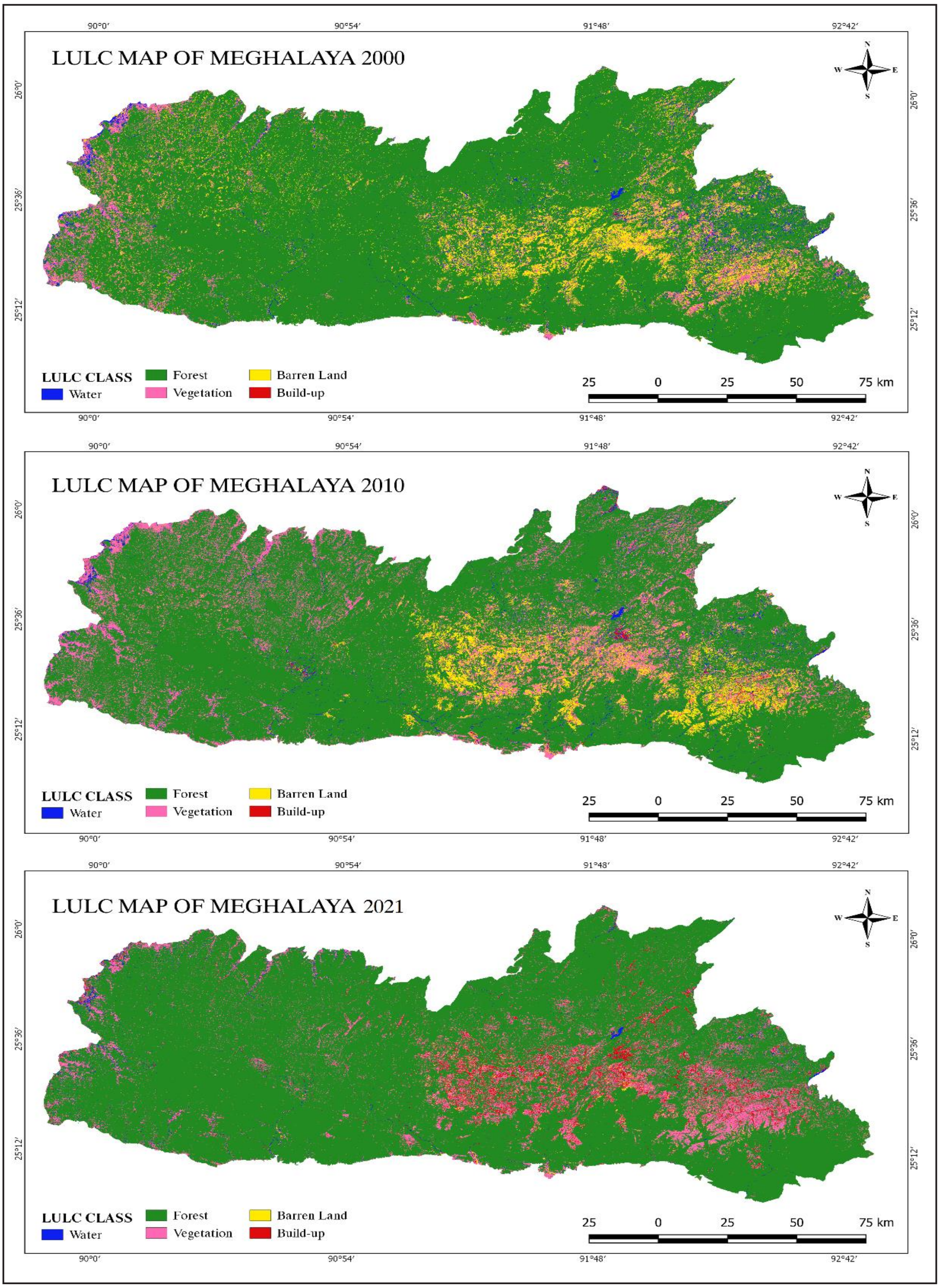
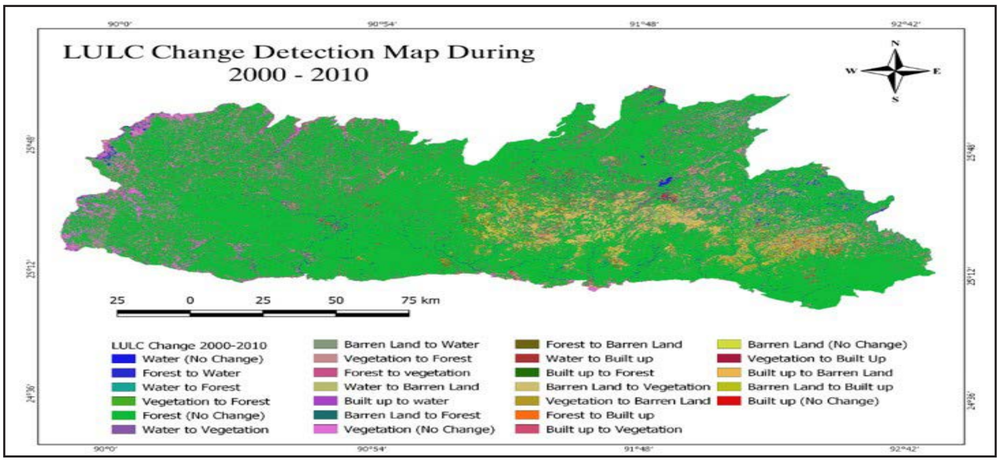
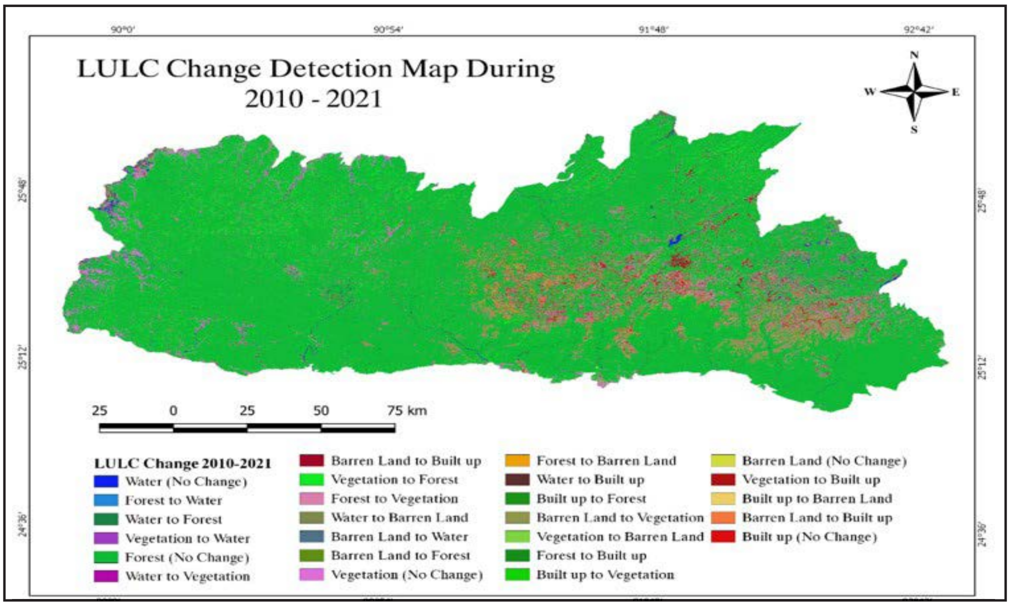


