International Journal of Advanced Technology and Engineering Exploration, Vol 10(109)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2022.10100536

Research Article

Machine learning based UHI data assessment to model the relationship between LULC and LST: case study of Srinagar City, Jammu & Kashmir, India

Mujtaba Shafi[1*], Amit Jain[2] and Majid Zaman[3]
Research Scholar, University Institute of Computing, Chandigarh University, Gharuan, Mohali, Punjab, India[1]
Professor, University Institute of Computing, Chandigarh University, Gharuan, Mohali, Punjab, India[2]
Scientist-E, Directorate of IT & SS, University of Kashmir, Hazratbal, Srinagar[3]

Received: 12-December-2022; Revised: 21-December-2023; Accepted: 23-December-2023

Abstract
The pressing issue of global climate change is being rigorously examined, with urban heat islands (UHIs) identified as a contributing factor. A UHI is a city or town area exhibiting a temperature variance from its surrounding environment. Researchers employ various relative UHI parameters to model UHI data and predict temperature fluctuations. For the study area of Srinagar City, Jammu & Kashmir, India, land surface temperature (LST) data and its correlated parameters were sourced from satellite imagery. The LST data for the region were analyzed to understand the development and investigate the UHI effect and its variations. Utilizing an 8-day revisit period during the busiest season month, the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite provided LST data from 2001 to 2021. To address the large dataset, which includes 16 samples of LST data per square kilometer each year measured in Kelvin (K), different machine learning (ML) techniques were employed to establish associations for UHI modeling. These clusters were then compared with established scientific categories in the study area. The long short-term memory neural network (LSTM) model, using time-series data, predicted changes in urbanized areas, vegetation cover, wetlands, and other factors. For land use land cover (LULC) prediction, the neural network (NN) model outperformed all others, with Regression (R) = 0.897, Validation = 0.912, and Training = 0.931, and mean squared error (MSE) = 2.012, Validation = 0.191, and Training = 1.124. This paper strives to identify and analyze the relationship between LULC change and LST variations in the context of urbanization. Initially, it examines the correlations between LST and variables such as vegetation, man-made structures, and agriculture, employing built-up indices within each LULC category and vegetation cover. Subsequently, it assesses the impacts of LULC change and urbanization on UHI using hot spot and urban landscape analyses. Finally, it proposes a model employing non-parametric regression to predict future urban climate trends, considering anticipated changes in land cover and land use.

Keywords
Land use land cover, Land surface temperature, Urban heat islands, Machine learning.

1.Introduction
1.1Background
In recent years, urban living has dramatically changed for humanity. At the beginning of the twentieth century, only 10% of the world's population lived in metropolitan areas. It has now reached nearly 50% of the present population, and in the upcoming years, this percentage of the urban population will increase even further. Urbanization is seen as a significant human endeavor that denotes the transformation of a naturally vegetated and terrain into a built-up and impenetrable one. The usage of materials, including concrete, marble, tiles, roof sheeting, and bitumen, is a significant contributor to this impermeable area [1]. A broad spectrum of professionals has expressed interest in studying urbanization. The subject's interdisciplinary scope piques the interest of ecologists, urban planners and civil engineers, sociologists, geographers, administrators, and, eventually, the general public [2]. This is due to the large number of activities and processes that occur in urban ecosystems daily [3]. Although an adequate definition of urban sprawl is debatable, there is general agreement that it is defined by an uncontrolled pattern of expansion, driven by a variety of mechanisms and resulting in inefficient resource allocation [4, 5].

*Author for correspondence

The temperature variance between urban and neighboring neighboring zoneshas been widely observed, and this phenomenon has been dubbed the "urban heat island (UHI) effect" [6]. It is a global problem that threatens the sustainability and livability of our urban environments. Addressing this complexity in monitoring and reviewing urban planning and management processes and practices cannot be overstated. Spatial analysis and “geographic information systems” (GIS) have been mainly used in the last two decades to identify, interpret, analyze, and simulate land use change processes. Satellite imaging delivers consistent data across large areas at various geographical scales [7–9].

The land surface temperature (LST) produced from landsat images is utilized research and analyze the UHI effect, which makes considerable use of satellite data. The surface urban heat island (SUHI) effect is the name given to the UHI effect when it is researched using LST. LST is a crucial element in understanding urban climate since it is one of the fundamental factors governing the Earth's physical, chemical, and biological processes [10]. LST regulates surface heat and water exchange with the atmosphere and may be calculated from the effective radiating temperature of the Earth's surface. Many heat balance parameter estimation, and climatological monitoring studies have made use of LST [11–13].

1.2Challenges

Since the study area, i.e., Srinagar city, has a tremendous amount of cloud cover during the winter, it can impede moderate resolution imaging spectroradiometer (MODIS) from collecting UHI data. This, throws a challenge for a researcher to work on structured and voluminous data.

1.3Motivation

Since significant urban development, Srinagar city of Jammu and Kashmir, India, has seen a tremendous urban expansion. In order todetermine the short- and long-term magnitudes of land use land cover (LULC) and LST alteration in this region, techniques based on recurrent neural network (RNN) architecture, namely long short term memory (LSTM), can be practical, which ultimately shall be helpful to urban planners and politicians to lessen the impact of UHI [14]. As a result, this study examines the link between LST under various LULC categories, calculates the historical data change of LULC and its influence on LST, and conducts a correlation analysis of various land cover metrics with LST for the years 2001, 2011, and 2021. The LSTM algorithm has also been used in this study to estimate the LULC and LST situation for Srinagar city in 2024 [15, 16].

1.4Objectives

The following are the main objectives of the proposed study.

- To develop a predictive framework for structured voluminous UHI datasets
- To validate the performance of the proposed framework with the state-of-the-art parameters.

1.5Contribution

In this study, an effort has been made to pinpoint and examine the relationship between LULC change and changes in LST in the context of urbanization. Initially, emphasis had been given to studying the relationship between LST and vegetation, man-made features, and agriculture using built-up indices within each LULC category and vegetation cover. The impacts of LULC change and urbanization in UHI are then assessed using hot spot analysis and urban landscape analysis. Finally, a proposed model uses non-parametric regression to predict future trends in an urban environment, taking into account expected changes in land cover and land use. The findings of urban sprawl and mitigating the effects of UHI.

The structure of this article is organized as follows: Section 1 provides a brief overview of urbanization and various related topics. Section 2 summarizes the literature review. The data and methodology adopted are detailed in Section 3. Section 4 presents the findings, and the discussion of the study is elaborated in Section 5. Finally, Section 6 presents the conclusion and outlines future work.

2.Review of literature

Several studies have been conducted to understand UHI phenomena, establish the relationship between LULC and LST, and modelling UHI data. Studies [17–20] looked at 32 Chinese cities using LST data and discovered that the annual average UHI effect was higher during the day than at night. An insightful information was provided about how the UHI was evaluated on a regional level in China [21]. Additionally, based on LST, authors [22] created an algorithm to map the global intensity of UHI. This study uses relative LST as a UHI indicator, continuing a recent trend. The temperature difference inside the urban area will henceforth be referred to as relative LST [23]. Similar to this, other research [24–26] has looked at a wide range of urban

characteristics and elements that contribute to the development and deepening of the UHI. These variables may be categorized into three primary groups, namely environmental, socioeconomic, and urban morphology variables, as shown in *Table 1*. The study views climate, meteorological parameters, and land features as uncontrolled elements that regulate environmental parameters.

Table 1 Three categories are used to summarize the factors affecting UHI. Climate conditions, meteorological variables, and geographical qualities all contribute to the environment. While the complicated intricacy of urban function governs socioeconomic aspects, urban elements' geometric performance and design determine urban morphology

Category	Factors	References
Urban Category	Pavements Water bodies Open Spaces Vegetation	[27, 28]
Socio-economic	Infrastructure Buildings Built-up ratio Heat Appliances Transportation Building height, width aspect ratio(AR)	[29, 30]
Environmental Factors	Geographical features Forest cover Vegetation Population densities Land use Long and short-wave radiations	[31]

To mimic UHI, researchers [32] have used various physics-based modeling techniques. Urban canopy models (UCMs) are the most well-known physics-based approaches to UHI evaluation. The "weather research and forecasting (WRF)" examines how urbanization affects local climate use UCMs. This method can predict how urbanization will affect local climate change by capturing exchanges between the land surface and the atmosphere. To describe the UHI in Toronto, Canada, three distinct UCMs were assessed for effectiveness. The researchers [33] concluded that the more complicated UCMs (multi-layer models), primarily because these models do not replicate the variety of urban architecture, do not reliably forecast air temperature.

