Discover Environment

Research

Improving spatial modeling framework to assess the impact of vegetation cover, water content, and surface temperature on urban heat island effects in Johannesburg, South Africa

Eskinder Gidey[1,2] · **Paidamwoyo Mhangara**[1]

Received: 26 January 2025 / Accepted: 2 June 2025
Published online: 15 June 2025

Abstract
The urban heat island has become a major environmental and socioeconomic concern in the City of Johannesburg Metropolitan Municipality (CoJ), South Africa. This research investigates vegetation characteristics in reducing urban heat islands (UHI) in the CoJ using multispectral and thermal remote sensing data from 1993 to 2023. The primary indices processed—including the Modified Soil Adjusted Vegetation Index (MSAVI2), Normalized Difference Moisture Index (NDMI), and Urban Heat Island (UHI)—were all analyzed using ArcGIS Pro v3.1.4183 geospatial software. A regression model determined the type and degree of relationship among the indices. Results showed that vegetation cover and water content status were below average, with MSAVI2 = 0.18 and NDMI = – 0.15. The surface temperature was found to be 17.08 °C. A significant, positive statistical relationship ($R^2 = 0.36$; $p < 0.01$; n = 1,048,575) existed between MSAVI2 and NDMI, indicating that MSAVI2 tends to increase in proportion to an increase in NDMI. A weakly negative statistical relationship existed between MSAVI2 and UHI ($R^2 = 0.08$; $p < 0.01$; n = 1,048,575), showing that UHI tends to decrease as MSAVI2 increases. Furthermore, a strongly negative, statistically significant relationship existed between NDMI and UHI ($R^2 = 0.54$; $p < 0.01$; n = 1,048,575), demonstrating that UHI decreases when NDMI rises. Creating environmentally and economically sustainable parks and green urban spaces can help mitigate the effects of urban heat islands.

Highlights

- Vegetation condition indicators were consistently low: MSAVI2 = 0.18, NDVI = 0.16, Pv = 0.33, NDMI = – 0.15
- Mean UHI = 17.08 °C, LST = 24.56 °C, SMI = 0.40 indicate heat stress and sparse vegetation
- UHI and LST declined from 1993 to 2023 as SMI improved
- Significant correlation ($p < 0.01$) among vegetation, water content, and surface temperature.

Keywords Urban heat island · Vegetation Vigor · Multispectral and thermal remote sensing · City of Johannesburg · South Africa

✉ Eskinder Gidey, eskinder.gidey@mu.edu.et | [1]School of Geography, Archaeology and Environmental Studies, Faculty of Science, University of Witwatersrand, Private Bag: 2050, Johannesburg, Gauteng, South Africa. [2]Department of Land Resources Management and Environmental Protection, College of Dryland Agriculture and Natural Resources, Mekelle University, P.O. Box: 231, Mekelle, Ethiopia.

1 Introduction

Tropical regions exhibit significant variability in urban climate patterns between wet and dry seasons [2]. Consequently, the interaction between surface and air temperatures becomes increasingly complex. Rapid urban population growth is expected to exacerbate the global issue of urban heat island (UHI) impact [11]. Researchers have yet to thoroughly investigate the efficiency of certain solutions in reducing UHI compared to areas covered by degraded vegetation [32]. Currently, one of the major environmental issues affecting sustainable development goals is the deterioration of vegetation, including forests, shrubs, grass, and crops. Degraded vegetation promotes global warming due to decreased surface albedo and increased greenhouse gases, which is a determining factor [42]. According to the United Nations' Sustainable Development Goals SDGs 2015, the objectives of SDG 15 include protecting, restoring, and promoting sustainable use of terrestrial ecosystems, sustainably managing forests, combating desertification, and halting and reversing land degradation and biodiversity loss. Conversely, the spatial extent of vegetation degradation has declined due to anthropogenic and natural-induced factors, resulting in significant social, economic, and environmental impacts. These impacts include decreased agricultural production and productivity, wildfires, biodiversity loss, decreased income or GDP growth, depletion of water resources, and losses in ecosystem services [23, 30]. For instance, drylands exposed to degradation cover approximately 46.2% of the world's surface area, affecting three billion people [23]. Additionally, Rivera-Marin et al. [30] reported that about 1.4 billion people worldwide, 74% of whom are impoverished, are affected. The increase in drylands may cause a significant reduction in soil moisture, affecting vegetation cover development. However, Africa, the second-largest continent, accounting for approximately 20.3% of the world's total land mass [45], has received insufficient attention.

Studies indicate that settlement expansion can affect vegetation cover extent, potentially transforming vegetated lands into built-up areas. This change accelerates soil erosion hazards and heat waves, disrupting vegetation and soil health [17, 46]. As a result, ecosystem services may suffer significantly. Vegetation's nutrient uptake is also impeded, impairing its health [17]. For instance, high wind erosion and decreased organic activity can reduce soil quality, hindering agriculture. Projections indicate decreased vegetation productivity in Africa, specifically in Mozambique, Namibia, Botswana, Zimbabwe, and Zambia, and increased productivity in certain regions of Madagascar and South Africa [17]. Over 90% of South African biomes show a stable trend in vegetation productivity from 1982 to 2015, whereas only 10% of the Karoo has a significantly increasing trend [39]. The same author emphasized understanding the nature, scope, and rate of change in drylands at various temporal and spatial scales.

Urban areas have expanded steadily over the past century, exacerbating surface urban heat island consequences (e.g., human health and comfort due to increased air pollution, diminished nighttime cooling, and warmer daytime temperatures) and continuously losing ecological services [36], such as clean water, nutritious food, plant drugs, and genetic resources. Environmental challenges like poor air quality and depleted water supplies, caused by rapid urbanization, have affected more people [2]. High UHI results from population growth, urban expansion, and low moisture content, leading to poor vegetation cover. Lawal et al. [17] reported that global vegetation cover is predicted to be adversely impacted by climate change, but the severity of its effects on southern Africa's vegetation biomes is poorly understood. Consequently, little is known about vegetation degradation's repercussions. One factor is inadequate research on UHI [36].

The Normalized Difference Vegetation Index (NDVI) is commonly used to describe spatial patterns in temperate and subtropical climates. Mirzabaev et al. [23] reported that vegetation degradation magnitude is challenging due to its diversity and complexity. Remote sensing vegetation indices, such as MSAVI2, EVI, NDVI, PV, and NDMI, efficiently monitor phenological growth, vegetation greenness, and moisture content. However, studies point out that NDVI is the most popular vegetation index. NDVI often reaches its limits at high leaf area indexlevels, making it susceptible to atmospheric disturbances, scaling issues, and the effects of soil background. In sparse canopies, NDVI is sensitive to canopy cover variations, but this sensitivity diminishes in intermediate and dense canopies [41]. The scientific community is increasingly focusing on advanced and reliable vegetation indices (e.g., MSAVI2) over NDVI for monitoring vegetation health. The Burn Area Index (BAI), estimated from remotely sensed data, helps depict vegetation stress due to elevated temperatures (e.g., fire). Venter et al. [40] investigated how Landsat-derived vegetation trends can be applied to assess land degradation and restoration in South Africa. Similarly, Choudhury et al. [7] generated the land surface temperature of the Asansol-Durgapur development region, finding an increase in LST by 0.06 °C and 0.43 °C per year during winter and summer seasons. The increase in LST causes significant environmental harm. Monitoring vegetation degradation and dynamics using earth observation imagery is crucial for obtaining up-to-date information and developing effective environmental policy measures [28, 40]. Vegetation degradation intensification aggravates land surface temperature (LST) incidence, resulting in elevated urban heat island (UHI) and greenhouse gas effects. LST and UHI have received attention due to their detrimental effects on ecology, climate, and public health, relying on vegetation cover, soil moisture, and albedo.

Land Surface Temperature (LST), also known as radiometric temperature, measures the emission of thermal infrared radiation [42] or the warmth of the Earth's surface. It is also referred to as the temperature of the soil and plant canopy. A 5% vegetation cover can decrease the surface temperature by up to 0.5 °C due to the shading effect (cooling). However, few studies have effectively addressed the issue of retrieving surface emissivity over diverse metropolitan surfaces [20]. Stemn and Kumi-Boateng [37] link land use change dynamics to increased surface temperature and Urban Heat Island (UHI). Factors such as vegetation and soil moisture content affect emissivity, making it a crucial consideration. UHI is severe at night, even when daytime temperatures have been higher [21], due to the area's thermal characteristics. Healthy vegetation cover can modify the climate and absorb carbon dioxide (CO_2) by altering the surface energy balance between the Earth's surface and atmosphere. The vegetation cover of Johannesburg Metropolitan City (CoJ), Gauteng Province, South Africa, is susceptible to increasing air and surface temperatures. In the CoJ, the urban heat island problem has emerged as a significant environmental and socioeconomic concern [5].

Therefore, continuously monitoring LST dynamics is necessary to maintain human health [28] and urban vegetation dynamics. Understanding vegetation's spatial distribution and temporal evolution is scientifically and economically valuable. However, gaps remain in our understanding of how LST and UHI evolve and relate to underlying surfaces at various temperatures [43]. There is a lack of comprehensive studies in South Africa, particularly in Johannesburg, that assess the combined effects of vegetation cover, water content, and surface temperature on Urban Heat Island (UHI) dynamics. Most existing research over-relies on NDVI, neglecting more sensitive indices like MSAVI2, NDMI, and PV, and fails to integrate these with LST and UHI in a unified spatial modeling framework. Additionally, African urban environments, especially dryland regions, remain underexplored in global UHI and vegetation degradation research, limiting context-specific mitigation strategies.

This study addresses the lack of research on integrating vegetation conditions and their roles in mitigating urban heat islands in South Africa. By combining reliable remote sensing spectral indices, this study focuses on analyzing vegetation cover and its role in reducing surface urban heat islands. The objectives of this study were to: (i) Evaluate vegetation cover status and water content using MSAVI2, NDVI, PV, and NDMI; (ii) Estimate spatial patterns of LST, UHI, and Soil Moisture Index (SMI) conditions; and (iii) Determine relationships among MSAVI2, NDVI, PV, NDMI, SMI, LST, and UHI using regression models. Reliable results on vegetation dynamics are essential for developing successful forest protection programs. The findings of this study can enhance existing environmental planning and management practices, particularly in combating desertification and promoting biodiversity. Additionally, public health professionals can access our scientific findings on UHI's impact on human health and how vegetation can help lower UHI.