Fig. 3: Land use Land Cover Map of Meghalaya for the Years 2000, 2010, and 2021

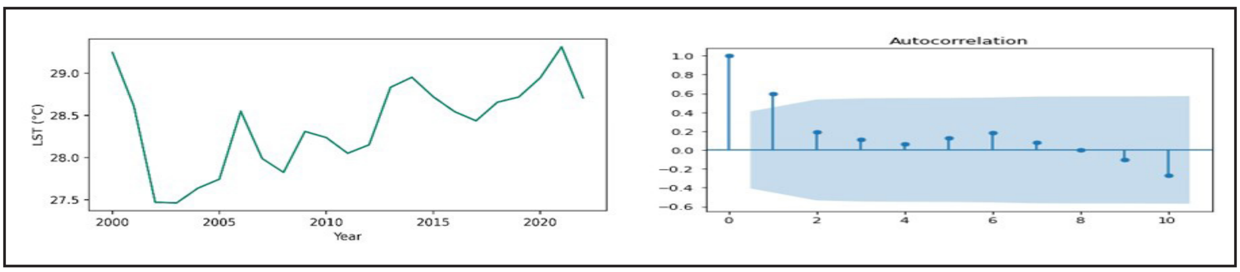


(A)



(B)

Fig. 4: The change detection map of Meghalaya is presented in two subgraphs (A) represents change detection map from 2000 to 2010 and (B) represents the change detection map from 2010 to 2021.



(A) (B)

Fig. 5: Represents two graphs; sub-graphs (A) represent the trend of LST of Meghalaya from 2000 to 2021, and (B) represent the Autocorrelation plot of LST.

The green color of Fig. 3 represents forest; the blue color represents water; the pink color represents vegetation; the yellow color represents barren land; and the red color represents buildup area. The total area cover in each class is presented in Table 2. To evaluate the authenticity of the results we also conducted accuracy measurements present in Table 3.

To evaluate the authenticity of the LULC map four evaluation measures has been utilized viz. user accuracy, producer accuracy, kappa and overall accuracy. Higher accuracy indicates the effectiveness of a classification algorithm in correctly identifying instances of a particular class among all the instances that truly belong to that class in reality.

Our next aim is to create a change detection map for 2000 to 2010 and 2010 to 2021. A change detection map indicates areas where changes have occurred between two or more periods (Fig. 4). These changes can include alterations in land cover, land use, infrastructure, vegetation cover, water bodies, and other features on the Earth's surface. Change detection maps are typically created by comparing remote sensing data, such as satellite imagery or aerial photographs, acquired at different time points. The process involves image classification to identify areas where significant changes have taken place. The outcome of this observation is to understand the maximum conversion of area, direction of conversion, and rate of the conversion. To justify the hypothesis, a change detection map is an important parameter. The maximum changes in land cover in the last 22 years presented in Fig.8 and Fig. 9 represent the percentage of area that converted to build-up area at the same period.

To understand the relationship between changing land cover and LST in the past 22 years we further conducted a trend analysis test on LST. The LST data pattern over the specified period is depicted in Fig. 5(A). We conducted a Mann-Kendall test to investigate the trend in LST. This non-parametric test is particularly useful when data is not normally distributed. It helps us understand whether there is a monotonic upward, downward, or no trend in the data. Additionally, an autocorrelation plot (Fig. 5(B)) was employed to detect correlations among data points in the LST dataset. This step is crucial for determining whether the basic Mann-Kendall test suffices or if a modified Mann-Kendall test is necessary for analysis. The statistics of the test are presented in Table 4.