Similar attempts have been made to arm decision-makers with the resources they need to develop effective UHI mitigation plans. For instance, researchers [34] gave policymakers an analytical tool when combining four specified mitigation methods. Several researchers recently put forth a paradigm for selecting the best variety of UHI mitigation measures, including performance evaluation models, simulation models, and genetic algorithm (GA) refinement in the selection process [35]. However, these tools are beneficial for creating mitigation plans and don't offer a quick evaluation of how decisions on urban design may affect UHI. Adopting mitigation techniques would usually require planners' proactive use of extra UHI impact mitigation measures (e.g., green roofs) [36, 37]. Urban planners can equip themselves with a tool that can assist them in examining the effect of their typical urban design decisions on UHI and determine what can be done to minimize the drawbacks by altering the design before they get to the stage where they wish to apply specialized solutions for the reduction of UHI [38].

Apart from these, recent developments in machine learning (ML) modeling techniques and data mining platforms have proven to be helpful in data extraction from a range of readily accessible datasets [39, 40]. In [41], unsupervised ML was used to anticipate China's pollutant concentrations. They [41] developed a methodology to model sustainability performance at multiple different temporal and spatial scales by integrating data-driven modeling techniques and building energy simulation strategies. The findings indicate that by including ML techniques in the current simulation workflow, there is a sizeable, significant chance to increase the accuracy of energy models. To the authors' best knowledge [42, 43], few recent researchers have employed data-driven approaches to simulate UHIs despite the methods' promise. For instance, to look into how temperatures throughout the night affect energy usage at the metropolitan level, authors

developed an ML model by using supervised algorithms based on open-source information available in Berlin City. Similarly, using unsupervised learning and statistical techniques, the researcher evaluated urban planning in terms of UHI changes [44].

The paper also highlights ML modeling techniques and data mining platforms to extract data from readily available datasets. Studies have used unsupervised and supervised ML to anticipate pollutant concentrations, simulate energy performance, and evaluate urban planning regarding UHI changes. However, only a few studies have attempted to offer decision-makers and urban planners practical results. The current methodologies do not generally incorporate socio-economic factors such as population density or traffic flow, and they do not frequently use publicly accessible datasets, which further limits their appeal.

The use of ML techniques to comprehend the interplay between various urban characteristics has a lot of promise, as can be shown from the review above Nevertheless, only a small number of research have attempted to offer decision-makers and urban planners results that are simple to put into practice. Instead of emphasizing explicability and interpretability, the tendency in that study area to focus more on prediction and accuracy. As it is known that ML techniques can execute a vast number of record features, socio-economic factors like population density or traffic flow are not generally incorporated [45]. The fact that the current methodologies don't frequently use publicly accessible datasets may further limit their appeal. In order to have a bird's eye view regarding the studies conducted using ML techniques, we have enumerated all the studies in *Table 2*, highlighting methods used, results, advantages, and limitations.

Table 2 Comparative analysis of ML techniques in terms of results obtained, advantages, and limitations used in existing works

S. No.	ML methods used	Results	Advantages	Limitations
1	Regression analysis [39]	Normalized difference built-up index (NDBI) , normalized difference vegetation index (NDVI) are effective indicators	Quantifies LULC impacts on LST	Limited to specific study area
2	Artificial neural network (ANN) [40]	High agreement between predicted and real LULC maps	Reliable predictions for future LULC. Reduced subjectivity and uncertainty	Limited to specific study area
3	Random forest(RF) [41]	R2: 0.56-0.71, mean absolute error (MAE): 0.05-0.07	Accurate models at the city level	Low generalizability between cities
4	LSTM Support Vector Regression (SVR) [42]	Low errors MAE, mean squared error (MSE) for LST estimation Higher errors (MAE, MSE) for LST estimation	Higher accuracy and reliability compared to SVR	Limited to specific study area
5	Support vector machine (SVM) ANN RF [43]	High classification performance Predicts LST and LULC distributions Predicts SUHI variations	Outperforms other classification techniques Utilizes various modeling approaches and prediction tools Evaluates overall performance using R2 and MSE	Suitable for remote sensing applications Depends on quality and availability of input data Limited to specific study area
6	ML ANN cellular automata (CA) [44]	Accurate prediction of LULC changes Accurate mapping of surface temperature changes Future LULC prediction accuracy of 89.2%	Provides insights for minimizing extreme UHI effects Demonstrates high prediction accuracy A broad review of UHI and thermal data Highlights progress in multi-sensor image optimization	Limited to specific study area Limited to specific study area Focuses on specific methodologies and applications

3.Methods

3.1Dataset and its description

The initial step is to access and download MODISs high-resolution satellite imageries. The ten images were clipped using the boundary shape file (the coordinates of Srinagar city) retrieve the LST data within the chosen coordinates. The data for relative parameters (elevation and aspect) have been acquired from the digital elevation model (DEM) tool. (30 m shuttle radar topography mission (SRTM) and 1 km global topographic (GTOPO) MODIS data have been resampled to 30 m to quantify the impact of resolution in predicting UHI. As such, the UHI information for Srinagar city has been computed at two spatial details- the coarse resolution (1 km grid) and the finer resolution (30 m grid). LST data was evaluated in this way to assess the urbanization of Srinagar city and help in analyzing the UHI impact and its variation over Srinagar city. Spring, summer, autumn, and winter are the four seasons in the study area (Srinagar City). The LST data was taken during the peak month of the season and the 8-day revisit time [46]. The following (*Table 3*) describes the approach used to gather LST data from 2001 to 2021:

Table 3 Data extraction time interval

Season	Month	Day (8-day visit period)
Winter	January	1, 9, 17, 25
Spring	April	97, 105, 113, 121
Summer	August	217, 225, 233, 241
Autumn	November	305, 313, 321, 329

The Srinagar development authority (SDA) provided the research area's coordinates and boundaries, which covered urban and suburban cover areas (*Figure 1*). Three zones or regions make up the city of Srinagar. U, V, and W stand for urbanized area, vegetation index, and waterbodies/wetland. Based on Srinagar's geographic viability, geoscientists have classified the area. The categorization is utilized in the analysis without being altered because it is hardcoded in the dataset.

The dataset consists of 6000 records with 21 attributes, including year, feature identification (FID), LST in Kelvin at various times of the year (1-16), AR, elevation, and the target variable LULC.

The general description of the attributes is as follows:
LST: For the current investigation, LST from eight days separated by 1 km were employed. The study uses data spanning 19 years, from 2001 to 2021. Every year, data from four set periods (winter, spring, summer, and autumn) have been used [47].
Elevation: The advanced spaceborne thermal emission and reflection radiometer (ASTER) DEM map was used to determine the elevation of the research area's pixels. Standard ASTER DEM data packages are created with 30 million posts and have Z accuracies, typically between 10 million and 25 million root mean squared error (RMSE). The ASTER DEM sceneries including stereo-correlation, are processed automatically to create scene-based DEMs. Outliers and any remaining poor values are deleted, and then the selected data is summed to provide the final pixel value. Each pixel's elevation has been determined using the ASTER DEM. After correcting any remaining abnormalities, the data is divided into 1×1-degree tiles. Even after urbanization, the elevation of the Earth's surface remains relatively constant throughout time and does not significantly vary [47, 48].
LULC: There isn't a single, optimum approach for classifying LULC. Even when an objective numerical technique is utilized, there are several points of view in the categorization process, and the process itself has a tendency to be subjective. Since LULC patterns fluctuate to meet changing needs for natural resources, there is no logical reason to believe that a single complete list should be enough for more than a short period. Few consumers will be content with an inventory that does not satisfy most of their wants since each categorization is designed to fit the user's needs. Multiple categorization methods are included in the MODIS land cover type product, which describes the characteristics as determined by observations covering a year's worth of Terra data [47, 49]. The fundamental land cover system defines 17 land cover classes, including 11 natural vegetation classes, 3 developed and mosaicked land classes, and 3 non-vegetated land classes, as the international geosphere-biosphere programme (IGBP) specified. All of the LULC's descriptions are included in *Table 4.*

