MethodsX 13 (2024) 102915

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/methodsx

Modeling the spatially varying effects of biophysical factors on land surface temperature

Getahun Bekele Debele[a,b,*], Kassahun Ture Beketie[a]

[a] Center for Environmental Sciences, College of Natural and Computational Sciences, Addis Ababa University, PO Box 1176, Addis Ababa, Ethiopia
[b] Department of Geography and Environment, Debark University, PO Box 90, Debark, Ethiopia

ARTICLE INFO

Method name:
Geospatial modeling

Keywords:
Spatial non-stationarity
Land surface temperature
Biophysical factors
Ordinary least squares
Geographically weighted regression

ABSTRACT

A growing number of studies have investigated how land surface temperature (LST) is influenced by a variety of driving factors; however, little effort has been made to identify the dominant ones. The suggested method used the Upper Awash Basin (UAB), Ethiopia, as an example to explore the spatial heterogeneity and factors affecting LST, which is critical for selecting effective mitigation strategies to manage the thermal environment. The study employed two models: ordinary least squares (OLS) and geographically weighted regression (GWR). The OLS model was first used to capture the overall relationship between LST and some biophysical factors. The GWR was then utilized to investigate the spatial non-stationary relationships between LST and its influencing biophysical factors. Although the method was tested in UAB, Ethiopia, it can be applied in similar agroecosystems, to identify the dominant factors that influence LST and develop site-specific LST mitigation strategies.

- The OLS and GWR models investigated the spatial heterogeneities of the influencing factors and LST.
- Biophysical parameters such as enhanced vegetation index (EVI), modified normalized difference water index (MNDWI), normalized difference built-up index (NDBI), normalized difference bareness index (NDBaI), albedo and elevation were used as potential driving environmental factors of LST
- The models performance was computed using the adjusted coefficient of determination (adj. R^2), Akaike Information Criterion (AICc), and residual sum of squares (RSS).

Specifications table

Subject area:	Environmental Science
More specific subject area:	Land surface temperature
Name of your method:	Geospatial modeling
Name and reference of original method:	Debele, G.B. and Beketie, K.T., 2023. Studying the spatial non-stationary relationships of some physical parameters on the Earth's surface temperature using GWR in Upper Awash basin, Ethiopia. *Scientific African*, 23(2024), p.e02052, https://doi.org/10.1016/j.sciaf.2023.e02052
Resource availability:	Landsat 8 OLI/TIRS and Shuttle Radar Topography Mission (SRTM) datasets were freely accessed from the United States Geological Survey (USGS) portal https://earthexplorer.usgs.gov/

* Corresponding author.
E-mail address: getahun.bekele@aau.edu.et (G.B. Debele).

https://doi.org/10.1016/j.mex.2024.102915
Received 3 March 2024; Accepted 14 August 2024
Available online 16 August 2024

Method details

LST is a crucial parameter that reflects land–atmosphere interaction, and its intensity and spatial pattern are mostly influenced by the biophysical composition of the land surface. Understanding the intensity and the spatio-temporal variations of LST is vital to derive appropriate mitigation strategies to regulate the thermal environment [1]. LST derived from satellite remotely sensed thermal infrared (TIR) imagery has become an indispensable data source to study the spatial non-stationary relationships between LST and its influencing biophysical factors. This method offers several advantages, including cost-effectiveness, wide temporal and spatial coverage, and data characterized by reputable accuracy, surpassing the limitations of in-situ data sources, such as traditional weather stations [2]. The spatial variation of LST is complex and nonlinear. Thus, specific methods are required for evaluating this phenomenon. This phenomenon has been evaluated using a variety of statistical approaches, including global regression models such as ordinary least squares (OLS), spatial error model (SEM), and spatial lag model (SLM), which often overlook the issue of spatial heterogeneity. Recently, various studies have employed a strong statistical approach called Geographically Weighted Regression (GWR), which adds geographical distribution information to the regression parameters to examine spatial changes. Therefore, this article describes the methodology used to investigate the spatial non-stationary relationships between LST and its influencing environmental factors. Landsat 8 OLI/TIRS images (168/54 and 169/54) collected in January 2021 were utilized to extract LST and its influencing environmental factors, while elevation was computed using SRTM DEM data [3]. Before analysis, satellite images must be radiometrically corrected to remove noise from the sensor and atmosphere, which is done by converting digital numbers (DN) to radiance and reflectance using a spectral radiance model [4]. First, the Landsat-8 OLI imagery was converted to top-of-atmosphere (TOA) spectral radiance using the radiance rescaling factors provided in the metadata file using the equation given in Eq. (1).