Our next phase is to find out the factors affecting LST. Meanwhile, we performed the Person's correlation test to understand the linear correlation between the factors. The value of this test ranges from -1 to +1. A value closer to -1 indicates a negative correlation, and a value closer to +1 indicates a positive correlation between the two features (Fig. 6). This method is used to understand linear relationships. However, most of our data are nonlinear, so we further conducted a regression analysis to understand the nonlinear relationship. Fig. 7 presented regression analysis, where the x-axis presented the actual value and the y-axis presented the observed value. The red line represents the best-fit line; data points closer to the best-fit line indicate the prediction of the model is accurate. Further, to measure the accuracy of the model, we calculated the r-square and RMSE values of the model.

Table 4: Parameter statistics of Mann-Kendall test

Parameter	Value
h	True
p	<0.05
Tau	0.438
slope	0.061

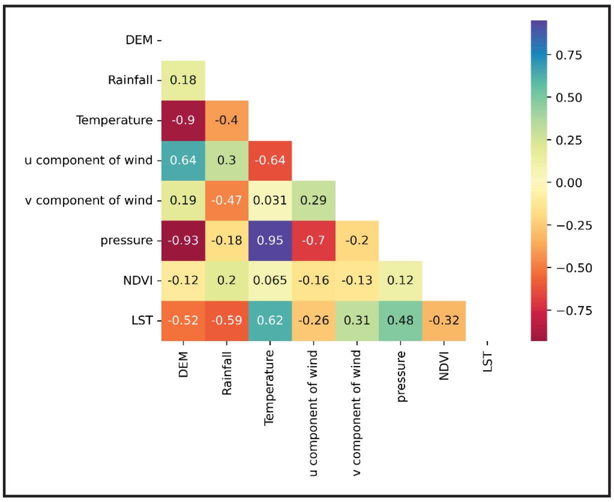
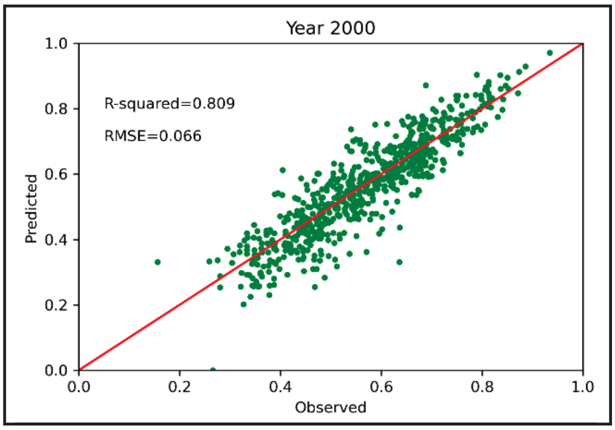


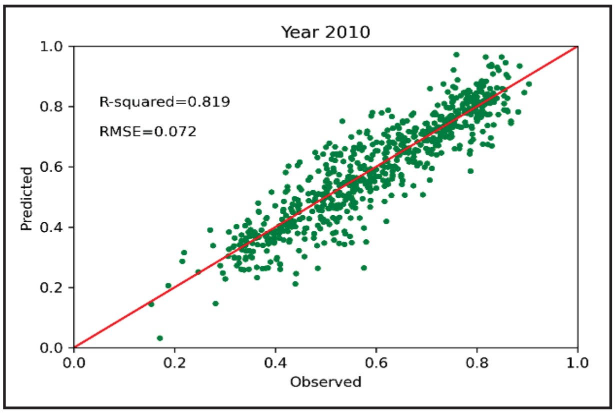
Fig. 6: Pearson's Correlation Map

DISCUSSION

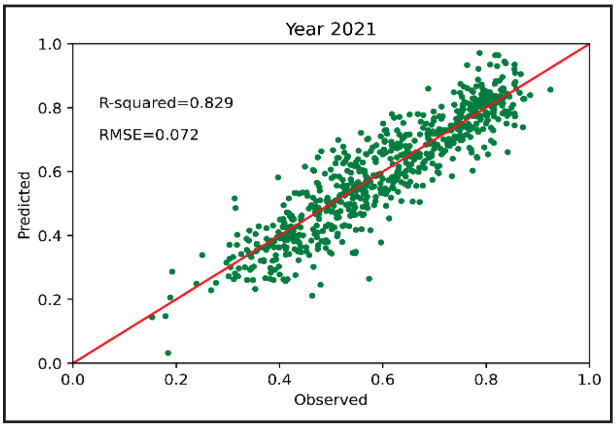
Processing voluminous satellite data can be incredibly challenging and time-consuming, necessitating platforms such as GEE for preprocessing and classification. We utilized random forest algorithm to generate an accurate land use land cover map of Meghalaya. Fig.3 depicts the classified map for 2000, 2010, and 2021, which illustrates each year's different land cover classes. The area calculation of the various land cover classes can be found in Table 2. These



(A)



(B)



(C)

Fig.7: Represents scattered graph, sub-graphs (A) represent observed vs. predicted graph for 2000, (B) represent observed vs. predicted graph for 2010, and (C) represent observed vs. predicted graph for 2021.

Between 2000 and 2010, Meghalaya experienced significant figures collectively indicate that the 22,400 km2 study area underwent significant changes. The overall water body decreased by 69.01% over the three years. 2000 water constituted 1.60% of the total area, but by 2021, it had been reduced to 0.50%. deforestation, with forest land being converted to vegetation land. This resulted in approximately 85.34% of the land being deforested in 2000, which decreased to 81.18% in 2010. The deforestation rate continued to decrease, and reforestation occurred between 2010 and 2021, resulting in an overall loss of 1.28% of forest land cover.

Furthermore, there was a 78% increase in vegetation land between 2000 and 2010, with an overall increase of 55.18% in vegetation land from 2000 to 2021. This suggests that some measures have been taken to reduce deforestation and promote vegetation growth.