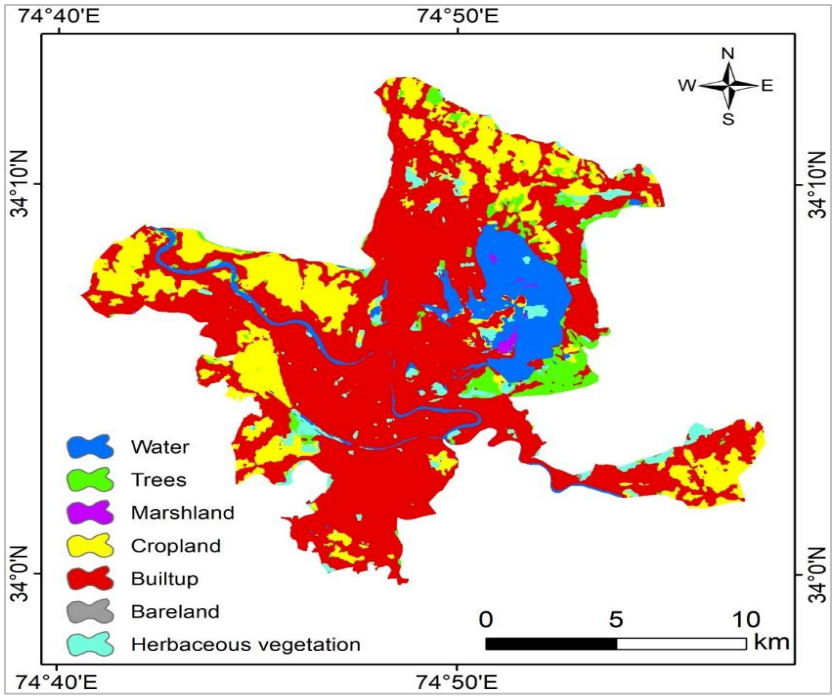

Figure 1 Urban built-up in the study area

Table 4 LULC classification as per international standards

LULC class value	Name	Description
1	Evergreen Needle leaf forests	Conifer trees (Tree cover > 60%)
2	Evergreen Broad leaf forests	Evergreen Broad leaf (Tree Cover > 60%)
3	Deciduous Needle leaf forests	Larch trees (Tree Cover > 60%)
4	Deciduous Broad leaf forests	Broadleaf trees (Tree Cover > 60%)
5	Mixed Forests	Neither Broad leaf nor Needle leaf (Tree Cover > 60%)
6	Closed Shrub-lands	Woody Perennials (Tree Cover > 60%)
7	Open Shrub-lands	Woody Perennials (Tree Cover > 10%-60%)
8	Woody Savannas	Tree Cover > 10%-60%
9	Savannas	Tree Cover > 10%-30%
10	Grass Lands	Herbaceous annuals
11	Permanent Wetlands	Inundated lands (30-60% water) > 10% Vegetation cover
12	Croplands	>60% area cultivated land
13	Urban and Built-up Lands	>30% including buildings and Vehicles
14	Natural Vegetation	Small Scale cultivation (40-60%) including natural tree, shrubs or herbaceous vegetation
15	Snow and Ice	>60% area covered by snow and ice > 10 months of the year
16	Barren	>60 % land is barren
17	Water bodies	>60% area covered by water bodies
255	Not Classified	Unclassified data

3.2Proposed model

The primary objective of this section is to suggest research that will use the LSTM technique to forecast the link between LST and LULC. The methodology adopted is depicted in *Figure 2*, and the study's block diagram is given in *Figure 3*. The LST data has been extracted from MODIS satellite imagery. It has been processed further to get the actual temperature by downscaling (DN × 0.02). The satellite images with 1718 defined projection have been clipped by a vector of Srinagar (boundary shape coordinates), and the LST data of the whole study area, i.e., Srinagar, have been processed. The data for relative parameters (elevation and aspect) have been acquired from the DEM Tool. (30 m SRTM and 1 km GTOPO) MODIS data have been resampled to 30 m to quantify the impact of resolution in predicting UHI. As such, the UHI information for Srinagar city would be computed at

two spatial details- the coarse resolution (1 km grid) and the finer resolution (30 m grid).

The approach used in this study automatically assesses the impact of LST on the extent of land cover by using the tabular LST data as input to the model. When utilizing LSTM, it has been shown that LSTM outperforms other forecasting-based models. It is constructed that the vanishing gradient problem is nearly eliminated while the training model is unchanged.

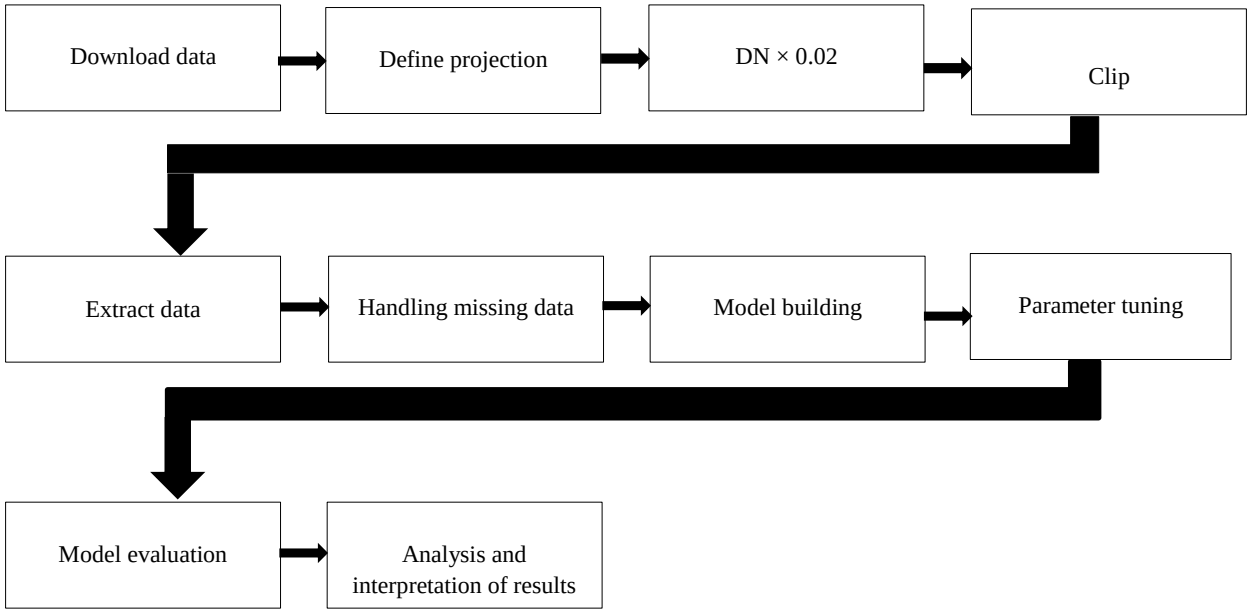

Figure 2 Methodology/flowchart depicting the flow of the work

LSTMs can deal with noise, distributed representations, and continuous values in addition to a long time lags inherent in specific problems. In contrast to the hidden Markov model (HMM), LSTMs do not require a fixed number of initial states (HMM). With LSTMs, we have access to various tuning options, including learning rates, input biases, and output biases. Thus, there is no requirement for fine-tuning. Advantages of LSTMs include a weight update complexity of O (1), like that of back propagation through time (BPTT), and the ability to learn from past examples. Unlike RNNs, LSTMs have a hidden layer that is a gated unit or gated cell, as shown in *Figure 4*. This is the primary distinction between the two architectures. It has four layers that work together to generate the cell's output and its current state. These two pieces of information are then relayed to the next covert level. When compared to RNNs, LSTMs have three logistic sigmoid gates and one tanh layer, while RNNs only have one tanh layer. As a result of the need to regulate the flow of data through a cell, gates have been implemented. They decide what data should be passed on to the next cell and what can be discarded. The output is typically a number between 0 and 1, where 0 indicates complete exclusion and 1 is complete inclusion [50, 51].

The proposed research presents a novel approach to forecasting the link between LULC and LST using LSTM. This study stands out in several aspects, offering unique contributions to the existing body of knowledge in the field.

Firstly, unlike previous studies that primarily rely on traditional statistical models or simple regression techniques, our approach leverages the power of LSTM, a type of RNN specifically designed to handle sequential data. By employing LSTM, we can capture the temporal dependencies and long-term patterns inherent in LULC and LST data, enabling more accurate and reliable forecasts.

Secondly, our research addresses the limitations of earlier studies that often overlook the impact of resolution in predicting UHI effects. We resampled the MODIS data to a finer resolution of 30 meters, allowing us to analyze the impact of resolution on UHI prediction quantitatively. This consideration of spatial details at different resolutions enhances the

precision and comprehensiveness of our findings. Furthermore, our approach incorporates additional features beyond LST and LULC, such as elevation and aspect, obtained from DEM tools. By combining these relative parameters, we can capture the topographic characteristics of the study area, which may further contribute to the understanding of the relationship between LULC and LST. This comprehensive feature set enhances the robustness and richness of our forecasting model.