2 Materials and methods

2.1 Study area

This research was conducted in Johannesburg, Gauteng Province, South Africa, a metropolitan city that serves as the country's financial and industrial hub. As the smallest province, it is the center of economic powerhouses and transportation [18, 47]. Mining is the primary industry in this province, one of the key sectors in South Africa [9]. Geographically, the province spans from 27°17′15″ S to 27°17′25″ S latitude and from 26°54′45″ E to 25°54′40″ E longitude (Fig. 1). The province comprises five districts: Tshwane (6296.2 km^2, 34.7%), West Rand (4084.5 km^2, 22.5%), Ekurhuleni (3458.6 km^2, 19%), Sedibeng (2686.9 km^2, 14.8%), and Johannesburg (1644.0 km^2, 9%). Additionally, the province has eight metropolitan municipalities [9]. The topography of Gauteng Province ranges from 931 to 1916 m above sea level (m.a.s.l.), with an average elevation of 1481 m.a.s.l. [9]. The study area is located in one of Gauteng Province's temperate climate zones, receiving up to 363 mm of mean annual rainfall [47]. The minimum and maximum temperatures range from 20.3 °C (69 °F) to 26.3 °C (79.3 °F), respectively, with an average temperature of 14.4 °C (57.9 °F). The study area experiences two primary climatic seasons: the dry season (April–October) and winter (May–August). According to the Soil and Terrain Database for Southern Africa (SOTERSAF version 1.0), nine dominant soil types exist in the study area. Lixisols are the most dominant, covering 6895.3 km^2 (37.9%), followed by Solonetz (68 km^2, 0.4%), Acrisols (3656.1 km^2, 20.1%), Leptosols (2924.2 km^2, 16.1%), Plinthosols (2458.3 km^2, 13.5%), Vertisols (1150.2 km^2, 6.3%), Luvisols (448.3 km^2, 2.5%), Nitisols

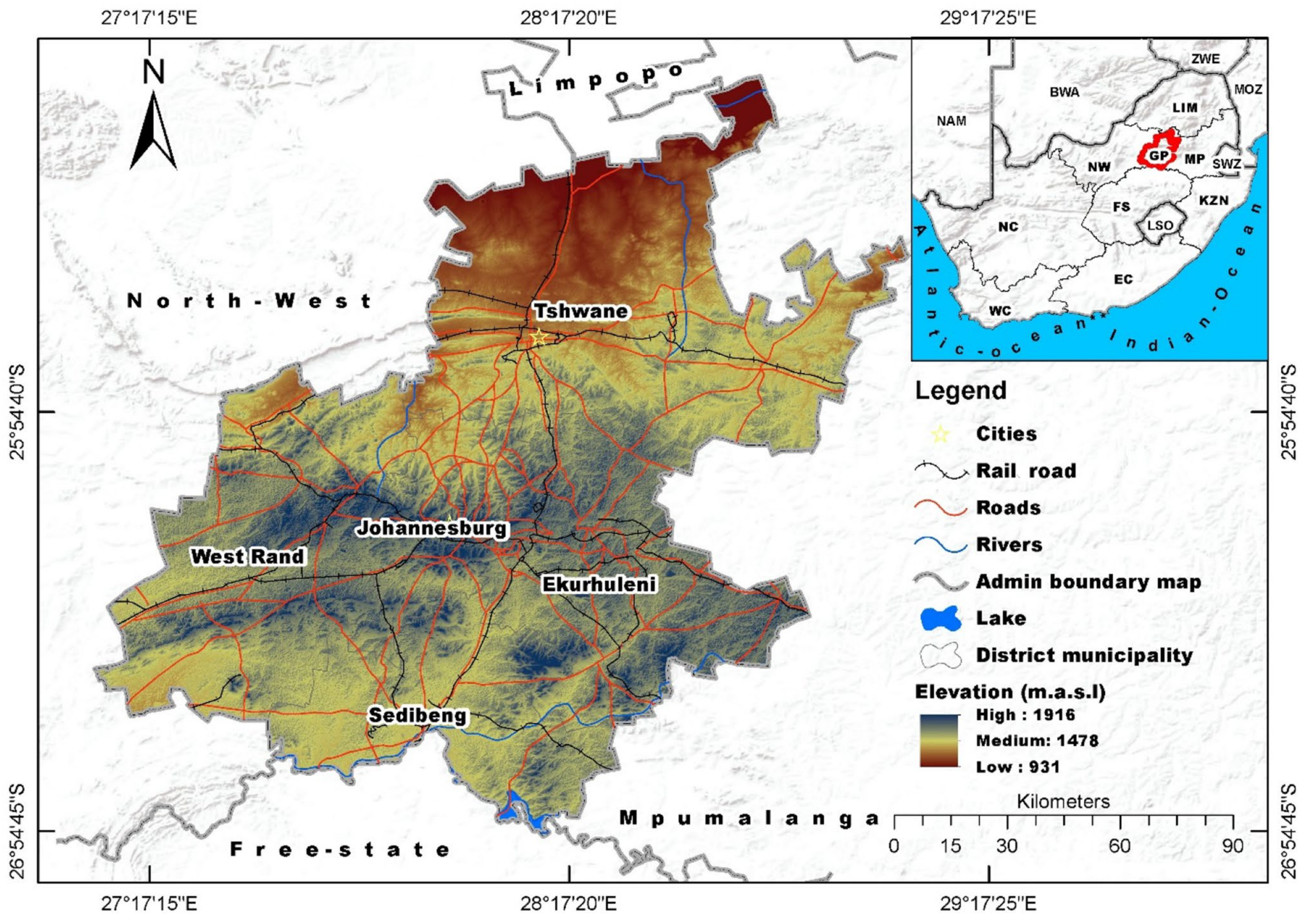


Fig. 1 Location of the study area

(441.3 km^2, 2.4%), and Regosols (100 km^2, 0.6%). Water bodies occupy 28.4 km^2 (0.2%) of the remaining land [9]. The Vaal Dam supplies water to approximately 13 million people in Gauteng Province and surrounding areas [27]. However, the province's crystalline natural rocks make it challenging to extract sufficient groundwater [34].

2.2 Image acquisition

We acquired Landsat multispectral and temporal imageries from the U.S. Geological Survey (USGS) for the summer seasons of 1993, 2000, 2013, and 2023. The imageries used were from the Thematic Mapper (TM), Operational Land Imager (OLI), and Thermal Infrared Sensor (TIRS) Collection-2 Level 1. We divided the study area into one scene (tile), specifically path 170/row 78 (Table 1). For quantification and analysis, we utilized the thermal, shortwave infrared (SWIR-1), near-infrared (NIR), and red bands from medium spatial resolution satellite imagery to calculate NSAVI2, NDVI, PV, NDMI, LSE, LST, UHI, and SMI.

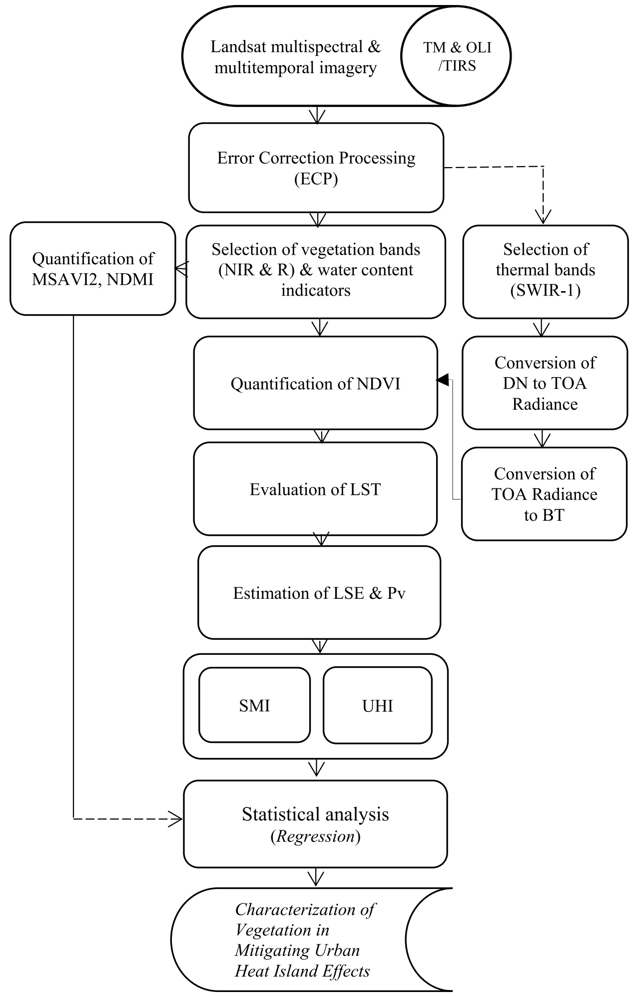
2.3 Image pre-processing, interpretation, and analysis

Earth observation datasets, such as Landsat imageries, are highly susceptible to atmospheric, radiometric, and geometric distortions (Fig. 2). For example, meteorological conditions like clouds, fog, and rain can significantly impact vegetation health and Urban Heat Island (UHI) values. Gidey and Mhangara [9] emphasized the importance of improving input quality through image pre-processing techniques to achieve better results. They also stressed the need to minimize geometry-related distortions by utilizing the World Geodetic System 1984 (WGS84) and the Universal Transverse Mercator (UTM) Zone 35 South projection. However, atmospheric and radiometric distortions affecting the surface and thermal reflectance of Landsat images in the study area reduced the effectiveness of haze and noise tools in ERDAS IMAGINE 2022

Table 1 Selected Landsat spectral bands for vegetation and surface temperature study

Year	Landsat scene identifier	Spectral band(s)	Pixel size (m)	Date and scene center time of acquisition
1993 (LS–4–5 TM C2 L1TP)	LT05_L1TP_170078_19930919_20200913_02_T1	Band 3—Red (0.63–0.69 µm) Band 4—Near Infrared (NIR) (0.76–0.90 µm) Band 5– Shortwave Infrared (SWIR– 1) (1.55–1.75 µm) Band 6 Thermal (10.40–12.50 µm)	30	1993-09-19 07:24:05.11238
2000 (LS–4–5 TM C2 L1TP)	LT05_L1TP_170078_20000906_20200906_02_T1			2000-09-06 07:40:27.2180690Z
2013 (LS–8 OLI–TIRS C2 L1TP)	LC08_L1TP_170078_20130910_20200913_02_T1	Band 4—Red (0.64–0.67 µm) Band 5—Near Infrared (NIR) (0.85–0.88 µm) Band 6—Shortwave Infrared (SWIR—1) (1.57–1.65 µm) Band 10—Thermal Infrared (TIRS—1) (10.6–11.19 µm)		2013-09-10 08:04:20.2303920Z
2023 (LS–8 OLI–TIRS C2 L1TP)	LC09_L1TP_170078_20230813_20230813_02_T1			2023-08-13 08:02:06.0463810Z

Fig. 2 Characterisation of vegetation in mitigating UHI



Version 16.7.0 Build 1416 software (Fig. 2). After applying necessary corrections, we conducted vegetation and surface temperature analysis using Eqs. 1–12 (Fig. 2).