Between 2000 and 2021, there was a notable reduction in barren land area, decreasing from 1197.84 km^2 to 56.80 km^2, marking a substantial decline. This reduction signifies a significant alteration in land cover, with approximately 95.25% of the barren land converted to different land cover types over the specified period.

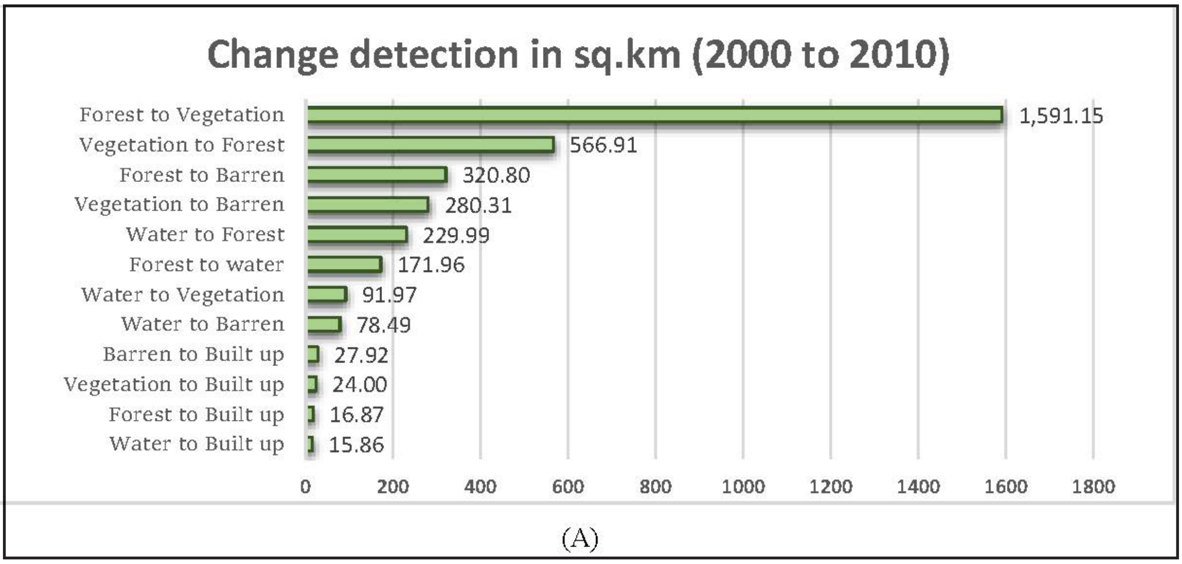
Table 2 illustrates a remarkable surge in the buildup area, particularly notable from 2010 to 2021, experiencing an approximate growth rate of 223.54%. Cumulatively, there was an overall increment of 3,785.45% in the buildup area from 2000 to 2021. This rapid increase indicates a trend towards unplanned urbanization, which in turn poses threats to biodiversity conservation efforts.

To ascertain the accuracy and reliability of the classified map, four accuracy measures were employed: user's accuracy, producer's accuracy, kappa coefficient, and overall accuracy. All accuracy metrics surpassed the 81% threshold, indicating the robustness and efficacy of the classification methodology utilized in this study.

The findings underscore the significant changes in land cover patterns over the studied period and emphasize the urgency of implementing sustainable land management practices to mitigate the adverse impacts of rapid urbanization on biodiversity and ecosystem services.

LULC Change Detection

The change detection maps, presented in Figures 4(A) and 4(B), depict the overall transition in land cover patterns for the periods 2000 to 2010 and 2010 to 2021, respectively.



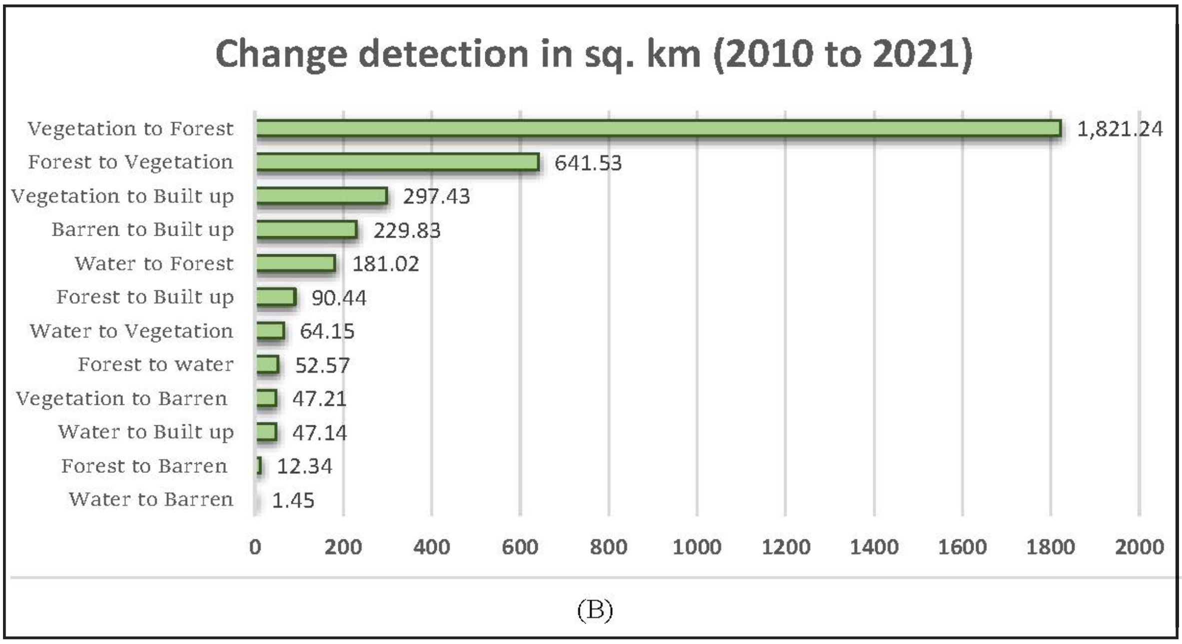


Fig. 8: Represents area (Km^2) conversion between classes with a bar graph; subgraphs (A) represent the change detection graph from 2000 to 2010, and (B) represent the change detection graph from 2010 to 2021