$$L_\lambda = M_L \times Q_{cal} + A_L \tag{1}$$

where L_λ is TOA spectral radiance in W/(m^2 × sr × μm), M_P is the band-specific multiplicative scaling factor from the metadata, A_L is the band-specific additive scaling factor from the metadata, and Q_{cal} is Level 1 pixel value in DN.

Landsat-8 OLI band data can also be converted to TOA planetary reflectance using the reflectance rescaling coefficients provided in the product metadata file, as shown in Eq. (2).

$$\rho_{\lambda'} = M_\rho \times Q_{cal} + A_\rho \tag{2}$$

where $\rho_{\lambda'}$ is TOA planetary reflectance in W/(m^2 × sr × μm), M_ρ is the band-specific multiplicative scaling factor from the metadata, A_ρ is the band-specific additive scaling factor from the metadata, and Q_{cal} is Level 1 pixel value in DN. The derived TOA reflectance is not true reflectance; solar elevation angle correction is required to obtain true TOA reflectance is as follows Eq. (3).

$$\rho_\lambda = \frac{\rho_{\lambda'}}{cos\left(\theta_{SZ}\right)} = \frac{\rho_{\lambda'}}{sin\left(\theta_{SE}\right)} \tag{3}$$

where ρ_λ is the true planetary reflectance and θ_{SE} is local sun elevation angle, θ_{SZ} is local solar zenith angle.

Extraction of biophysical factors

The biophysical properties of the earth's surface often considered as influencing factors in LST studies, which contain a large amount of spectral information and are easy to obtain from remoting sensing images. EVI, a modified version of the NDVI, was first introduced by Tucker [5] and has become one of the most widely used indexes for characterizing vegetation greenness. EVI was shown to have a negative relationship with LST [6], indicating that vegetation has a greater cooling capability and plays an important role in lowering the LST of its surrounding neighbors. Xu [7] introduced a modified normalized difference water index (MNDWI) to improve the features of open water bodies, which is now frequently used to evaluate the impact of water bodies on LST. The NDBI, NDBaI, and albedo are important biophysical factors that exacerbate the thermal environment. NDBI, initially developed by Zha et al. [8], is the most extensively used indices for characterizing and identifying the built-up area. NDBaI is also a well-known LULC indices for estimating the extent of barren lands in the region of study, and it is commonly used to evaluate the impact of bare soil on LST [9]. Albedo, a measure of the fraction of solar radiation reflected from the earth's surface, is one of the environmental factors influencing LST and is calculated as described by Liang [10]. The radiometrically corrected reflectance bands of Landsat-8 were used to calculate the selected five surface biophysical features (Table 1).

Retrieval of LST

The retrieval of LST using the mono-window algorithm involves the following steps. In the first step, the digital number (DN) of Landsat 8 OLI/TIRS band 10 images was converted to spectral radiance using Eq. (4) given by Ihlen & Zanter, [12].

$$L_\lambda = ML \times Q_{cal} + AL - O_i \tag{4}$$

where, L_λ is the Top-of-Atmosphere (TOA) spectral radiance in W/(m^2 × sr × μm), ML is the band-specific multiplicative scaling factor from the metadata, AL is the band-specific additive scaling factor from the metadata, Q_{cal} is Level 1 pixel value in DN and O_i is the correction of the thermal band.