2.3.1 Analysis of vegetation health

We conducted a pixel-by-pixel analysis of vegetation cover development and its water-holding content in the study area using four vegetation indices: Modified Soil Adjusted Vegetation Index 2 (MSAVI2), Normalized Difference Vegetation Index (NDVI), Proportion of Vegetation (PV), and Normalized Difference Moisture Index (NDMI) (Eqs. 1–5).

$$\mathrm{MSAVI2} = (2^{*}\ \mathrm{NIR}\ + 1 -\ \mathrm{sqrt}\ ((2^{*}\ \mathrm{NIR}\ + 1)^{2} - 8^{*}(\mathrm{NIR} -\ \mathrm{R})))\hat{}0.5/2 \tag{1}$$

In ArcGIS Pro raster calculator, the MSAVI2 was defined as follows (Eq. 2):

$$((2^{*}\mathrm{Float}\ (\varepsilon\mathrm{NIR}\varepsilon)) + 1 - (\mathrm{Power}\ (((\mathrm{Power}\ (((2^{*}\mathrm{Float}\ (\mathrm{NIR})) + 1), 2)) - 8^{*}(\mathrm{Float}(\mathrm{NIR}) -\ \mathrm{Float}(\varepsilon\mathrm{Red}\varepsilon))), 0.5)))/2 \tag{2}$$

Considering the influence of soil brightness in areas with sparse vegetation, we estimated the Modified Soil-Adjusted Vegetation Index 2 (MSAVI2). Wiesmair et al. [44] previously utilized this index to analyze plant cover. MSAVI2 values range from -1.0 to + 1.0. This study examined the estimated MSAVI2 values, which are presented in Table 2.

Vegetation indices, considered dimensionless indicators, range from – 1.0 to + 1.0. Positive values signify healthy vegetation conditions, while negative values indicate unhealthy vegetation. Non-vegetative areas, bare soils, and water bodies typically exhibit negative NDVI values [4]. Additionally, higher NDVI values indicate a denser presence of greenery or a greater proportion of vegetation, while lower NDVI values reflect vegetation that is under moisture stress. Healthy vegetation also

Table 2 Characterization of vegetation cover based on MSAVI2

MSAVI2 values	Description
– 1.0 to 0.2	Bare soil
0.2 to 0.4	Seed germination stage
0.4 to 0.6	Leaf development stage
> 0.6	Dense vegetation

Table 3 Vegetation cover status determination based on NDVI values [10]

NDVI values	Description
0.72 to 1.0	Very good
0.42 to 0.72	Good
0.22 to 0.42	Normal
0.12 to 0.22	Poor
–1.0 to 0.12	Very poor

Table 4 Determination of vegetation water holding or stress status[a]

NDMI value	Description
–1.0 to –0.8	Bare soil
–0.8 to –0.6	Non-existent canopy cover
–0.6 to –0.4	Very low canopy cover
–0.4 to –0.2	Low canopy cover that is either damp or dry
–0.2 to 0.0	Mid–low canopy cover with high levels of water stress-or-low canopy cover, low water stress
0.0 to 0.2	Average canopy cover under high levels of water stress
0.2 to 0.4	Mid–high canopy cover under high levels of water stress
0.4 to 0.6	High canopy cover with no water stress
0.6 to 0.8	There is no water stress and a very high canopy cover
0.8 to 1.0	No water stress or waterlogging, and complete canopy cover

[a] https://eos.com/make-an-analysis/ndmi/#:~:text=The%20NDMI%20range%20is%20%2D1,NDMI%20values%20could%20signal%20waterlogging;

enhances soil and water quality. As shown in Table 3 (Eq. 3), we utilized the Normalized Difference Vegetation Index (NDVI) to assess vegetation temperature, greenness, and health. This measurement provides accurate land surface temperature results.

$$\mathrm{NDVI} = ((\mathrm{NIR} - \mathrm{R})/(\mathrm{NIR} + \mathrm{R})) \tag{3}$$

Additionally, we used NDVI as a proxy to calculate the proportion of vegetation (PV) as shown in Eq. 4:

$$PV = Square\frac{\left(NDVI_i - NDVI_{min}\right)}{\left(NDVI_{max} - NDVI_{min}\right)} \tag{4}$$

where $NDVI_i = NDVIvalue, NDVI_{min} = theminimumNDVIvalue(lower), NDVI_{max} =$ the maximum NDVI value (higher).

We also assessed the vegetation moisture level using the Normalized Difference Moisture Index (NDMI) (Eq. 5). Water content deficiency has significantly impacted vegetation development, making this index essential for understanding the effect of Land Surface Temperature (LST) on vegetation [38, 41]. Furthermore, NDMI serves as a reliable indicator of dryness, as it evaluates various humidity levels in landscape components, including soils, rocks, and vegetation [33]. We calculated the vegetation water content in the research area using the thermal and near-infrared bands of Landsat 4–5 TM and OLI/TIRS (Eq. 5) and analyzed the results as presented in Table 4.

$$NDMI = ((\mathrm{NIR} - SWIR)/(\mathrm{NIR} + SWIR)) \tag{5}$$

where *NDMI* = Normalized Difference Moisture Index, *NIR* = Near Infrared band, SWIR = Shortwave Infrared.

The Normalized Difference Moisture Index (NDMI) is a sensitive indicator of changes in vegetation water content. With values ranging from − 1.0 to 1.0, this dimensionless index provides a clear measure of moisture levels. Specifically, lower NDMI values correspond to lower water content in vegetation, while higher values indicate higher water content.

2.3.2 Surface temperature

We analyzed the surface temperature of the study area using various indicators, including Land Surface Emissivity (LSE), Land Surface Temperature (LST), Soil Moisture Index (SMI), and Urban Heat Islands (UHI). Estimating Land Surface Temperature (LST) is a multi-step process, which involves the following five crucial steps:

i. Top of Atmospheric (TOA) spectral radiance
ii. Brightness Temperature (BT)
iii. Normalized Difference Vegetation Index (NDVI)
iv. Proportion of Vegetation (Pv)
v. Land Surface Emissivity (LSE)

2.3.3 Digital number (DN) value to top of atmosphere (TOA) radiance conversion

Converting Digital Number (DN) values to Top-of-Atmosphere (TOA) reflectance is a critical step in satellite image processing, enabling the calculation of spectral radiance for each pixel. In this study, we converted the thermal bands of Landsat 4–5 TM and OLI/TIRS to spectral radiance and brightness temperature (BT) using Eq. 6.

$$L_{\lambda} = ML \times Q_{calc} + AL \tag{6}$$

where L_{λ} = Top-of-Atmosphere (TOA) spectral radiance, *ML* = Multiplicative scaling factors, Q_{calc} = Quantized and calibrated pixel value, *AL* = Radiance additive factor for band X (band 6 for Landsat 4–5 TM, band 10 for Landsat 8 OLI/TIRS).

2.3.4 Top of atmosphere radiance to brightness temperature conversion

We converted the spectral radiance of brightness temperature (BT) from Kelvin to Celsius using the Landsat algorithm based on the Planck curve (Table 5), as expressed in Eq. 7:

$$BT = \frac{K2}{\ln\left(\frac{K1}{L_{\lambda}} + 1\right)} - 273.15 \tag{7}$$

where BT = Brightness temperature in Celsius, $K2\&K1$ = Calibration constant 2 and 1 for band 6 TM and 10 TIRS, L_{λ} = calculated TOA spectral radiance.

2.3.5 Land surface emissivity (LSE)

Land Surface Emissivity (LSE) analysis has numerous applications, including studying urban heat islands, vegetation health, agricultural productivity, and extreme heat events like wildfires. Emissivity is defined as the difference between a natural material's energy and that of a perfect blackbody at a specific temperature [15]. As global temperatures rise, public interest

Table 5 Landsat calibration constant values for bands 6 and 10

Sensor	Radiometric rescaling factor		Calibrated constant values	
	Multiplicative	Addition	K2	K1
Landsat 4– 5 TM (*Band 6*)	5.5375E–02 (i.e., 0.055375)	1.18243	1260.56	607.76
Landsat 8 OLI/TIRS (*Band 10*)	3.8000E–04 (i.e., 0.000334)	0.10000	1321.0789	774.8853

and awareness of climate change have grown [38]. Estimating land surface temperature requires information on surface emissivity. Satellite imagery, such as Landsat, can calculate this index to understand changes in emissivity caused by humidity, vegetation development, and other temporal characteristics [20]. Notably, LSE varies linearly with vegetation percentage [19]. In this study, we estimated the land surface emissivity (LSE) condition of the study area using Eq. 8.

$$LSE = 0.004 \times PV + 0.986 \tag{8}$$

We estimated Land Surface Emissivity (LSE) based on the proportion of vegetation (Pv) from NDVI, using constant soil and vegetation emissivity coefficients for both TM and OLI/TIRS satellite imageries. Specifically, we applied a standard deviation of 0.004 and a mean value of 0.986. However, for nighttime data, Konda et al. [16] recommended an emissivity value of 0.984 ± 0.004.