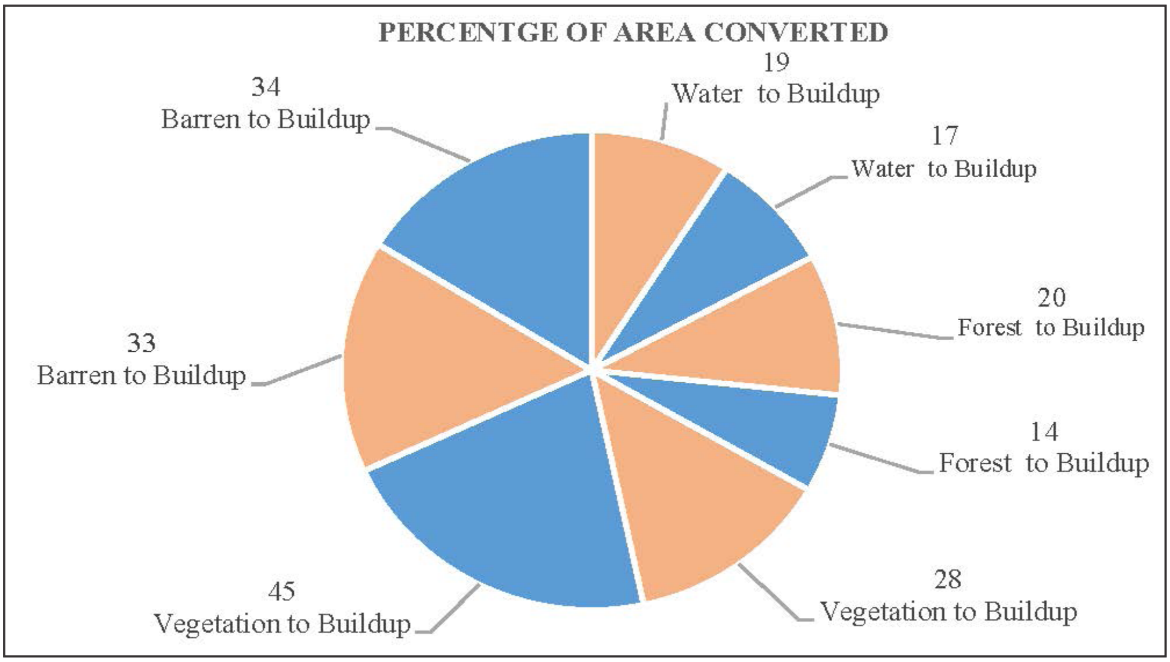


Fig. 9: Represents Percentage (%) of Area Vonversion from other Classes to Build-up

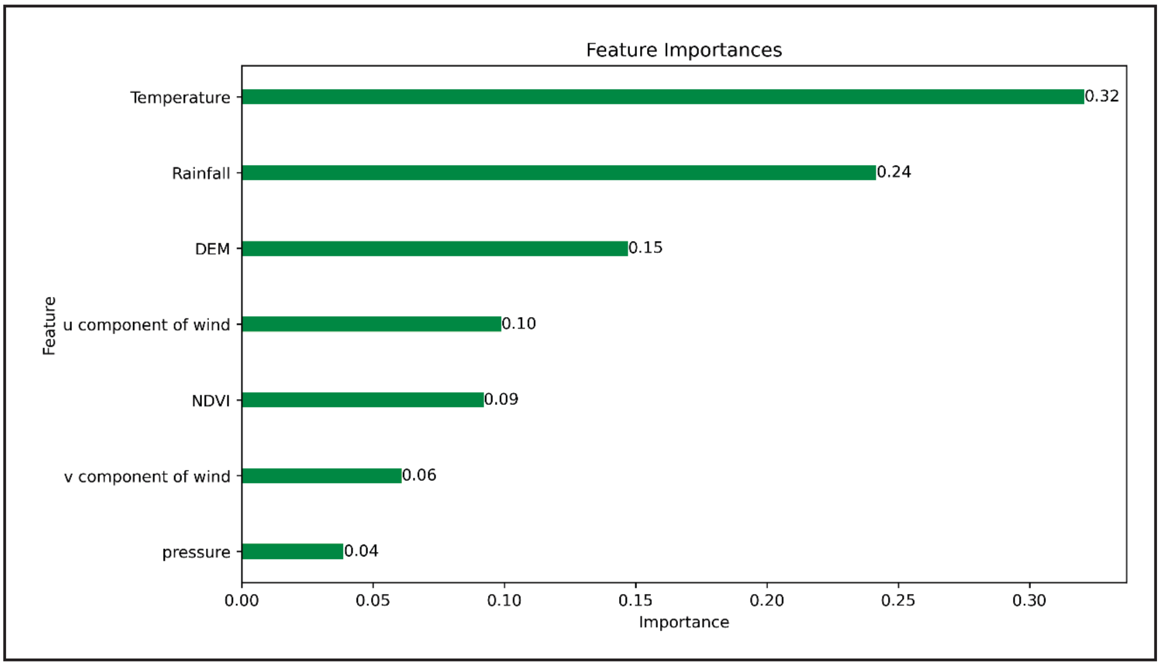


Fig.10: Feature Importance Graph

To discern significant changes, the data is further visualized through bar graphs in Figures 8(A) and 8(B).