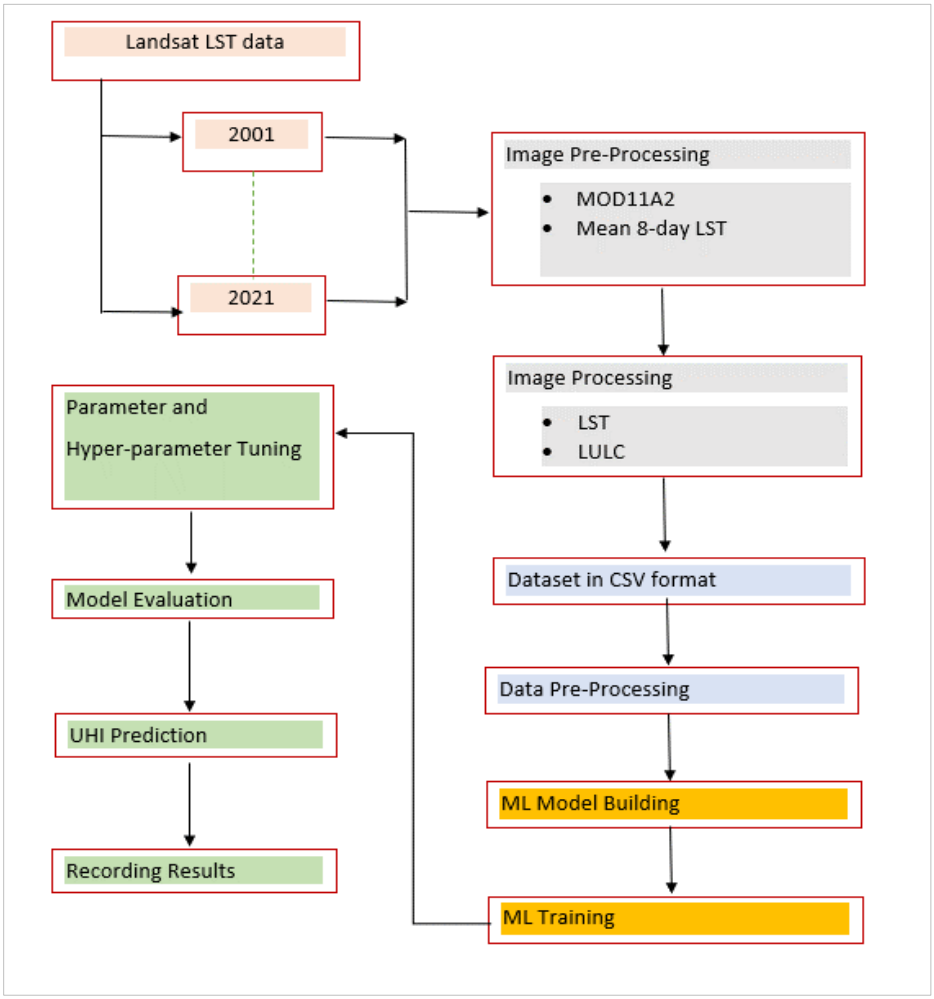

Figure 3 Block diagram of the study

Lastly, our research extends beyond mere prediction and emphasizes the experimental analysis of the relationship between LULC and LST. We conduct extensive experiments to investigate the impact of various factors on the LST dynamics, including different seasons, periods, and changes in LULC attributes. Through this analysis, we gain deeper insights into the complex interactions and dynamics between LULC and LST, thereby contributing to a more comprehensive understanding of the phenomena. The novel aspects of our approach, including the utilization of LSTM, consideration of resolution impact, incorporation of additional features, and the emphasis on experimental analysis, differentiate our study from previous works. These distinctive features enable us to provide more accurate predictions and a deeper understanding of the link between LULC and LST, thereby advancing the existing knowledge in the field.

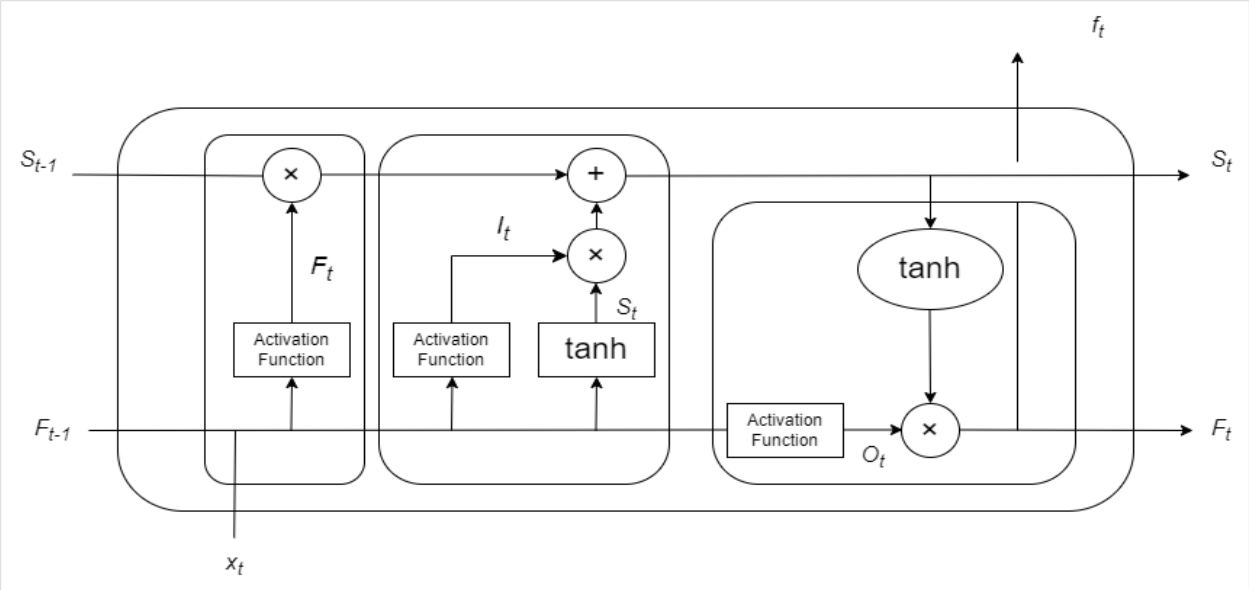

Figure 4 LSTM Architecture and hidden layers

The parameters and hyperparameters of a tuned LSTM model are given next.

Parameters:

- **Input data:** The input data is the sequence of data that is being processed.
- **Hidden state:** The hidden state represents the memory of the network. It is updated at each time step and used to make predictions.
- **Cell state:** The cell state is the long-term memory of the network. It is used to store relevant information from previous time steps.
- **Output:** The output of the LSTM model is the prediction made by the network.

Hyperparameters:

- **Number of LSTM units**: This is the number of LSTM cells in the network. It determines the network's power, and can affect its ability to learn complex patterns.
- **Learning rate:** This is the rate at which the network updates its parameters during training. A high learning rate can cause the network to converge too quickly and result in poor performance, while a low learning rate can result in slow convergence.
- **Dropout rate:** This is the fraction of input units randomly dropped out during training. It can help prevent overfitting by forcing the network to learn more robust features.
- **Batch size:** number of samples processed in each training batch, often referred to as the batch size, can significantly influence the training process. A larger batch size can lead to faster convergence as more samples are processed simultaneously, but it also requires more memory. Balancing the batch size is crucial to efficiently train a model within the available computational resources.
- **Number of epochs:** The number of times the network sees the entire training dataset determines of how much training the network receives and can affect the model's performance.

To ensure transparency and reproducibility, the values of the parameters and hyperparameters utilized in the LSTM model are presented in *Table 5*. This allows for a clear understanding and potential replication of the model's configuration.

Table 5 Values used for parameters and hyperparameters

Parameter	
Name	**Value**
Input data	Sequence of data
Hidden state	Updated at each step
Cell state	Long-term memory
Output	Prediction
Hyperparameters	

Name	Value	
Number of LSTM units	128	The values change iteration wise
Learning rate	0.001	
Dropout rate	0.2	
Batch size	32	
Number of epochs	50	
Number of LSTM units	128	
Learning rate	0.001	
Dropout rate	0.2	
Batch size	32	

3.2.1Working mechanism

The working mechanism of the proposed technique, including how experimentation was carried out to forecast the link between LULC and LST using LSTM, is described step by step as follows:

1) **Data acquisition**: The first step is to acquire the necessary data for your study. Regarding LULC and LST forecasting, that LST data has been extracted from MODIS satellite imageries, and data for relative parameters such as elevation and aspect have been acquired from DEM tools.
2) **Data pre-processing:** Once the data is acquired, it must be pre-processed to prepare it for the LSTM model. This pre-processing step may include cleaning data, handling missing values, normalizing the data, and splitting it into training and testing sets.
3) **LSTM architecture:** LSTM is a type of RNN architecture designed to overcome the vanishing gradient problem in traditional RNNs, enabling the modeling of long-term dependencies in sequential data. The LSTM network comprises memory cells as shown in *Figure 4*, each with a cell state and three multiplicative gates: input, forget, and output gates. These gates regulate the flow of information, allowing the network to selectively update and retain information over time. The input gate governs the flow of new information into the cell state, the forget gate controls the removal of irrelevant information from the cell state, and the output gate determines the information to be passed to the next time step. The ability to selectively store and retrieve information over extended sequences makes LSTMs adept at capturing complex patterns and dependencies in temporal data, such as in natural language processing and time series analysis. This architecture excels in handling long-range dependencies, mitigating the challenges posed by the short-term memory limitations of traditional RNNs, and has become a cornerstone in various applications requiring sequential modeling..
4) **Training the LSTM model:** The pre-processed data is used to prepare the LSTM model. The model is presented with input sequences, which would include the features related to LULC and LST and the corresponding target variable (e.g., LULC class or LST value). The model learns to predict the target variable based on the input sequences by adjusting its internal weights through backpropagation.
5) **Hyperparameter tuning:** Hyperparameters are configuration choices that determine the , behavior, and performance of the LSTM model. Examples of hyperparameters include the number of LSTM units, learning rate, dropout rate, batch size, and number of epochs. Hyperparameter tuning involves selecting the optimal values for these parameters to improve the model's performance. This can be done using techniques like grid search or random search.
6) **Performance measures:** The performance measures used to assess the model are defined as follows:

- **Accuracy:** Accuracy is a measure of the overall correctness of the model. It calculates the ratio of correct predictions to the total number of predictions made by the model. A higher accuracy indicates better performance. Accuracy computed as:

  Accuracy= Total Number of Predictions/Number of Correct Predictions
- **Cohen's kappa score**: Cohen's Kappa measures the agreement between two sets of categorical data, while accounting for the possibility of agreement occurring by chance. It is particularly useful when there is an imbalance in class distribution. Cohen's Kappa is calculated as (Equation 1):

  $\kappa = Po - Pe / 1 - Pe$ (1)

where Po is the observed agreement and Pe is the expected agreement

- **F1-score:** F1-score is the harmonic mean of precision and recall. It provides a balance between the two, making it a suitable metric for binary classification problems, especially when there is an imbalance between the classes. It is computed as (Equation 2):

  $F1 = 2 \times Precision \times Recall / Precision + Recall$ (2)

- **Error**: Error, complementary to accuracy, measures the proportion of incorrect predictions made by the model. A lower error rate corresponds to better model performance. Error is computed as (Equation 3):

  Error= FP+FN/TP+TN+FP+FN (3)

7) **Evaluation and validation:** The trained model also needs to be evaluated and validated. This involves using the testing set to assess the model's performance in forecasting the link between LULC and LST. Standard evaluation metrics for classification tasks include accuracy, precision, recall, and F1 score. While in contrast, regression tasks can be evaluated using metrics such as MSE or MAE.

3.2.2 Models used for comparison

The proposed model was compared with the following widely used algorithms in the field of LULC and LST prediction are as follows:

RF: RF is an ensemble learning method that combines multiple decision tree (DTs) to make predictions. It has been widely utilized in remote sensing and geospatial analysis tasks, including LULC and LST prediction. Its ability to handle complex relationships and capture nonlinear patterns makes it popular.

SVM: SVM is a supervised learning algorithm that can be used for classification and regression tasks. SVM has been successfully applied in various remote sensing applications, including land cover classification and temperature prediction. It is good at handling high-dimensional data and nonlinear relationships, making it suitable for LULC and LST prediction.

Convolutional neural networks (CNN): CNN is a deep learning algorithm shown remarkable success in image analysis tasks. CNN has been widely used in remote sensing applications for land cover classification and feature extraction. It can automatically learn hierarchical representations from input data and, is a potential candidate for LULC and LST prediction.

Decision tree (DT): DT are simple yet powerful ML models that can be used for classification and regression tasks. DT have been extensively applied in remote sensing for land cover mapping and LST estimation. Their interpretability and ease of use make them popular in geospatial analysis.

K-nearest neighbour (KNN): KNN is a non-parametric algorithm that makes predictions based on the similarity between input samples. KNN has been used in various remote sensing applications, including land cover classification and temperature estimation. Its simplicity and effectiveness in handling spatial relationships make it suitable for LULC and LST prediction.

4. Results

Only four LULC classes were obtained from the dataset's 17 LULC classes throughout the previous two decades (2001, 2011, and 2021) using the SVM technique. Thus, these 17 groups are subdivided into three categories: built-up regions, water bodies, and undeveloped terrain. To check the accuracy of the categorized LULC maps, 328 random points were collected for each year. The total accuracy (percentage) was assessed to be more than 91% for all years (*Table 6*). Furthermore, the kappa and F1 scores yielded promising results, with values better than 93 and 90, respectively [52].

Two distinct altering tendencies were detected, including a rise in urbanization and declines in undeveloped soil, green cover, and water bodies. From 2001 to 2021, the built-up area rose by 9.86%, with a 0.49% annual change rate. Between 2001 and 2021, vegetation covers and water bodies dropped by 3.16% and 0.94%, respectively, with a yearly decline rate of 0.158% and 0.04%, as shown in the below table (*Table 7*). As shown in *Table 7* there is a considerable amount of existing infrastructure, and unplanned growth has significantly altered the ratio of built-up areas to natural ones. From 2001-2007 (8.92%, 4.1%), 2008-2014, and 2015-2021, there was a drastic increase in the ratio of impervious surfaces to total land area (7.76%, 3.41%). From 2001 to 2021, the city expanded its built-up areas by 29.15 km^2, increasing the percentage of settled land in the metro area to 64.39% from its natural setting of lakes, forests, and open fields.

Table 6 Assessment of LULC classification

Years (Group)	Accuracy statistics in %			Accuracy	Cohens Kappa Score	F1 Score
	Built-up areas	Vegetation cover	Water bodies			
2001-2007	87.44%	91.49%	91.34%	91.22%	0.932	0.901
2008-2014	91.34%	89.33%	89.87%	92.16%	0.938	0.919
2015-2021	89.45%	85.66%	90.21%	91.43%	0.941	0.921

Table 7 Area change in LULC classification

LULC category	Geographical area in KMs			%age of Change	Annual change
	2001-2007	2008-2014	2015-2021	2021-2001	2021-2001
Built-up areas	72.59	83.44	88.12	8.92%	0.52%
Vegetation cover	42.92	32.98	29.23	-4.25%	-0.14%
Water bodies	6.67	7.23	7.47	-0.87%	-0.07%

4.1LST change variations across Srinagar City

The highest and lowest temperatures were collected from the SDA for 2001, 2010, and 2021, respectively, to cross-check the tabular data's projected LST values. SDA weather station data was used to calculate the differences. If the LST deviation is negative, then the assessed value of the data is higher than the observed value. Post deviation, on the other hand, shows that the estimated value is less than the actual value. Temperature extremes for 2021 were a maximum of -3.53 degrees Celsius and a minimum of -4.56 degrees Celsius. As shown in *Table 8*, the average maximum and minimum deviations for 2001, 2010, and 2021 were -2.14 °C, -2.52 °C, and -4.04 °C, respectively.

Table 8 Average maximum and minimum deviations of seasonal thermal bands

Year	LST					
Estimated vs. Recorded Temperatures (LST)	2001		2010		2021	
	Max	Min	Max	Min	Max	Min
Estimated (LST) ˚C	34.85 (308 K)	8.85 (282 K)	35.96 (309.11)	9.06 (282.2 K)	39.29 (312.1 K)	11.94 (285.1 K)
Observed/ Recorded LST ^{O}C	33.74 (306.8 K)	5.67 (278.8 K)	34.88 (308.1 K)	5.10 (278.2 K)	35.76 (308.9 K)	7.38 (280.5 K)
Deviation	-1.11 (272.1 K)	-3.18 (269.9 K)	-1.08 (272.1 K)	-3.96 (269.1 K)	-3.53 (272.6 K)	-4.56 (277.7 K)
Average Deviation	-2.14 (271.0 K)		-2.52 (270.6 K)		-4.04 (267.1 K)	

The more minor discrepancy between the estimated and LST measured may still be accepted and used for exploration in UHI studies, such as LST prediction, despite the limitations the real-world situation imposed on the estimated data values.

4.2LST class changes across Srinagar city

Several equations were used to compute LST's spatial and temporal distribution over the study period using MODIS thermal bands. The increasing trends in LST are shown graphically in *Figure 5*, which displays the annual distribution of LST from 2001 to 2021. With a 0.22°C yearly increase, the highest recorded temperature of 34.85 °C in 2001 will soar to 39.29 °C in 2022. Similarly, 2001 saw the lowest temperature at 8.85 °C, and 2021 will see it at 11.94 °C, a change of 0.09 °C annually.

The majority of the study area (68.76%) was subjected to temperatures between 22 and 25 degrees Celsius in 2001 (109.25 km^2). The percentage of the study area with temperatures between 23 and 27 degrees Celsius increased from 37.52% in 2010 to 65.24% in 2021, with 19.65% of the study area experiencing temperatures between 27 and 30 degrees Celsius. Under this scenario, surface temperatures in the study region are expected to rise dramatically. While only 1.99% and 5.36% of the area, respectively, experienced high temperatures in 2010 and 2021, it is that this percentage will have a significant increase in the years to come if the current trend continues. The explanation above makes it abundantly clear that the LST has risen dramatically over the past two decades [53].

The UHI effect has worsened the proliferation of LULC transitions, which have had a significant bearing on the distribution and concentration of vital critical locations. For this reason, UHI mitigation relies heavily on careful land use planning and management. This was demonstrated in [22], in which researchers discovered that land use and management alterations have an equal impact on heat transfer (surface temperatures).