2.3.6 Land surface temperature (LST)

In this study, we used the thermal bands of Landsat 4–5 and 8 OLI/TIRS to derive LST as follows (Eq. 9).

$$LST = (BT/(1 + (0.00115 \times BT/1.4388) \times \ln(LSE))) \tag{9}$$

Higher Land Surface Temperature (LST) values indicate warmer surface conditions, while lower LST values signify cooler conditions, collectively describing the thermal characteristics of the land surface. The accuracy of LST values derived from satellite images relies heavily on the quality of the thermal bands.

2.3.7 Soil moisture index (SMI)

Soil moisture plays a crucial role in vegetation growth, with moisture deficiencies significantly impacting both quantity and quality [25]. To assess the soil moisture status of the study area, we employed Moawad's [24] formula (Eq. 10), a widely used index that has been successfully applied in various studies (e.g., [22]).

$$SMI = \frac{LSTmax - LST}{LSTmax - LSTmin} \tag{10}$$

The Soil Moisture Index (SMI) is calculated using the maximum (LSTmax) and minimum (LSTmin) Land Surface Temperature (LST) values, ranging from 0 to 1, where values less than 1 indicate dry (stress) conditions, values close to 1 indicate wet conditions, and typically, desert regions exhibit low SMI levels [25].

2.3.8 Urban heat islands (UHI)

We evaluated the surface urban heat island (UHI) in the study area using Eq. 11, a technique also employed by Siddiqui et al. [35]. Alternatively, Wei and Wang [43] calculated UHI by analyzing temperature differences between urban and rural areas, as well as between heat islands, cold islands, and their surrounding environments.

$$UHI = LST_i - LST_{mean}/Stdev \tag{11}$$

where UHI = Urban Heat Island, LST_i = Land Surface Temperature Value of i– period, LST_{mean} = Mean Land Surface Temperature Value of i– period, $Stdev$ = Standard Deviation value of Land Surface Temperature for i– period.

2.3.9 Statistical relationship of vegetation, water content, and surface temperature

A linear regression model (Eq. 12) was employed to examine the statistical relationships among vegetation cover, vegetation water content, and surface temperature indices, including MSAVI2, NDMI, NDVI, PV, LST, UHI, and SMI. This approach has been successfully utilized in various studies [1, 10, 12, 14]. The model's results are represented by the coefficient of determination (R2), ranging from − 1 to + 1. To interpret the strengths of the regression model, we referred to Evans [8] (Table 6).

$$Y = a + bx + e \tag{12}$$

where Y dependent; x = explanatory factor; a = intercept; b = slope (regression coefficient); e = residual error.

Table 6 Linear regression model statistical values interpretation [8]

Degree of statistical power (s)	Coefficient of determination (R^2)
Very strong (perfect)	0.64–1.00
Strong	0.36–0.63
Moderate	0.16–0.35
Weak	0.04–0.15
Very weak	0.00–0.03

3 Results and discussion

3.1 Analysis of vegetation cover, vegetation water content, and surface temperature

Table 7 and Fig. 3 present the vegetation cover characteristics, water content, and surface temperature from 1993 to 2023. The study area's mean MSAVI2 values over the past three decades were 0.16, 0.18, 0.19, and 0.18, respectively (Table 7; Fig. 3). These results indicate that the vegetation health status of the study area is characterized by average canopy-covered vegetation under high levels of water stress, exposed to high levels of bare soil and surface temperature. Healthy vegetation can mitigate the impacts of surface temperature by reflecting a large portion of incoming solar energy. Additionally, LST values are higher in sparsely vegetated, arid, and heavily developed environments [32]. This type of

Table 7 Vegetation and its water content status in the CoJ from 1993–2023

Year	Statistics	Vegetation cover status			Water content	Surface temperature condition		
		MSAVI2	NDVI	Pv	NDMI	UHI (°C)	LST (°C)	SMI
1993	Min	− 1.94	− 0.52	0.06	− 0.60	− 0.19	5.50	0.09
	Max	0.86	0.75	0.79	0.61	38.79	44.48	0.40
	x̄	0.16	0.09	0.31	− 0.27	23.8	29.48	0.21
	SD	0.14	0.09	0.05	0.10	3.20	3.20	0.03
	CV%	87.92	97.36	17.61	− 36.73	13.46	10.86	12.10
2000	Min	− 2.68	− 0.62	0.04	− 0.82	− 7.51	2.41	0.20
	Max	0.88	0.78	0.79	0.64	29.25	39.18	0.47
	x̄	0.18	0.11	0.31	− 0.25	16.67	26.59	0.29
	SD	0.14	0.09	0.05	0.09	2.92	2.92	0.02
	CV%	77.00	81.15	16.02	− 37.11	17.54	10.99	7.30
2013	Min	− 0.54	− 0.21	0.23	− 0.73	− 2.86	4.60	− 0.10
	Max	0.71	0.54	0.9	0.66	58.72	66.19	0.63
	x̄	0.19	0.11	0.46	− 0.04	20.13	27.60	0.37
	SD	0.07	0.05	0.04	0.06	2.49	2.49	0.03
	CV%	37.74	41.88	8.39	− 133.69	12.36	9.02	8.00
2023	Min	− 0.75	−0.27	0.00	− 0.39	− 8.14	− 1.32	0.62
	Max	0.68	0.52	0.87	0.39	19.31	26.13	0.92
	x̄	0.18	0.10	0.22	− 0.03	7.74	14.56	0.75
	SD	0.08	0.05	0.05	0.06	2.49	2.49	0.03
	CV%	43.73	47.91	22.93	−176.04	32.19	17.11	3.67
1993–2023	Min	− 0.86	− 0.15	0.14	− 0.37	− 1.93	5.54	0.31
	Max	0.73	0.59	0.76	0.26	25.97	33.45	0.58
	x̄	0.18	0.16	0.33	− 0.15	17.08	24.56	0.40
	SD	0.09	0.05	0.04	0.07	2.41	2.41	0.02
	CV%	0.52	0.32	0.12	− 0.45	0.14	0.10	0.06

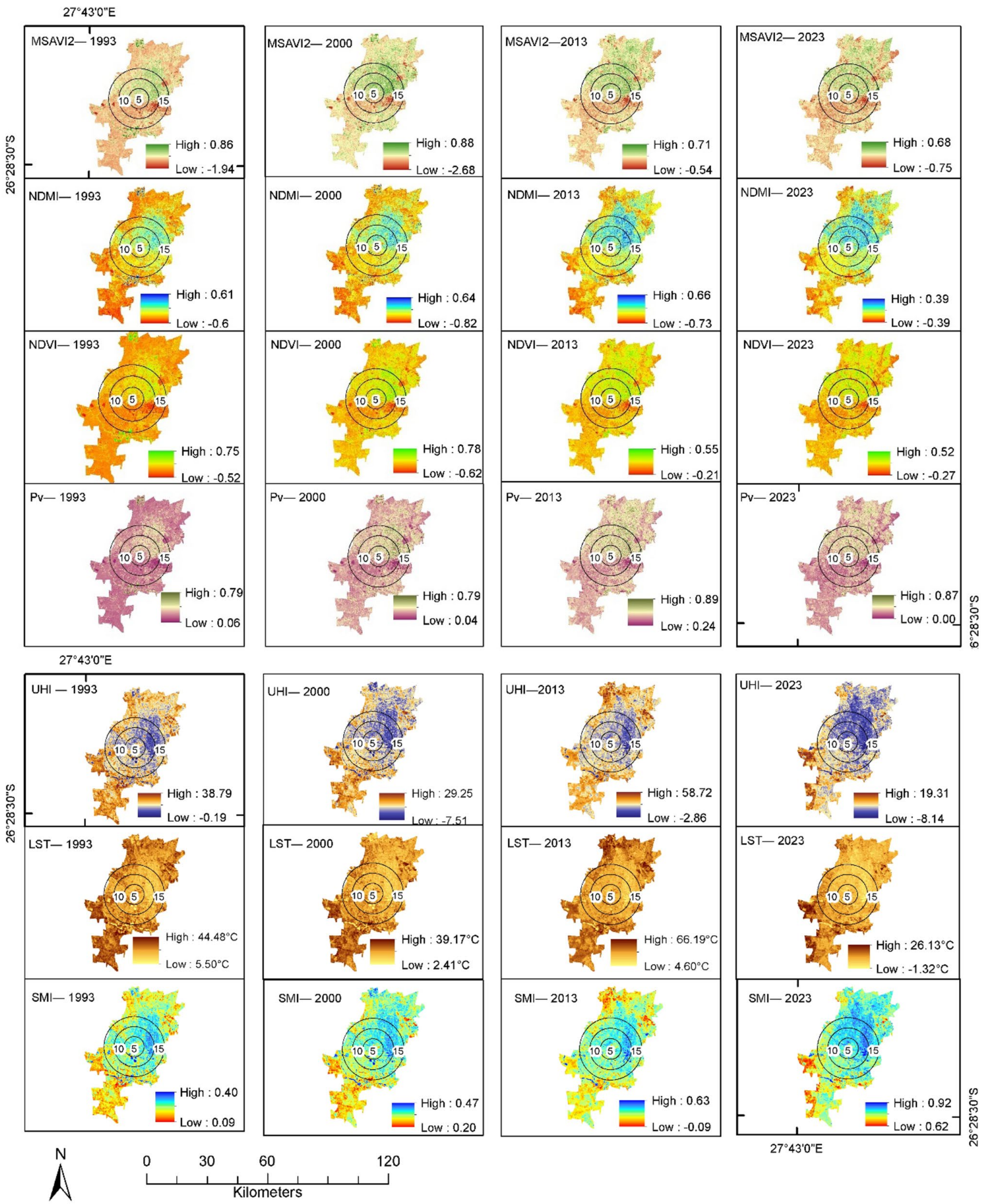


Fig. 3 A spatio-temporal trends of vegetation and surface temperature in CoJ from 1993 to 2023

vegetation cools the air about two to four times less than dense vegetation cover, reducing the natural cooling system. The NDMI confirms this finding, with mean values for CoJ in 1993, 2000, 2013, and 2023 being − 0.27, − 0.25, − 0.04, and − 0.03, respectively (Table 7). These values indicate low water content in the vegetation. The NDVI also depicted lower vegetation coverage, with mean values for CoJ in 1990, 2000, 2013, and 2023 being 0.09, 0.11, 0.11, and 0.10, respectively. Furthermore, the fraction of vegetation calculated from the NDVI over the study years shows lower PV values.