Fig. 8(A) illustrates that from 2000 to 2010, a considerable area of approximately 1591.15 km^2 underwent a transition from forest to vegetation land, marking the most substantial change observed. Additionally, a notable conversion of approximately 566.91 km^2 from vegetation to forest land was noted. Other significant land cover changes include forest to barren land, vegetation to barren land, water to vegetation, and water to barren land. Minor changes were observed in the remaining land cover classes.

Conversely, in Fig. 8(B), covering the period from 2010 to 2021, notable transformations in land cover are depicted. The most prominent change observed was approximately 1821.34 km^2 between vegetation and forest areas. Conversely, approximately 641.52 km^2 of land transitioned from forest to vegetation. Additionally, approximately 297.43 km^2 of vegetation land was converted into build-up areas, while approximately 229.83 km^2 of barren land was transformed into build-up areas. Notable transitions observed in the 2010 to 2021 change detection plot include forest to build-up, water to vegetation, vegetation to barren, forest to barren, and water to barren land.

Further to analyze urbanization trends over the past two decades, we utilized pie chart (Figure 9). From 2000 to 2010, approximately 33% of barren land was converted to buildup areas, and from 2010 to 2021, this increased slightly to 34%, indicating improved land utilization. However, vegetation land conversion to buildup areas rose from 28% in the first decade to 45% in the second, raising concerns about food security and regional surface temperature. Additionally, 20% of forest land was converted to buildup areas in the first decade, with a slight decrease to 14% in the second. Despite the reduction, further decreases are necessary to protect forests, which are crucial for biodiversity and climate regulation.

LST Trend its Impacting Factors

This study aims to comprehend the LST trend and its correlation with urbanization and deforestation in Meghalaya. Initially, to ascertain the applicability of the Mann-Kendall test, we conducted an auto-correlation test. Fig. 5(A) and 5(B) illustrate the yearly LST and autocorrelation plots, respectively, with Fig. 5(B) indicating a slight correlation among data points. Subsequently, the Hamid and Rao Mann-Kendall test was employed, yielding a true h value and $p<0.05$, indicating a discernible trend. Tau and positive slope values further confirmed an increasing trend in LST data in the last 22 years. Notably, our study reveals that the recent rise in LST is directly linked to the augmented conversion of forest or vegetated land into built-up areas, aligning with findings published by (Alam *et al.*, 2022; Saha *et al.*, 2024).

The literature survey indicates that DEM, rainfall, temperature, and NDVI are correlated with LST. Additionally, we included pressure, the u-component of wind, and the v-component of wind in our analysis. The Pearson correlation graph (Fig. 6) illustrates that DEM, rainfall, NDVI, and the u-component of wind are negatively correlated with LST, while the remaining features show positive correlations. This finding aligns with previous studies (Feng *et al.*, 2019; Khandelwal *et al.*, 2018), which reported negative correlations between LST and DEM and NDVI. The negative relationship between LST and NDVI justified that the increasing LST trend is correlated with the deforestation or vegetation loss of a region. Further to examine the non-linear relationships, we conducted a Random Forest regression test. The scatter plot provides a visual representation, with R-square values ranging from 0.815 to 0.823, indicating that the models capture a significant portion of the data variability. Moreover, lower RMSE values suggest minimal prediction error. Additionally, the feature importance graph (Fig. 10) reveals that temperature, rainfall, DEM, and the u-component of wind, along with NDVI, are the most important variables in model building. However, the v-component of wind and pressure also significantly contribute to enhancing model performance.

Increasing LST and its Impact on our Food Security

The ramifications of elevated LST are profound, particularly impacting agriculture and ecosystems. Elevated temperatures induce heat stress in crops, leading to reduced photosynthesis, hindered growth, and decreased yields. Shifts in growing seasons disrupt planting and harvesting schedules, while intensified LST accelerates evapotranspiration rates, exacerbating water stress in plants and reducing water availability for crops. Furthermore, heightened temperatures expedite soil degradation processes such as erosion and desertification, compromising soil fertility and long-term crop health.

Changes in LST also disrupt ecosystems, altering habitats and affecting biodiversity. Urgent action is required to mitigate these impacts, including reducing greenhouse gas emissions, adopting improved agricultural practices, and advocating for sustainable land management. Our research in Meghalaya reveals a concerning trend: LST has been steadily increasing at an annual rate of 0.05%, attributed to rapid unplanned urbanization, deforestation, and vegetation area depletion. However, potential remedies such as increasing NDVI through reforestation efforts and enhancing rainfall can effectively mitigate rising LST and its adverse effects.

These findings significantly impact stakeholders engaged in policymaking, urban planning, environmental conservation, and land management in Meghalaya. They underscore the necessity of implementing strategic interventions to counteract the detrimental effects of escalating LST. Measures such as restraining urban expansion, promoting sustainable land utilization practices, and facilitating vegetation restoration efforts emerge as pivotal strategies to mitigate adverse consequences on local ecosystems and agriculture.

Future studies should focus on long-term monitoring of surface temperature trends and their socio-economic impacts, providing deeper insights into environmental dynamics and aiding in formulating effective mitigation strategies.

CONCLUSION

This study presents a comprehensive analysis of the increasing LST trend in Meghalaya over the past 22 years, highlighting its correlation with unplanned urbanization, deforestation and changing climate patterns. Our observations from the LULC maps between 2000 and 2010 indicate a high rate of deforestation and a significant expansion of agricultural land. Furthermore, LULC change detection reveals substantial conversion of vegetation and barren land into built-up areas, alongside a notable shrinkage of vegetation and water bodies over the last two decades.