4.3Predicted LST and LULC for 2025

The year 2025 prospects were calculated by extrapolating the trend. The urbanization growth, as the data reflects, continues without any planned action. 67% (62.27% in 2021) of the expanding urban areas will be concentrated in the Srinagar region, with built-up areas replacing both undeveloped land and vegetation cover. Additionally, vegetation cover decreased by a significant 11.23 percentage points

compared to the 2021 average of 15.54 percentage points. The percentage of developed land in the simulated LULC scenario would increase by 25.88% compared to the baseline year of 2001, followed by a dramatic decrease of 10.23% in undeveloped land, 8.34% in plant cover, and 2.41% in water bodies.

A city's ecological services, urban health, and thermal elements may all suffer if plant cover is reduced and urbanization is accelerated. Suppose the current trend of uncontrolled urban development continues, it will exacerbate adverse health, economic, and ecological impacts in the study area. A more environmentally sustainable Srinagar could be achieved through measures such as proper land-use planning, the protection of water bodies, afforestation, and an increase in urban greenery [54]. Based on the historical LST datasets, this study made predictions for 2025. Maximum and minimum LSTs of 45.21 °C and 18.35 °C, respectively, were predicted by the scenario to be found in the Srinagar area, with higher temperatures concentrated in urban areas.

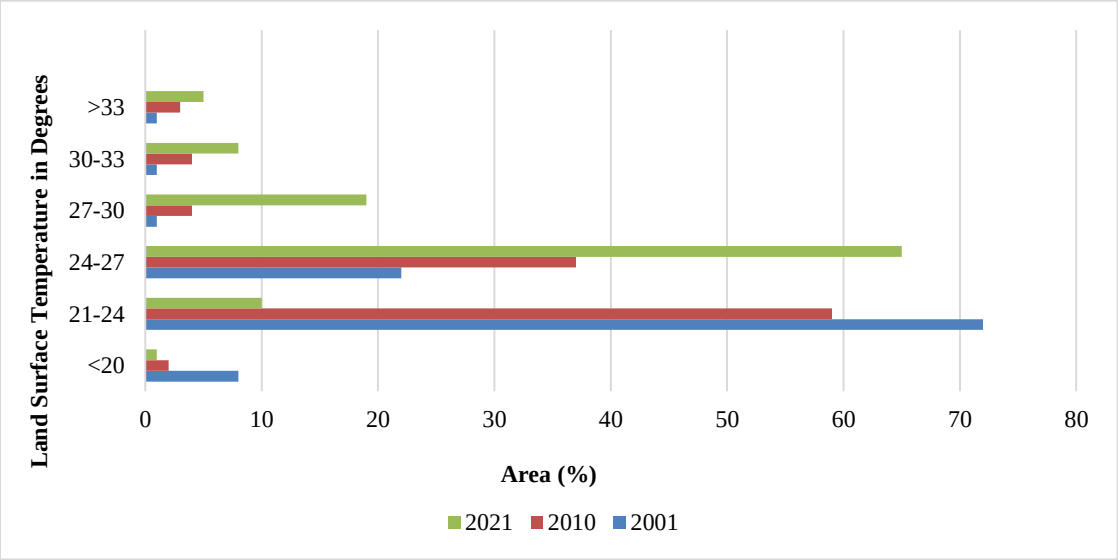

Figure 5 LST distribution of Srinagar city from 2001-2021

5. Discussion

This research aims to determine the impact of LST data on LULC classes and probable LULC and LST change scenarios in Srinagar, Jammu & Kashmir, India. The study looked at the effects of LST temperatures on various LULC classes, and it estimated that if the trend of developing urban infrastructure from 2001 to 2021 continues, urban areas will have grown by 23% by 2025.

5.1Overall analysis and impact of the results

The results of the projections show that LST has been increasing over the past few decades (2001–2021), with urban areas being the primary cause of this trend. Further UHI impacts will be exacerbated by expanding urban footprints and a shocking reduction in plant cover. Changes in the greenhouse effect, global warming, and surface features would be possible causes of the temperature increase in a world without urbanization. The expected LST reflected the threats of the current trend's warming, including amplified UHI effects. Increases in energy use and its allied ill impact air pollution contribute to the threat that UHI poses. Rising levels of greenhouse gases are particularly harmful to human health, as well as to the city's ecological responsibility and urban health standards. Our research found significant changes in LULC patterns over the last 20 years, with a substantial increase in built-up regions and a decrease in plant cover and water bodies. *Table 7* summarizes the percentage changes and annual change rates for each LULC group. Our data show that built-up areas have increased by 9.86%, while vegetation cover and water bodies have decreased by 3.16% and 0.94%, respectively.

5.2Comparative analysis

In this section, we compare our models to those of other ensemble and traditional methods like DT, k nearest neighbor (KNN), SVM, RF[55], logistic regression (LR), and neural network (NN)[56,57] that have been used on the same dataset in this study

(*Table 9*). Total performance was calculated by applying both sets of methods to the same collection of data from the Kashmir province. Data and time requirements for these ensemble approaches were shown to be higher when compared to the LSTM model. The proposed hybrid model is defined by its MSE value being between (2.012 and 0.189) percent, while the total performance of these traditional and ensemble ML methods is between (31% and 80%).

Table 9 Comparison of algorithms on the UHI dataset of Srinagar City

Algorithms	**DT**	**KNN**	**SVM**	**RF**	**LR**	**NN**	**LSTM**
Cohen's kappa	0.562	0.192	0.047	0.671	0.648	0.571	Regression (R):-Testing: 0.897%
F1-Score	0.612	0.234	0.165	0.765	0.832	0.639	Validation: 0.912%
Error	0.276	0.723	0.791	0.318	0.318	0.298	Training: 0.931%
Accuracy	0.751	0.360	0.310	0.802	0.789	0.791	MSE:-Testing: 2.012%
							Validation: 0.189%
							Training: 1.124%

When applied to time series data, LSTM effectively separates it into conclusive and error sequence. Several studies have shown off performance levels when using classical and ensemble methods. When these models are considered, the performance accuracy remains in the middle range (31–80%). DT, KNN, SVM, RF, LR, and NN are just some of the algorithms whose performance results are displayed in Table 9. Predictive solid performance has been achieved by RF, LR, and NN, as demonstrated by their respective Cohen's kappa and F1 score values. The prediction accuracy, with data mentioned, for KNN and SVM has plummeted to 0.28 and 0.29, respectively. Therefore, these two algorithms do not work with the data that has been provided. Once the LSTM model was put into place, however, it was shown that the accuracy statistics improved, with the MAE also decreasing slightly. The implementation's finding that resolution stagnation is falsely improved the LSTM model's output. The SVM approach achieved an overall accuracy of more than 91%, a kappa score greater than 93%, and an F1 score greater than 90%.

The exceptional performance of SVM is attributed to their capability to manage nonlinear data and high-dimensional feature spaces, characteristics commonly found in remote sensing datasets. Additionally, the SVM algorithm can efficiently handle small sample sizes and noisy data, conditions frequently encountered in remote sensing applications.

Based on our findings, we have detected significant changes in the LULC patterns over the last two decades. To further elaborate on the results, we calculated by the percentage changes in the LULC categories and their annual change rates, presented in *Table 7*. Our results indicate that the built-up areas have increased by 9.86%, whereas vegetation cover and water bodies have decreased by 3.16% and 0.94%, respectively. These changes have occurred at an annual rate of 0.49%, 0.158%, and 0.04%, respectively.

To provide a better experience of these results, we have discussed the reasons behind the observed changes, which include urbanization, infrastructure development, and unplanned growth. These factors have led to a rise in the ratio of built-up areas to natural ones, as reflected in *Table 7*. Our results show that the city has expanded its built-up areas by 29.15 km^2 over the last two decades, which has resulted in a shift from its natural setting of lakes, forests, and open fields to a more urbanized environment.

In addition, to present a comprehensive evaluation of our proposed LSTM based model and demonstrate its effectiveness compared to other state-of-the-art methods, we conducted a comparative study with five widely used algorithms i.e., RF, SVM, CNN, DT, and KNN in the LULC and LST prediction [55–57]. These algorithms were carefully selected based on their relevance and popularity in remote sensing and geospatial analysis tasks.

Each of these algorithms has been extensively applied in remote sensing for various tasks, such as land cover classification and temperature estimation. By comparing our LSTM based model against these established algorithms, we aimed to gain insights into its performance and distinguish its unique contributions. The comparative analysis was conducted on multiple datasets, encompassing diverse geographical regions and temporal scales. Evaluation metrics such as accuracy, precision, recall, F1 score, and computational efficiency were utilized to assess the performance of each algorithm. Furthermore, we considered the interpretability and ease of implementation in real-world scenarios. Our LSTM based model exhibited superior performance

in terms of accuracy and precision, outperforming the other algorithms in most of the evaluation metrics. Its ability to capture temporal dependencies and handle long-term memory made it particularly effective in LULC and LST prediction tasks. Additionally, the model demonstrated promising results regarding interpretability and computational efficiency, making it a practical choice for real-world applications. Overall, the comparative analysis highlighted the strengths and weaknesses of different methodologies, emphasizing the novelty and superiority of our proposed LSTM based model in addressing the challenges associated with LULC and LST prediction. By showcasing its comparative advantages, the position of this innovative approach is established, pushing the boundaries of current research in this field.