In 1993, the mean MSAVI2 values at 5, 10, and 15 km from the center of CoJ were − 0.20, − 0.44, and − 0.54, respectively. The MSAVI2 values within a 5-km radius ranged from − 1.15 to 0.75, which are lower than those within 10 and 15 km radii by − 0.18 and − 0.13, respectively. This indicates bare soil in the study area, as assessed by MSAVI2. The NDMI also showed that the mean vegetation water content within a 5 km radius was more severe than in the 10 and 15 km radius. The mean NDMI values within a 5 km radius were − 0.14, whereas those within the 10 and 15 km radii were 0.04 and − 0.01, respectively. This illustrates that the area has mid-low to shallow canopy with high levels of water stress. The mean NDVI values within a 5 km radius were 0.11, with similar values within 10 and 15 km radius. The NDVI coverage varied from − 0.38 to 0.60 within a 5 km radius, − 0.48 to 0.71 within a 10 km radius, and − 0.52 to 0.75 within a 15 km radius. The Pv values were lower within a 5 km radius (0.38) compared to the 10 and 15 km radius. The Pv values ranged from 0.10 to 0.66 within a 5 km radius, 0.07 to 0.76 within a 10 km radius, and 0.06 to 0.79 within a 15 km radius.

In 2000, we estimated the mean MSAVI2 values within 5, 10, and 15 km radii of the study area to be − 0.66, 0.60, and − 0.51, respectively. The MSAVI2 values within a 5-km radius ranged from − 2.06 to 0.74. Similarly, within the 10 km radius, the maximum and minimum MSAVI2 values were between − 2.08 and 0.88, indicating better vegetation health than the 5 and 15 km radio. The lowest and highest MSAVI2 values in the 15-km radius were − 1.84 and 0.81, respectively.

We also found the mean NDMI values within the three radii to be − 0.03, − 0.13, and − 0.03, respectively, indicating mid-low canopy cover with high levels of water stress. The NDMI values within a 5-km radius ranged from − 0.52 to 0.45, suggesting extremely low to high canopy cover with varying levels of water stress. Similarly, the minimum and maximum NDMI values within a 10-km radius ranged from − 0.82 to 0.57, respectively, indicating bare soil, high canopy cover, and lack of vegetation water stress. The NDMI values within a 15-km radius ranged from − 0.67 to 0.64, indicating either non-existent canopy or high canopy cover with vegetation and no water stress.

The mean NDVI values within the three radii were 0.03, 0.12, and 0.10, respectively, classifying the areas as having very low vegetation coverage. The NDVI values within the 5-km radius ranged from − 0.53 to 0.59, demonstrating varying vegetation coverage from very poor to good. Similarly, NDVI values within the 10-km radius ranged from − 0.55 to 0.78, and within the 15-km radius, from − 0.49 to 0.69. We measured the mean Pv values, including information about leaves, stalks, and branches, within the three radii to be 0.34, 0.42, and 0.39, respectively. The Pv values within the 5-km radius ranged from 0.05 to 0.63, within the 10 km radius from 0.05 to 0.79, and within the 15 km radius from 0.06 to 0.71.

In 2013, we estimated the mean MSAVI2 values within 5, 10, and 15 km radii of the study area to be 0.07, 0.17, and 0.08, respectively. These values indicate that the study area is highly characterized by bare soil. However, the minimum and maximum MSAVI2 values within a 5-km radius ranged from -0.54 to 0.67, revealing that the area is characterized by both bare soil and dense vegetation. This contrast significantly influences the availability of vegetation water content and counteracts the impact of surface temperature. During the same period, we measured the mean NDMI values across the three radii to be − 0.09, − 0.01, and 0.00, respectively. These observations revealed that the canopy cover ranges from extremely low to extremely high, with no water stress under extremely high canopy cover. This is critical for controlling surface temperatures and vegetation growth. The mean NDVI values across the 5, 10, and 15 km radii were 0.15, 0.19, and 0.14, respectively. The lowest and highest mean NDVI values within a 5-km radius were − 0.21 and 0.50. Similarly, the minimum and maximum NDVI values in the 10 and 15 km radii ranged from − 0.16 to 0.54 and − 0.20 to 0.48, respectively. These findings indicate that the area is characterized by very poor to good vegetation cover. We estimated the mean Pv values within the 5, 10, and 15 km radii to be 0.54, 0.58, and 0.53, respectively. The Pv values within the 5 km radius ranged from 0.23 to 0.85, within the 10 km radius from 0.27 to 0.90, and within the 15 km radius from 0.24 to 0.83.

In 2023, we estimated the mean MSAVI2 values within 5, 10, and 15 km radius to be 0.11, 0.17, and 0.08, respectively, indicating bare soil in the study area. However, further investigation within a 5 km radius revealed MSAVI2 values ranging from − 0.32 to 0.55, characterized by bare soil and under-leafed plants. Similarly, the minimum and maximum values within 10 and 15 km radii ranged from − 0.32 to 0.66 and − 0.49 to 0.65, respectively, indicating a range from bare soil to dense vegetation. We measured the mean NDMI values across the 5, 10, and 15 km radii to be − 0.04, − 0.03, and − 0.03, respectively. Overall, the study area's vegetation water content in 2023 was characterized by low canopy cover or mid-low canopy cover with significant water stress. However, NDMI values within a 5-km radius ranged from − 0.32 to 0.25, indicating low to mid-high canopy cover under high water stress. Similarly, NDMI values within 10 and 15 km radii ranged from − 0.36 to 0.31 and − 0.35 to 0.30, respectively. The mean NDVI values within 5, 10, and 15 km radii were

0.12, 0.18, and 0.14, respectively. The lowest and highest NDVI values within a 5-km radius ranged from − 0.14 to 0.38, indicating very poor to poor vegetation status. Similarly, NDVI values within 10 and 15 km radii ranged from − 0.14 to 0.50 and − 0.20 to 0.48, respectively. We measured the mean Pv values across the three radii to be 0.33, 0.44, and 0.41, respectively. The Pv values within a 5-km radius ranged from 0.04 to 0.61, within a 10 km radius from 0.04 to 0.83, and within a 15 km radius from 0.02 to 0.79.

In general, the vegetation health and water content in CoJ, measured by robust remote sensing indices from 1993 to 2023, are characterized as below normal conditions, with MSAVI2 = 0.18, NDVI = 0.16, Pv = 0.33, and NDMI = − 0.15 (Table 7).

Table 7 shows that the mean UHI and LST of the study area in 1993 were 23.80 and 29.48 °C, respectively, higher than the total study years. The SMI was also lower (0.21) than the other years, demonstrating that increases in UHI and LST significantly impact SMI. In 2000, UHI and LST declined by 7.13 and 2.89 °C compared to 1993, respectively, with a slight increase in SMI from 0.21 to 0.29. In 2013, the mean UHI, LST, and SMI increased from their 2000 status by 3.46 °C, 1.01 °C, and 0.08 °C, respectively (Table 7). Furthermore, the mean UHI and LST significantly dropped in 2023 by 12.39 and 13.04 °C, respectively, while SMI increased by 0.38. The SMI condition of the study area was highly characterized by dryness during the periods 1993, 2000, and 2013, with SMI values of 0.21, 0.29, and 0.37, respectively. However, the situation improved to 0.75 by 2023, conducive to vegetation development (Table 7). The study area's overall mean UHI, LST, and SMI from 1993 to 2023 were 17.08 °C, 24.56 °C, and 0.40 °C, respectively (Table 7). The mean spatial distribution of UHI at a 5 km radius covered an area of 78.5 km^2 (4.72%) and was higher (21.67 °C) than the 10 and 15 km radius (20.45 °C) by 1.23 and 4.64 °C, respectively, from the center of CoJ. The minimum and maximum UHI values at a 5 km radius were 11.79 and 31.55 °C, respectively. The UHI concentrations within a 10-km radius ranged from 8.56 to 32.33 °C, covering an area of 235 km^2 (14.2%) of the entire study site. The 15-km radius had the lowest UHI concentration, with a maximum value of 34.26 km^2 and a minimum value of -0.19 °C.

The mean LST value within a 5-km radius was 27.36 °C, with minimum and maximum LST values of 17.5 and 37.2 °C, respectively. This value exceeded the 10 km radius LST concentration by 1.23 °C. The LST concentrations within a 10 km radius ranged from 14.2 to 38 °C, and within a 15 km radius, from 5.5 to 39.9 °C. However, there was a slight variation of 4.64 °C from the LST concentration within a 5 km radius. The SMI values within a 5-km radius fell between 0.15 and 0.3. In the 10 and 15 km radius, these values varied from 0.14 to 0.33 and 0.14 to 0.40, respectively, indicating a dry soil moisture status. The mean SMI values for the 5, 10, and 15 km radius were 0.23, 0.24, and 0.26, respectively. This may directly or indirectly affect the water cycle [31] and exacerbate drought on various scales.