Correlation analyses revealed that DEM, rainfall, and NDVI are negatively correlated with LST, aligning with findings from other studies. Additionally, we incorporated wind components and pressure features into our study. To examine the nonlinear correlation of these variables with LST, we conducted a regression analysis. The results provided valuable insights, indicating that all variables are crucial for accurate LST prediction. Among these, temperature, rainfall, and DEM emerged as the most significant factors.

Overall, this study underscores the increasing LST trend as a consequence of unplanned urbanization and the reduction of forest and vegetation land in Meghalaya. These findings suggest the urgent need for more sustainable development plans for the future of the state. Our results could serve as a valuable asset for government authorities and policymakers to aid in creating better development strategies. While this study focused on data from the last two decades, incorporating more extensive datasets could further enhance analysis and decision-making processes.

ACKNOWLEDGEMENT

The study is carried out without any external funding.

REFERENCES

Afaq, Y., Manocha, A., 2021. Analysis on change detection techniques for remote sensing applications: A review. Ecol. Inform. 63. https://doi.org/10.1016/j.ecoinf.2021.101310

Alam, H. mim E., Arafat, M.Y., Ahmed, K.T., Uddin, M.N., 2022. Temporal Variation of Land Surface Temperature in Response to Changes in Vegetation Index of Bhawal National Park, Bangladesh, in: Lecture Notes in Civil Engineering. pp. 329–337. https://doi.org/10.1007/978-981-16-5543-2_27

Amini, S., Saber, M., Rabiei-Dastjerdi, H., Homayouni, S., 2022. Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series. Remote Sens. 14. https://doi.org/10.3390/rs14112654

Bhuyan, M., Singh, B., Vid, S., Jeganathan, C., 2023. Analysing the spatio-temporal patterns of vegetation dynamics and their responses to climatic parameters in Meghalaya from 2001 to 2020. Environ. Monit. Assess. 195. https://doi.org/10.1007/s10661-022-10685-6

Chakraborty, T., Saikia, U.S., 2022. Climate change implications on groundwater resources in East Khasi Hills, with special reference to greater Shillong, Meghalaya, India. pp. 3–17. https://doi.org/10.1016/b978-0-323-85378-1.00001-5

Chughtai, A.H., Abbasi, H., Karas, I.R., 2021. A review on change detection method and accuracy assessment for land use land cover. Remote Sens. Appl. Soc. Environ. 22, 100482. https://doi.org/10.1016/j.rsase.2021.100482

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educ. Psychol. Meas. 20, 37–46. https://doi.org/10.1177/001316446002000104

Duan, H., Shang, C., Yang, K., Luo, Y., 2022. Dynamic Response of Surface Water Temperature in Urban Lakes under Different Climate Scenarios—A Case Study in Dianchi Lake, China. Int. J. Environ. Res. Public Health 19. https://doi.org/10.3390/ijerph191912142

ESRI, 2016. ArcMap 10.3. ESRI, Redlands, California.

Ezimand, K., Aghighi, H., Ashourloo, D., Shakiba, A., 2024. The analysis of the spatio-temporal changes and prediction of built-up lands and urban heat islands using multi-temporal satellite imagery. Sustain. Cities Soc. 103. https://doi.org/10.1016/j.scs.2024.105231

Feng, Y., Gao, C., Tong, X., Chen, S., Lei, Z., Wang, J., 2019. Spatial patterns of land surface temperature and their influencing factors: A case study in Suzhou, China. Remote Sens. 11. https://doi.org/10.3390/rs11020182

Fratello, M., Tagliaferri, R., 2018. Decision trees and random forests, in: Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics. pp. 374–383. https://doi.org/10.1016/B978-0-12-809633-8.20337-3

Gao, B.-C., 1995. Normalized difference water index for remote sensing of vegetation liquid water from space, in: Imaging Spectrometry. p. 225. https://doi.org/10.1117/12.210877

Gao, B.C., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gondwe, J.F., Lin, S., Munthali, R.M., 2021. Analysis of Land Use and Land Cover Changes in Urban Areas Using Remote Sensing: Case of Blantyre City. Discret. Dyn. Nat. Soc. 2021, 1–17. https://doi.org/10.1155/2021/8011565

Guha, S., 2021. Dynamic seasonal analysis on LST-NDVI relationship and ecological health of Raipur City, India. Ecosyst. Heal. Sustain. https://doi.org/10.1080/20964129.2021.1927852

Ham, J., Yoo, H.J., Kim, J., Lee, B., 2022. Vowel speech recognition from rat electroencephalography using long short-term memory neural network. PLoS One 17. https://doi.org/10.1371/journal.pone.0270405

Hamed, K.H., Ramachandra Rao, A., 1998. A modified Mann-Kendall trend test for autocorrelated data. J. Hydrol. 204, 182–196. https://doi.org/10.1016/S0022-1694(97)00125-X

Ho, T.K., 1995. Random decision forests, in: Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. pp. 278–282. https://doi.org/10.1109/ICDAR.1995.598994

Hossain, M.S., Khan, M.A.H., Oluwajuwon, T.V., Biswas, J., Rubaiot Abdullah, S.M., Tanvir, M.S.S.I., Munira, S., Chowdhury, M.N.A., 2023. Spatiotemporal change detection of land use land cover (LULC) in Fashiakhali wildlife sanctuary (FKWS) impact area, Bangladesh, employing multispectral images and GIS. Model. Earth Syst. Environ. 9, 3151–3173. https://doi.org/10.1007/s40808-022-01653-7

Huang, S., Tang, L., Hupy, J.P., Wang, Y., Shao, G., 2021. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. J. For. Res. https://doi.org/10.1007/s11676-020-01155-1