5.3Limitations

The study area, Srinagar city, experiences significant cloud/snow cover during the winter, which can impede the creation of a complete UHI dataset. This poses a challenge for researchers to work with structured and voluminous data. Retrieving LST remains a challenging task, although it is modeling key critical parameter in modeling the UHI effect.

A complete list of abbreviations is shown in *Appendix I.*

6. Conclusion and future work

Extensive research is underway utilizing ML and DL for UHI data assessment to model the relationship between LULC and LST. The proposed study identified that inconsistencies in the dataset and inaccurate delineation of wetlands by geoscientists could contribute to the wide range of performance predictions generated by the algorithms. To establish the dataset's base on LST, other relevant criteria such as AR and elevation must be considered. Regression or correlation analyses can be performed to test the correlation between independent variables and LST data. The study aims to clarify the effect of LST data on LULC classes and potential LULC and LST change scenarios in Srinagar city, Jammu and Kashmir, India. The impact of LST temperatures on various LULC classes was examined, and predictions indicate that if the trend of increasing urban infrastructure from 2001 to 2021 continues, urban areas will have expanded by 23% by 2025. This expansion is expected to destroy 9.2% of vegetated land and 3.1% of aquatic bodies. Additionally, the study found that temperatures have risen by an average of 1.89°C over the past two decades and are expected to increase further by 2.01°C between 2021 and 2025.

To mitigate the effects of UHI and regulate the urban microclimate, the study suggests building vertical infrastructure, implementing a plantation strategy, and developing urban green parks on previously undeveloped land. Urban planners should consider the amount of sunlight reaching the ground and its effect on reducing surface temperatures. Strategies such as using light-colored roofs, planting large canopies of shade trees, and creating vertical and horizontal green spaces inside buildings are recommended to reduce the effects of UHI.

The study reveals notable changes in the LULC patterns over the past 20 years, with a considerable increase in built-up areas and a decrease in vegetation cover and water bodies. Specifically, the findings indicate that built-up areas have grown by 9.86%, while vegetation cover and water bodies have declined by 3.16% and 0.94%, respectively. The annual rates of change for these categories are 0.49%, 0.158%, and 0.04%, respectively. To better explain these results, the contributing factors, such as urbanization, infrastructure development, and unregulated expansion, have been explored. These factors have led to an increase in the proportion of built-up areas compared to natural ones. The results demonstrate that the city has expanded its built-up areas by 29.15 km² during the last two decades, transforming from a natural setting of lakes, forests, and open fields to a more urbanized landscape. The findings of this research will assist policymakers in limiting urban sprawl and mitigating the effects of UHI.

Acknowledgment

This article uses MODIS data.

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contribution statement

Mujtaba Shafi: Conceptualization, investigation, data curation, design, writing – original draft, writing – review and editing. **Amit Jain:** Design, supervision, analysis and interpretation of results, writing – review and editing.
Majid Zaman: Supervision, investigated challenges and draft manuscript preparation, writing – review, and editing.

References

[1] Chakroborty S, Al RA, Al KA. Monitoring water quality based on community perception in the northwest region of Bangladesh. In 1st international student research conference-2020 (pp.1-8).

[2] Dar I, Qadir J, Shukla A. Estimation of LST from multi-sensor thermal remote sensing data and evaluating the influence of sensor characteristics. Annals of GIS. 2019; 25(3):263-81.
[3] Kafy AA, Rahman AF, Al RA, Akter KS, Raikwar V, Jahir DM, et al. Assessment and prediction of seasonal land surface temperature change using multi-temporal landsat images and their impacts on agricultural yields in Rajshahi, Bangladesh. Environmental Challenges. 2021; 4:1-21.
[4] Mishra K, Prasad P. Automatic extraction of water bodies from landsat imagery using perceptron model. Journal of Computational Environmental Sciences. 2015; 2015:1-10.
[5] Latham JS, He C, Alinovi L, Digregorio A, Kalensky Z. FAO methodologies for land cover classification and mapping. Linking People, Place, and Policy: a GIScience Approach. 2002:283-316.
[6] Qiao Z, Tian G, Zhang L, Xu X. Influences of urban expansion on urban heat island in Beijing during 1989–2010. Advances in Meteorology. 2014; 2014:1-12.
[7] Shao B, Zhang M, Mi Q, Xiang N. Prediction and visualization for urban heat island simulation. Transactions on Edutainment VI. 2011:1-11.
[8] Naim MN, Kafy AA. Assessment of urban thermal field variance index and defining the relationship between land cover and surface temperature in Chattogram city: a remote sensing and statistical approach. Environmental Challenges. 2021; 4:100107.
[9] Zhou J, Chen Y, Zhang X, Zhan W. Modelling the diurnal variations of urban heat islands with multi-source satellite data. International Journal of Remote Sensing. 2013; 34(21):7568-88.
[10] Oke TR. The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society. 1982; 108(455):1-24.
[11] Yonezawa G. Generation of DEM for Urban transformation of Hanoi, Vietnam. Kyoto Working Papers on Area Studies: G-COE Series. 2009; 60:1-10.
[12] Rousta I, Sarif MO, Gupta RD, Olafsson H, Ranagalage M, Murayama Y, et al. Spatiotemporal analysis of land use/land cover and its effects on surface urban heat island using landsat data: a case study of Metropolitan City Tehran (1988–2018). Sustainability. 2018; 10(12):1-25.
[13] Song C, Woodcock CE, Seto KC, Lenney MP, Macomber SA. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sensing of Environment. 2001; 75(2):230-44.
[14] Li H, Zhou Y, Jia G, Zhao K, Dong J. Quantifying the response of surface urban heat island to urbanization using the annual temperature cycle model. Geoscience Frontiers. 2022; 13(1):101141.
[15] Giannopoulou K, Livada I, Santamouris M, Saliari M, Assimakopoulos M, Caouris YG. On the characteristics of the summer urban heat island in Athens, Greece. Sustainable Cities and Society. 2011; 1(1):16-28.
[16] Jiménez-muñoz JC, Sobrino JA. A generalized single-channel method for retrieving land surface temperature from remote sensing data. Journal of Geophysical Research: Atmospheres. 2003; 108(D22):1-9.
[17] Dimoudi A, Zoras S, Kantzioura A, Stogiannou X, Kosmopoulos P, Pallas C. Use of cool materials and other bioclimatic interventions in outdoor places in order to mitigate the urban heat island in a medium size city in Greece. Sustainable Cities and Society. 2014; 13:89-96.
[18] Mirzaei PA, Haghighat F, Nakhaie AA, Yagouti A, Giguère M, Keusseyan R, et al. Indoor thermal condition in urban heat Island–development of a predictive tool. Building and Environment. 2012; 57:7-17.
[19] Amiri R, Weng Q, Alimohammadi A, Alavipanah SK. Spatial–temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. Remote Sensing of Environment. 2009; 113(12):2606-17.
[20] Prata AJ. Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 1. theory. Journal of Geophysical Research: Atmospheres. 1993; 98(D9):16689-702.
[21] Zha Y, Gao J, Ni S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing. 2003; 24(3):583-94.
[22] Zhang H, Qi ZF, Ye XY, Cai YB, Ma WC, Chen MN. Analysis of land use/land cover change, population shift, and their effects on spatiotemporal patterns of urban heat islands in metropolitan Shanghai, China. Applied Geography. 2013; 44:121-33.
[23] Weng Q. A remote sensing? GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. International Journal of Remote Sensing. 2001; 22(10):1999-2014.
[24] Weng Q. Fractal analysis of satellite-detected urban heat island effect. Photogrammetric Engineering & Remote Sensing. 2003; 69(5):555-66.
[25] Hachem S, Duguay CR, Allard M. Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain. The Cryosphere. 2012; 6(1):51-69.
[26] Perini K, Magliocco A. Effects of vegetation, urban density, building height, and atmospheric conditions on local temperatures and thermal comfort. Urban Forestry & Urban Greening. 2014; 13(3):495-506.
[27] Lee L, Chen L, Wang X, Zhao J. Use of Landsat TM/ETM+ data to analyze urban heat island and its relationship with land use/cover change. In international conference on remote sensing, environment and transportation engineering 2011 (pp. 922-7). IEEE.
[28] Bogoliubova A, Tymków P. Accuracy assessment of automatic image processing for land cover