In 2000, the UHI concentration within a 5 km radius of CoJ ranged from 0.5 to 25.7 °C, which was lower than the 10 km radius by an average value of − 3.7 °C and higher than the 15 km radius by 3.0 °C. The UHI content was high within a 10-km radius, varying between 0.48 and 28.9 °C. In contrast, the UHI concentration within a 15-km radius ranged from − 7.5 to 27.7 °C. The mean LST value of CoJ was higher within a 10 km radius than within a 5 and 15 km radius by 3.7 and 6.7 °C, respectively. The lowest and highest LST values within this radius were 15.0 and 39 °C, respectively. The 5 km radius of LST values was higher than the 15 km radius by 3.03 °C, ranging from 10 to 36.0 °C. In contrast, the 15-km radius had a low LST status, with values ranging from 2.4 to 38.0 °C. The mean SMI status within 5, 10, and 15 km radii ranged from 0.30 to 0.31, respectively. The maximum and minimum SMI values within a 5-km radius ranged from 0.23 to 0.41. Similarly, the SMI values within a 10-km radius ranged from 0.20 to 0.38, and within a 15-km radius, from 0.21 to 0.47. This indicates significant soil moisture stress in the study area.

In 2013, researchers found that the UHI values within a 5 km radius were 17.46 °C, exceeding those within a 10 and 15 km radius by 5.61 and 5.21 °C, respectively. The lowest and highest UHI values within this radius ranged from 6.96 to 27.96 °C. Conversely, the minimum and maximum UHI values within a 10 and 15 km radius were -2.86 to 26.57 and − 2.15 to 26.65 °C, respectively. The mean UHI values for the 10 and 15 km radii were 11.85 and 12.25 °C, respectively. The LST values within a 5 km radius of the study area were higher than those within 10 and 15 km by 5.61 and 5.21 °C, respectively, mirroring the mean UHI values since they were derived from LST. The mean LST values for the 5, 10, and 15 km radii were 24.93, 19.32, and 19.72 °C, respectively. Following the 5 km radius, the second-highest LST accumulation in the study area was within a 15 km radius, comparable to the 10-km radius. The SMI values in the 5 km radius were lower than those in the 10 and 15 km radii. The mean SMI value for the 5 km radius was 0.4, while those for 10 and 15 km were 0.46. The minimum and maximum SMI values for the 5 km radius were 0.27 and 0.52. In contrast, SMI values within a 10 and 15 km radius ranged from 0.29 to 0.63.

In 2023, the mean UHI values ranged from 4.32 to 5.33 °C. Notably, we observed higher UHI values in a 10 km radius compared to those in 5 and 15 km areas, likely due to sparse vegetation cover. Within this radius, UHI values ranged from − 5.18 to 15.85 °C, indicating a 1.01 and 1.03 °C increase in mean UHI values. The average UHI values for 5, 10, and

15 km radii were 4.32, 5.33, and 4.30 °C, respectively. We also found that the CoJ's LST values within a 10 km radius were 1.01 and 1.03 °C higher than those within 5 and 15 km radius, with a mean LST value of 12.15 °C. The SMI values across 5, 10, and 15 km radii indicated similar moisture stress, ranging from 0.78 to 0.79. The minimum and maximum values for each radius were: 5 km (0.69–0.88), 10 km (0.66–0.89), and 15 km (0.66–0.92). Compared to previous years, we observed lower UHI, LST, and SMI values in 2023, indicating consistent wetness throughout all radii. Typically, SMI declines when UHI and LST increase, leading to decreased healthy vegetation cover. Anthropogenic factors, such as high energy use due to population growth, exacerbating LST, substantially impacting vegetation growth and development. Analyzing the relationship between spectral indices allows for a quantitative evaluation of changes in vegetation quality [29].

3.2 Statistical relationship between vegetation, vegetation water content, and surface temperature

Figure 4a–t displays the statistical relationships among vegetation, vegetation water content, and surface temperature. Our findings indicate that the correlation coefficients among these indices were statistically significant at $p < 0.01$; $n = 1,048,575$. For instance, the correlation coefficient between NDVI and MSAVI2 was statistically significant ($R^2 = 0.99$; $p < 0.01$) (Fig. 4a), suggesting that the regression model can account for 99.99% of the variation in MSAVI2. This indicates a significant concordance between NDVI and MSAVI2, essential for defining vegetation and assessing young plant health. A positive relationship exists between NDVI and MSAVI2, where increasing NDVI values tend to increase MSAVI2 values. Pravalie et al. [29] reported lower R^2 values (0.15 and 0.92) for the statistical relationship between MSAVI2 and NDVI for the years 1990 and 2011. However, our study suggests that considering mean long-term values of the two indices provides the best indication of their relationship. A substantial correlation ($R^2 = 0.36$; $p < 0.01$) exists between NDMI and MSAVI2 (Fig. 4b). The regression model explains 36% of the variation in MSAVI2, indicating that MSAVI2 tends to increase with NDMI. Viswambharan et al. [41] reported NDMI values ranging from -0.71 to 0.75 in 2001 and − 0.33 to 0.40 in 2021. We also found a high and positive correlation between NDVI and NDMI ($R^2 = 0.39$; $p < 0.01$) (Fig. 4c). Since NDVI is dependent on NDMI, our study suggests that NDVI tends to increase as NDMI increases. The statistical correlation coefficients for Pv-NDVI and Pv-MSAVI2 are both positive and significant ($R^2 = 0.99$; $p < 0.01$) (Fig. 4d and e). Similarly, a statistically significant and positive relationship ($R^2 = 0.40$; $p < 0.01$) exists between Pv and NDMI (Fig. 4f), indicating that the regression model explains 40% of the NDMI variation. As a result, NDMI tends to increase with Pv.

Figure 4g shows a weakly negative correlation ($R^2 = 0.08$; $p < 0.01$) between MSAVI2 and LST, indicating that the regression model can account for 8% of the variation in LST. This negative relationship suggests that as MSAVI2 increases, LST tends to diminish. Similarly, a weakly negative statistical correlation ($R^2 = 0.08$; $p < 0.01$) exists between NDVI and LST (Fig. 4h). This demonstrates that the regression model explains 8% of the variation in LST, influenced by the significant impact of vegetation amount in the study area. Viswambharan et al. [41] reported a weakly negative correlation between NDVI and LST in 2001 ($R^2 = 0.37$) and 2021 ($R^2 = 0.16$). Typically, LST declines as NDVI values increase. Gidey et al. [10] found weak and negative links between NDVI and LST, with R^2 values ranging from 0.40 to 0.62. Guha et al. [13] also reported a negative correlation between LST and NDVI. Adeyeri et al. [3] noted that built-up area expansion accelerates increased LST and decreased vegetation cover. Morabito et al. [26] reported a positive and highly significant linear relationship between built-up surfaces and LST ($p < 0.01$). The increase in LST affects the urban microclimate and intensifies the urban heat island effect [26], due to thermal transfer between buildings and the atmosphere [6]. This study proves that vegetation cover tends to decline as LST rises, as healthy vegetation with sufficient moisture content can reduce LST in the Metropolitan CoJ. However, seasonality can vary the level of relationship between LST and NDVI. Based on this relationship, we can determine the soil moisture content [25] of the study area.

Figure 4i shows a high and negative statistical correlation ($R^2 = 0.54$; $p < 0.01$) between NDMI and LST, indicating that the regression model can account for 54% of the variation in LST. As NDMI increases, LST tends to decrease. Similarly, a statistically significant weakly negative relationship ($R^2 = 0.09$; $p < 0.01$) exists between Pv and LST (Fig. 4j), suggesting that LST decreases as Pv increases. A weakly negative statistical relationship ($R^2 = 0.08$; $p < 0.01$) exists between MSAVI2 and UHI (Fig. 4k), with the regression model explaining 8% of the variation in UHI. As MSAVI2 increases, UHI tends to decrease. Similarly, a statistically significant weakly negative relationship ($R^2 = 0.08$; $p < 0.01$) exists between NDVI and UHI (Fig. 4l), indicating that UHI decreases as NDVI increases. Figure 4m shows a strongly negative and statistically significant relationship ($R^2 = 0.54$; $p < 0.01$) between NDMI and UHI. The regression model can account for 54% of the variation in UHI, demonstrating that UHI declines as NDMI increases. Furthermore, a weakly negative but statistically significant relationship ($R^2 = 0.09$; $p < 0.01$) exists between Pv and UHI (Fig. 4n), with the regression model explaining 9% of the variation in UHI.