Jiang, H., Peng, M., Zhong, Y., Xie, H., Hao, Z., Lin, J., Ma, X., Hu, X., 2022. A Survey on Deep Learning-Based Change Detection from High-Resolution Remote Sensing Images. Remote Sens. https://doi.org/10.3390/rs14071552

Kanagasundaram, G., Dissanayake, K., Samarasuriya, C., 2022. Remote Sensing and GIS Approach to Monitor the Land-Use and Land-Cover Change in Kaduwela Metropolitan Area, in: MERCon 2022 - Moratuwa Engineering Research Conference, Proceedings. https://doi.org/10.1109/MERCon55799.2022.9906148

Kaur, H., Tyagi, S., Mehta, M., Singh, D., 2023. Time series (2001/2002–2021) analysis of Earth observation data using Google Earth Engine (GEE) for detecting changes in land use land cover (LULC) with specific reference to forest cover in East Godavari Region, Andhra Pradesh, India. J. Earth Syst. Sci. 132. https://doi.org/10.1007/s12040-023-02099-w

Keerthi Naidu, B.N., Chundeli, F.A., 2023. Assessing LULC changes and LST through NDVI and NDBI spatial indicators: a case of Bengaluru, India. GeoJournal 88, 4335–4350. https://doi.org/10.1007/s10708-023-10862-1

Khandelwal, S., Goyal, R., Kaul, N., Mathew, A., 2018. Assessment of land surface temperature variation due to change in elevation of area surrounding Jaipur, India. Egypt. J. Remote Sens. Sp. Sci. https://doi.org/10.1016/j.ejrs.2017.01.005

Lamare, R.E., Singh, T.S., Singh, O.P., 2019. Land use land cover change in East Jaintia Hills, Meghalaya in relation to limestone mining. Environ. Ecol.

Moisa, M.B., Gabissa, B.T., Hinkosa, L.B., Dejene, I.N., Gemeda, D.O., 2022. Analysis of land surface temperature using Geospatial technologies in Gida Kiremu, Limu, and Amuru District, Western Ethiopia. Artif. Intell. Agric. 6, 90–99. https://doi.org/10.1016/j.aiia.2022.06.002

Pearson, K., 1904. On the Theory of Contingency and Its Relation to Association and Normal Correlation. Drapers' Co. Res. Mem. Biometric Ser. I.

Prokop, P., 2020. Remote sensing of severely degraded land: Detection of long-term land-use changes using high-resolution satellite images on the Meghalaya Plateau, northeast India. Remote Sens. Appl. Soc. Environ. https://doi.org/10.1016/j.rsase.2020.100432

Rogers-Bennett, L., Catton, C.A., 2022. Cascading impacts of a climate-driven ecosystem transition intensifies population vulnerabilities and fishery collapse. Front. Clim. 4. https://doi.org/10.3389/fclim.2022.908708

Saha, J., Ria, S.S., Sultana, J., Shima, U.A., Hasan Seyam, M.M., Rahman, M.M., 2024. Assessing seasonal dynamics of land surface temperature (LST) and land use land cover (LULC) in Bhairab, Kishoreganj, Bangladesh: A geospatial analysis from 2008 to 2023. Case Stud. Chem. Environ. Eng. 9. https://doi.org/10.1016/j.cscee.2023.100560

Sajan, B., Kanga, S., Singh, S.K., Mishra, V.N., Durin, B., 2023. Spatial variations of LST and NDVI in Muzaffarpur district, Bihar using Google earth engine (GEE) during 1990-2020. J. Agrometeorol. 25, 262–267. https://doi.org/10.54386/jam.v25i2.2155

Shah, A., Ayushi, S., Bhumika, R., Bhandary, K.P., Krishnan, R., Vindhya, M., 2023. Change Detection of Terrain Using Satellite Imagery, in: IEEE International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems, ICAECIS 2023 - Proceedings. pp. 453–458. https://doi.org/10.1109/ICAECIS58353.2023.10170208

Singh, P.P., Chakraborty, T., Dermann, A., Dermann, F., Adhikari, D., Gurung, P.B., Barik, S.K., Bauhus, J., Fassnacht, F.E., Dey, D.C., Rösch, C., Saha, S., 2020. Assessing restoration potential of fragmented and degraded fagaceae forests in Meghalaya, North-East India. Forests 11. https://doi.org/10.3390/F11091008

Staszewski, T., 2023. CHARACTERISTICS OF CLIMATE CHANGE IN THE YEARS 1850-2100. Inżynieria Bezpieczeństwa Obiektów Antropog. 1–6. https://doi.org/10.37105/iboa.170

Suri, A., Jena, P., Azad, S., 2023. Rainfall Wet Spells Variability Across Temperature Mean-Change Years in the Northwestern Himalayan Region. Earth Sp. Sci. https://doi.org/10.1029/2022EA002568

Vohra, R., Kumar, A., Jain, R., Hemanth, D.J., 2024. Analysis and prediction of land surface temperature with increasing urbanisation using satellite imagery. Heliyon 10. https://doi.org/10.1016/j.heliyon.2024.e40378

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int. J. Remote Sens. 24, 583–594. https://doi.org/10.1080/01431160304987

Zhang, Y., Li, X., Zhang, K., Wang, L., Cheng, S., Song, P., 2023. A Simple Real LST Reconstruction Method Combining Thermal Infrared and Microwave Remote Sensing Based on Temperature Conservation. Remote Sens. 15. https://doi.org/10.3390/rs15123033

Zheng, X., Zhang, D., Su, J., CHEN, H., 2022. Prediction of carbon emission based on energy consumption structure by statistical forecasting. p. 236. https://doi.org/10.1117/12.2628222