classification of St. Petersburg protected area. Acta Scientiarum Polonorum. Geodesia et Descriptio Terrarum. 2014; 13(1-2):5-22.
[29] Peron F, De MMM, Spinazzè F, Mazzali U. An analysis of the urban heat island of Venice mainland. Sustainable Cities and Society. 2015; 19:300-9.
[30] Wong MS, Nichol JE. Spatial variability of frontal area index and its relationship with urban heat island intensity. International Journal of Remote Sensing. 2013; 34(3):885-96.
[31] Yankovich KS, Yankovich EP, Baranovskiy NV. Classification of vegetation to estimate forest fire danger using landsat 8 images: case study. Mathematical Problems in Engineering. 2019; 2019:1-15.
[32] Tran DX, Pla F, Latorre-carmona P, Myint SW, Caetano M, Kieu HV. Characterizing the relationship between land use land cover change and land surface temperature. ISPRS Journal of Photogrammetry and Remote Sensing. 2017; 124:119-32.
[33] Tan J, Yu D, Li Q, Tan X, Zhou W. Spatial relationship between land-use/land-cover change and land surface temperature in the Dongting Lake area, China. Scientific Reports. 2020; 10(1):1-9.
[34] Kafy AA, Dey NN, Al RA, Rahaman ZA, Nasher NR, Bhatt A. Modeling the relationship between land use/land cover and land surface temperature in Dhaka, Bangladesh using CA-ANN algorithm. Environmental Challenges. 2021; 4:100190.
[35] Bokaie M, Zarkesh MK, Arasteh PD, Hosseini A. Assessment of urban heat island based on the relationship between land surface temperature and land use/land cover in Tehran. Sustainable Cities and Society. 2016; 23:94-104.
[36] Zhao ZQ, He BJ, Li LG, Wang HB, Darko A. Profile and concentric zonal analysis of relationships between land use/land cover and land surface temperature: case study of Shenyang, China. Energy and Buildings. 2017; 155:282-95.
[37] Weng Q, Lu D, Schubring J. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment. 2004; 89(4):467-83.
[38] Hamstead ZA, Kremer P, Larondelle N, Mcphearson T, Haase D. Classification of the heterogeneous structure of urban landscapes (STURLA) as an indicator of landscape function applied to surface temperature in New York City. Ecological Indicators. 2016; 70:574-85.
[39] Sun Q, Wu Z, Tan J. The relationship between land surface temperature and land use/land cover in Guangzhou, China. Environmental Earth Sciences. 2012; 65:1687-94.
[40] Guha S, Govil H, Gill N, Dey A. Analytical study on the relationship between land surface temperature and land use/land cover indices. Annals of GIS. 2020; 26(2):201-16.
[41] Irons JR, Dwyer JL, Barsi JA. The next Landsat satellite: the Landsat data continuity mission. Remote Sensing of Environment. 2012; 122:11-21.
[42] Acosta MP, Vahdatikhaki F, Santos J, Hammad A, Dorée AG. How to bring UHI to the urban planning table? a data-driven modeling approach. Sustainable Cities and Society. 2021; 71:102948.
[43] Nengroo ZA, Bhat MS, Kuchay NA. Measuring urban sprawl of Srinagar city, Jammu and Kashmir, India. Journal of Urban Management. 2017; 6(2):45-55.
[44] Liu P, Jia S, Han R, Liu Y, Lu X, Zhang H. RS and GIS supported urban LULC and UHI change simulation and assessment. Journal of Sensors. 2020; 2020:1-7.
[45] Dijoo ZK. Urban heat island effect concept and its assessment using satellite-based remote sensing data. Geographic Information Science for Land Resource Management. 2021: 81-98.
[46] Hu L, Brunsell NA. The impact of temporal aggregation of land surface temperature data for surface urban heat island (SUHI) monitoring. Remote Sensing of Environment. 2013; 134:162-74.
[47] Saputra MH, Lee HS. Prediction of land use and land cover changes for north Sumatra, Indonesia, using an artificial-neural-network-based cellular automaton. Sustainability. 2019; 11(11):1-16.
[48] Fu P, Weng Q. A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with landsat imagery. Remote sensing of Environment. 2016; 175:205-14.
[49] Stathopoulou M, Cartalis C, Petrakis M. Integrating corine land cover data and landsat TM for surface emissivity definition: application to the urban area of Athens, Greece. International Journal of Remote Sensing. 2007; 28(15):3291-304.
[50] Aslam B, Maqsoom A, Khalid N, Ullah F, Sepasgozar S. Urban overheating assessment through prediction of surface temperatures: a case study of Karachi, Pakistan. ISPRS International Journal of Geo-Information. 2021; 10(8):1-16.
[51] Li XM, Ma Y, Leng ZH, Zhang J, Lu XX. High-accuracy remote sensing water depth retrieval for coral islands and reefs based on LSTM neural network. Journal of Coastal Research. 2020; 102(SI):21-32.
[52] Feizizadeh B, Blaschke T, Nazmfar H, Akbari E, Kohbanani HR. Monitoring land surface temperature relationship to land use/land cover from satellite imagery in Maraqeh County, Iran. Journal of Environmental Planning and Management. 2013; 56(9):1290-315.
[53] Walawender JP, Szymanowski M, Hajto MJ, Bokwa A. Land surface temperature patterns in the urban agglomeration of Krakow (Poland) derived from landsat-7/ETM+ data. Pure and Applied Geophysics. 2014; 171:913-40.
[54] Mccarville D, Buenemann M, Bleiweiss M, Barsi J. Atmospheric correction of landsat thermal infrared data: a calculator based on North American Regional Reanalysis (NARR) data. In proceedings of the American society for photogrammetry and remote sensing conference, Milwaukee, WI, USA 2011:1-12.

[55] Sameh S, Zarzoura F, El-mewafi M. Automated mapping of urban heat island to predict land surface temperature and land use/cover change using machine learning algorithms: Mansoura City. International Journal of Geoinformatics. 2022; 18(6):1-21.
[56] Vergara DC, Blanco AC, Marciano JJJ, Meneses III SF, Borlongan NJ, Sabuito AJ. Assessing and modelling urban heat island in Baguio city using landsat imagery and machine learning. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2023; 48:457-64.
[57] Lan T, Peng J, Liu Y, Zhao Y, Dong J, Jiang S, et al. The future of China's urban heat island effects: a machine learning based scenario analysis on climatic-socioeconomic policies. Urban Climate. 2023; 49:101463.

Mujtaba Shafi is a Research Scholar at the University Institute of Computing, Chandigarh University, Punjab, India. He holds a Master's degree in Computer Applications (MCA). He also has a Bachelor's in Engineering (B.E) from Bangalore University, India.
Email: mujtabashafi@gmail.com

Dr. Amit Jain is an Associate Dean of Academic Affairs at Chandigarh University, Mohali. He has 18+ years of teaching experience in educational institutions across India. He is a member of several professional societies, including IEEE, CSI, IAENG, STRA, and IAOIP. He is currently working in the field of Data Sciences and is supervising 05 Ph.D. Research Scholars. His research interest includes Algorithms, Machine learning, Network Security, and Artificial Neural Networks. He currently has 15+ publications in renowned International Journals. Some of them are in Scopus Indexed Journals. Further, he has 20+ publications in several National and International Conferences. He has also published a Book Chapter in an internationally renowned publishing company, CRC, CRC, CRC, CRC Press. He a reviewer for various book titles from reputed publishers, including McGraw Hill, Pearson Education, Elsevier, etc.
Email: amit_jainci@yahoo.com

Dr. Majid Zaman is Scientist "E" in the Directorate of Information Technology & Support System, University of Kashmir, J&K, India. He holds a Ph.D. in Computer Science from the University of Kashmir, Srinagar, J&K, and an M.S. in Software Systems from the Birla Institute of Technology and Science (BITS), Pilani, India. In addition, he received a Bachelor's in Computer Science Engineering from BAMU, Mumbai, India.
Email: zamanmajid@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	AR	Aspect Ratio
2	ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
3	BPTT	Back Propagation Through Time
4	CA	Cellular Automata
5	CNN	Convolutional Neural Networks
6	DEM	Digital Elevation Model
7	DT	Decision Tree
8	FID	Feature Identification
9	GA	Genetic Algorithm
10	GIS	Geographic Information Systems
11	GTOPO	Global Topographic
12	HMM	Hidden Markov Model
13	IGBP	International Geosphere-Biosphere Programme
14	KNN	K Nearest Neighbor
15	LR	Logistic Regression
16	LST	Land Surface Temperature
17	LSTM	Long Short Term Memory
18	LULC	Land Use Land Cover
19	MAE	Mean Absolute Error
20	ML	Machine Learning
21	MODIS	Moderate Resolution Imaging Spectroradiometer
22	MSE	Mean Squared Error
23	NN	Neural Network
24	NDBI	Normalized Difference Built-up Index
25	NDVI	Normalized Difference Vegetation Index
26	RF	Random Forest
27	RMSE	Root Mean Squared Error
28	RNN	Recurrent Neural Network
29	SDA	Srinagar Development Authority
30	SRTM	Shuttle Radar Topography Mission
31	SUHI	Surface Urban Heat Island
32	SVM	Support Vector Machine
33	SVR	Support Vector Regression
34	UCM	Urban Canopy Models
35	UHI	Urban Heat Island
36	WRF	Weather Research and Forecasting