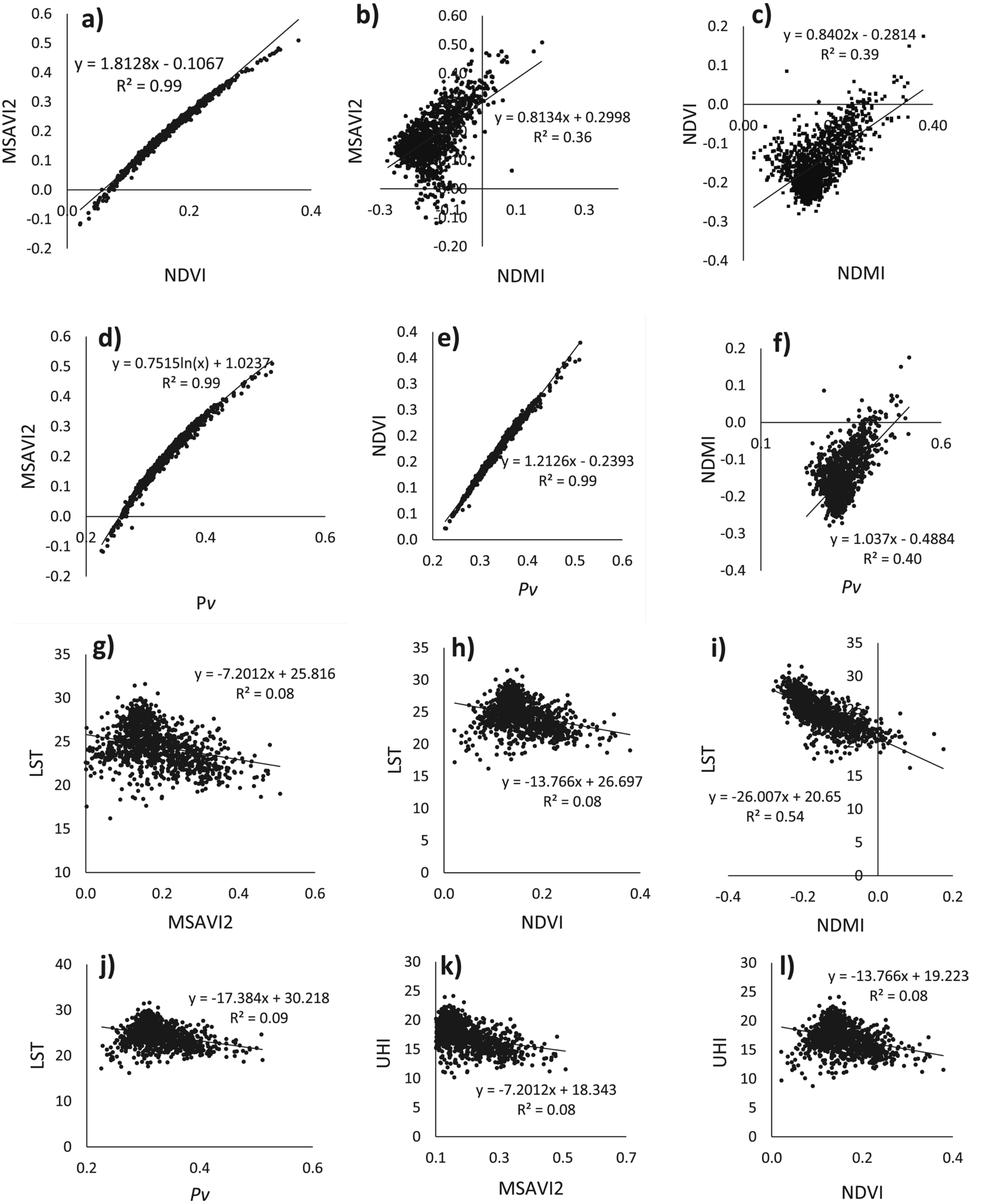


Fig. 4 **a-t** Statistical relationships among vegetation cover, vegetation water content, and surface temperature

The model also shows that UHI tends to decrease as Pv increases. Figure 4o illustrates a positive but weak statistical relationship between MSAVI2 and SMI ($R^2 = 0.07$; $p < 0.01$). The regression model can account for 7% of the variation in SMI, indicating that SMI tends to increase with MSAVI2. A weak but positive statistical relationship

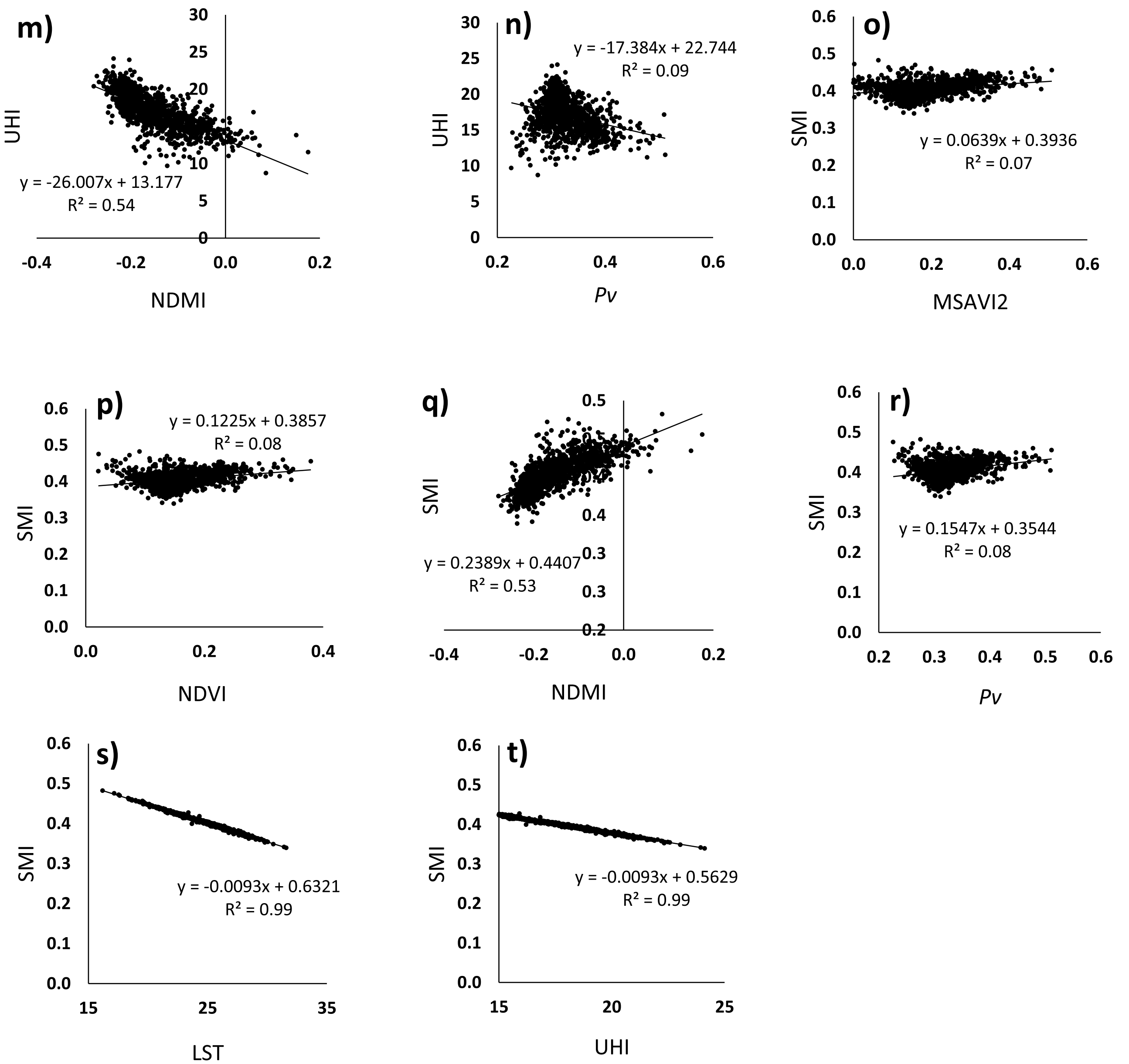


Fig. 4 (continued)

($R^2 = 0.08$) exists between NDVI and SMI (Fig. 4p), with the regression model explaining 8% of the variation in SMI. This suggests that SMI increases along with NDVI. Furthermore, a strong positive and statistically significant correlation ($R^2 = 0.53$; $p < 0.01$) exists between NDMI and SMI (Fig. 4q). The regression model can account for 53% of the variation, demonstrating that SMI increases with NDMI. A weakly positive correlation between Pv and SMI is statistically significant ($R^2 = 0.08$; $p < 0.01$) (Fig. 4r), indicating that the regression model can account for 8% of the variation in SMI. This means that SMI tends to increase with Pv. Figure 4s shows a statistically significant negative relationship ($R^2 = 0.99$; $p < 0.01$) between LST and SMI, with the regression model explaining 99.99% of the variation in SMI. The negative correlation indicates that SMI decreases as LST increases. Figure 4t shows a significant and strong statistical correlation ($R^2 = 0.99$; $p < 0.01$) between UHI and SMI. The negative correlation indicates that SMI tends to decrease as UHI increases.

4 Conclusions

This study models and analyzes vegetation cover and its role in mitigating surface urban heat islands. Robust remote sensing indices from 1993 to 2023 reveal that the overall vegetation health and water content in the City of Johannesburg (CoJ) are below average, with MSAVI2 = 0.18, NDVI = 0.16, Pv = 0.33, and NDMI = – 0.15. Similarly, the mean surface conditions are characterized by UHI = 17.08 °C, LST = 24.56 °C, and SMI = 0.40. These findings indicate that the study area is marked by sparse canopy cover, significant water stress, and widespread bare soil exposure, all contributing to the urban heat island effect. Dry, sparsely vegetated, and densely developed areas exhibit higher land surface temperatures and UHI intensity. From 1993 to 2023, vegetation indices (MSAVI2, NDVI, NDMI, and Pv) remained consistently low across 5, 10, and 15 km radii. As a result, UHI and LST peaked in 1993 but declined significantly by 2023, aligning with improvements in SMI values. Spatial analysis shows that areas with limited vegetation and dense urbanization have higher LST and UHI, while greener zones provide natural cooling. Healthy vegetation mitigates surface temperatures and enhances vegetation water content. Notably, a statistically significant relationship ($p < 0.01$; n = 1,048,575) exists among vegetation indices, vegetation water content, and surface temperature. The analysis suggests that developing parks and green urban areas can offer an environmentally and economically viable strategy for mitigating urban heat island effects. These findings underscore the need to increase green spaces and vegetation to combat urban heat islands, thereby improving the city's resilience to heat stress and enhancing residents' quality of life.

Acknowledgements We are indebted to the support of the University of the Witwatersrand School of Geography, Archaeology, and Environmental Studies (GAES), Johannesburg– South Africa.

Author contributions E.G. and P.M. conducted the research and submitted the manuscript for publication.

Funding: This research was funded by the University of Witwatersrand, School of Geography, Archaeology, and Environmental Studies (GAES), Johannesburg– South Africa.

Data availability Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate No ethical approval is required. The authors hereby declare that they have provided informed consent to participate in this research and publication process.

Consent for publication The authors have been fully informed and have agreed to publish the manuscript.

Competing interests The authors declare no competing interests.

References

1. Abulibdeh A. Analysis of urban heat island characteristics and mitigation strategies for eight arid and semi–arid gulf region cities. Environ Earth Sci. 2021;80:1–26.
2. Acero JA, González-Asensio B. Influence of vegetation on the morning land surface temperature in a tropical humid urban area. Urban Clim. 2018;26:231–43.
3. Adeyeri OE, Akinsanola AA, Ishola KA. Investigating surface urban heat island characteristics over Abuja, Nigeria: relationship between land surface temperature and multiple vegetation indices. Remote Sens Appl Soc Environ. 2017;7:57–68.
4. Amiri M, Pourghasemi HR. Mapping the NDVI and monitoring of its changes using Google Earth Engine and Sentinel-2 images. In: Computers in earth and environmental sciences. Amsterdam: Elsevier; 2022. p. 127–36.

5. Brou YT, Shikwambana L, Sivakumar V. Dynamics of Climatic and Vegetation Parameters in Urban and Township Areas: A Case Study Over the City of Johannesburg and Alexandra Township in South Africa. *In Proceedings of the Earth and Environmental Sciences International Webinar Conference.* Cham: Springer International Publishing. 2021; pp. 431–448
6. Chen W, Hu C, Liu L, Zhang J, Wang K, Liu S. Characteristics of built-up land change in resource-based cities and their impact on land surface temperature—Taking Wu'an as an example. Eco Inform. 2022;68:101582.
7. Choudhury D, Das K, Das A. Assessment of land use land cover changes and its impact on variations of land surface temperature in Asansol-Durgapur Development Region. Egypt J Remote Sens Space Sci. 2019;22(2):203–18.
8. Evans JD. Straightforward statistics for the behavioral sciences. Thomson Brooks/Cole Publishing Co. 1996.
9. Gidey E, Mhangara P. An application of machine–learning model for analyzing the impact of land– use change on surface water resources in Gauteng Province South Africa. Remote Sens. 2023;15(16):4092. https://doi.org/10.3390/rs15164092.
10. Gidey E, Dikinya O, Sebego R, Segosebe E, Zenebe A. Using drought indices to model the statistical relationships between meteorological and agricultural drought in Raya and its environs, Northern Ethiopia. Earth Syst Environ. 2018;2:265–79.
11. Gomez-Martinez F, De Beurs KM, Koch J, Widener J. Multi-temporal land surface temperature and vegetation greenness in urban green spaces of Puebla. Mexico Land. 2021;10(2):155.
12. Guha S. Dynamic seasonal analysis on LST– NDVI relationship and ecological health of Raipur City India. Ecosyst Health Sustain. 2021;7(1):1927852.
13. Guha S, Govil H, Dey A, Gill N. Analytical study of land surface temperature with NDVI and NDBI using Landsat 8 OLI and TIRS data in Florence and Naples city Italy. Eur J Remote Sens. 2018;51(1):667–78.
14. Hishe S, Gidey E, Zenebe A, Girma A, Dikinya O, Sebego R, Lyimo J. Urban heat island and ecological condition modeling using thermal remote sensing in Tigray-Northern Ethiopia. Model Earth Syst Environ. 2023;10:1–15.
15. Jin M, Liang S. An improved land surface emissivity parameter for land surface models using global remote sensing observations. J Clim. 2006;19(12):2867–81.
16. Konda M, Imasato N, Nishi K, Toda T. Measurement of the sea surface emissivity. J Oceanogr. 1994;50:17–30.
17. Lawal S, Lennard C, Hewitson B. Response of southern African vegetation to climate change at 1.5 and 2.0 global warming above the pre– industrial level. Clim Serv. 2019;16:100134.
18. Landau LB, Gindrey V. Migration and population trends in Gauteng province 1996–2055. Migration Studies Working Paper Series, 2008; 42.
19. Li ZL, Tang BH, Wu H, Ren H, Yan G, Wan Z, Sobrino JA. Satellite-derived land surface temperature: current status and perspectives. Remote Sens Environ. 2013;131:14–37.
20. Mallick J, Singh CK, Shashtri S, Rahman A, Mukherjee S. Land surface emissivity retrieval based on moisture index from LANDSAT TM satellite data over heterogeneous surfaces of Delhi city. Int J Appl Earth Obs Geoinf. 2012;19:348–58.
21. Mathew A, Khandelwal S, Kaul N. Analysis of diurnal surface temperature variations for the assessment of surface urban heat island effect over Indian cities. Energy Build. 2018;159:271–95.
22. Medida SK, Rani PP, Kumar GS, Sireesha PG, Kranthi KC, Vinusha V, Hossain A. Detection of water deficit conditions in different soils by comparative analysis of standard precipitation index and normalized difference vegetation index. Heliyon. 2023;9(4): e15093.
23. Mirzabaev A, Wu J, Evans J, García-Oliva F, Hussein IAG, Iqbal MH, Kimutai J, Knowles T, Meza F, Nedjraoui D, Tena F, Türkeş M, Vázquez RJ, Weltz M. Desertification. In: Shukla PR, Skea J, Calvo Buendia E, Masson-Delmotte V, Pörtner H-O, Roberts DC, Zhai P, Slade R, Connors S, van Diemen R, Ferrat M, Haughey E, Luz S, Neogi S, Pathak M, Petzold J, Portugal Pereira J, Vyas P, Huntley E, Kissick K, Belkacemi M, Malley J, editors. Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems; 2019. In press.
24. Moawad BM. Geoscience general tool package. Max– Planck Institute für Chemie, Mainz, Germany. 2012.
25. Mohamed ES, Ali A, El-Shirbeny M, Abutaleb K, Shaddad SM. Mapping soil moisture and their correlation with crop pattern using remotely sensed data in arid region. Egypt J Remote Sens Space Sci. 2020;23(3):347–53.
26. Morabito M, Crisci A, Messeri A, Orlandini S, Raschi A, Maracchi G, Munafò M. The impact of built-up surfaces on land surface temperatures in Italian urban areas. Sci Total Environ. 2016;551:317–26.
27. Obaid A, Adam E, Ali KA. Land use and land cover change in the vaal dam catchment, South Africa: a study based on remote sensing and time series analysis. Geomatics. 2023;3(1):205–20.
28. Orimoloye IR, Ololade OO, Mazinyo SP, Kalumba AM, Ekundayo OY, Busayo ET, Nel W. Spatial assessment of drought severity in Cape Town area South Africa. Heliyon. 2019;5(7): e02148.
29. Pravalie R, Sîrodoev I, Peptenatu D. Changes in the forest ecosystems in areas impacted by aridization in south-western Romania. J Environ Health Sci Eng. 2014;12:1–15.
30. Rivera-Marin D, Dash J, Ogutu B. The use of remote sensing for desertification studies: a review. J Arid Environ. 2022;206:104829.
31. Saha A, Patil M, Goyal VC, Rathore DS. Assessment and impact of soil moisture index in agricultural drought estimation using remote sensing and GIS techniques. Multidiscip Digit Pub Inst Proc. 2018;7(1):2.
32. Schwaab J, Meier R, Mussetti G, Seneviratne S, Bürgi C, Davin EL. The role of urban trees in reducing land surface temperatures in European cities. Nat Commun. 2021;12(1):6763.
33. Seeyan S, Merkel B, Abo R. Investigation of the relationship between groundwater level fluctuation and vegetation cover by using NDVI for Shaqlawa Basin, Kurdistan Region-Iraq. J Geogr Geol. 2014;6(3):187.
34. Shikwambana L, Kganyago M, Mhangara P. Temporal analysis of changes in anthropogenic emissions and urban heat islands during COVID– 19 restrictions in Gauteng province, South Africa. Aerosol Air Qual Res. 2021;21(9):200437. https://doi.org/10.4209/aaqr.200437.
35. Sidiqui P, Huete A, Devadas R. Spatio– temporal mapping and monitoring of Urban Heat Island patterns over Sydney, Australia using MODIS and Landsat– 8. In 2016 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA). IEEE. 2016; (pp. 217–221).
36. Simwanda M, Ranagalage M, Estoque RC, Murayama Y. Spatial analysis of surface urban heat islands in four rapidly growing African cities. Remote Sens. 2019;11(14):1645.

37. Stemn E, Kumi-Boateng B. Modelling of land surface temperature changes as determinant of urban heat island and risk of heat– related conditions in the Wassa West Mining Area of Ghana. Model Earth Syst Environ. 2020;6(3):1727–40.
38. Taloor AK, Manhas DS, Kothyari GC. Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data. Appl Comput Geosci. 2021;9:100051.
39. Timm Hoffman M, Skowno A, Bell W, Mashele S. Long-term changes in land use, land cover and vegetation in the Karoo drylands of South Africa: implications for degradation monitoring. Afr J Range Forage Sci. 2018;35(3–4):209–21.
40. Venter ZS, Scott SL, Desmet PG, Hoffman MT. Application of Landsat– derived vegetation trends over South Africa: potential for monitoring land degradation and restoration. Ecol Ind. 2020;113:106206.
41. Viswambharan S, Kumaramkandath IT, Tali JA. A geospatial approach in monitoring the variations on surface soil moisture and vegetation water content: a case study of Palakkad District, Kerala India. Environ Earth Sci. 2022;81(20):494.
42. Wan Mohd Jaafar WS, Abdul Maulud KN, Muhmad Kamarulzaman AM, Raihan A, Md Sah S, Ahmad A, Razzaq Khan W. The influence of deforestation on land surface temperature—A case study of Perak and Kedah Malaysia. Forests. 2020;11(6):670.
43. Wei X, Wang XJ. Analyzing the spatial distribution of LST and its relationship with underlying surfaces in different months by classification and intersection. Front Environ Sci. 2022;10:872282.
44. Wiesmair M, Feilhauer H, Magiera A, Otte A, Waldhardt R. Estimating vegetation cover from high–resolution satellite data to assess grassland degradation in the Georgian Caucasus. Mt Res Dev. 2016;36(1):56–65.
45. Yao R, Cao J, Wang L, Zhang W, Wu X. Urbanization effects on vegetation cover in major African cities during 2001–2017. Int J Appl Earth Obs Geoinf. 2019;75:44–53.
46. Zoungrana BJ, Conrad C, Thiel M, Amekudzi LK, Da ED. MODIS NDVI trends and fractional land cover change for improved assessments of vegetation degradation in Burkina Faso, West Africa. J Arid Environ. 2018;153:66–75.
47. Zubaidi SL, Ortega–Martorell S, Al–Bugharbee H, Olier I, Hashim KS, Gharghan SK, Al Khaddar R. Urban water demand prediction for a city that suffers from climate change and population growth: Gauteng province case study. Water. 2020;12(7):1885.