Secondly, using Eq. (5), sensor spectral radiance is converted into sensor brightness

$$T_B = \frac{K_2}{ln\left(\frac{K_1}{P\lambda} + 1\right)} \tag{5}$$

Table 1
Equation for Surface biophysical parameter calculations.

Biophysical parameters	Equation	Description
EVI	[*] $EVI = G * \frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+(C1*\rho_{RED}-C2*\rho_{BLUE})+L}$	ρ_{NIR}, ρ_{RED}, ρ_{BLUE} represent reflectance values of Bands 2, 4, and 5 of the Landsat-8 OLI/TIRS
MNDWI	$MNDWI = \frac{\rho_{Green}-\rho_{SWIR1}}{\rho_{Green}+\rho_{SWIR1}}$	ρ_{Green}, ρ_{SWIR1} represent reflectance values of Bands 3 and 6 of the Landsat-8 OLI/TIRS
NDBI	$NDBI = \frac{\rho_{SWIR1}-\rho_{NIR}}{\rho_{SWIR1}+\rho_{NIR}}$	ρ_{SWIR1}, ρ_{NIR} represent reflectance values of Bands 5 and 6 of the Landsat-8 OLI/TIRS
NDBaI	$NDBaI = \frac{\rho_{SWIR1}-\rho_{TIRS}}{\rho_{SWIR1}+\rho_{TIRS}}$	ρ_{SWIR1}, ρ_{TIRS} represent reflectance values of Bands 5 and 10 of the Landsat-8 OLI/TIRS
Albedo	$Albedo = \frac{0.356\ \alpha 2+0.130\ \alpha 4+0.373\ \alpha 5+0.085\ \alpha\ 6+0.072\ \alpha 7-0.0018}{0.356+0.130+0.373+0.085+0.072}$	α represent reflectance values of Bands 2,4,5,6 and 7

[*] L is a soil adjustment factor, and C1 and C2 are coefficients used to correct aerosol scattering in the red band by the use of the blue band. In general, $G = 2.5$, C1=6.0, C2=7.5, and $L = 1$ [11].

where, T_B is effective at-sensor brightness temperature in Kelvin (K), and K_1 and K_2 are band-specific thermal conversion constants from the metadata. For Landsat-8 TIRS band 10, K_1 was 774.8853 and K_2 was 1321.0789, $P\lambda$ is sensor spectral radiance and ln is the natural logarithm.

The land surface is composed of different objects that have different emissivities. Since the land surface emissivity (ε) varies greatly with surface features, it is vital to determine it before computing LST [13]. The NDVI threshold approach is one of the most popular emissivity extraction techniques currently in use due to its simplicity of application [14]. Thus, ε is calculated as given in Eq. (6) given by Sobrino et al. [15].

$$\mathcal{E} = 0.004P_V + 0.986 \tag{6}$$

where, $\mathcal{E}$ is emissivity and P_V is the proportion of the vegetation, which may be estimated using the formula provided by Carlson and Ripley [16] and stated in Eq. (7).

$$P_V = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right)^2 \tag{7}$$

where, $NDVI_{max}$ is a fully vegetated land-covers $NDVI$ value 0.5; $NDVI_{min}$ is a non-vegetated land-covers $NDVI$ value 0.2 [14].

Finally, Eq. (8) was applied to correct the computed radiant surface temperature for emissivity .

$$LST = \frac{T_B}{\left\{1 + \left[\left(\frac{\lambda T_B}{\rho}\right) ln\varepsilon\lambda\right]\right\}} - 273.15 \tag{8}$$

where, LST is land surface temperature (in degree Celsius); T_B is the radiant surface temperature (in kelvin); λ is emitted radiance wavelength (10.8 μm); ρ is calculated as $\rho = h\frac{c}{\sigma}$ =1.438 × 10^{-2} mK, ρ is the Boltzmann constant (1.38 × 10^{-23} J/K), h is Planck's constant (6.62 × 10^{-34} J/s), and c is velocity of light (2.998 × 10^{8} m s^{-1}); and ε is land surface emissivity.

Spatial association of LST

To evaluate the relationships between LST and selected biophysical factors, 5446 sample points in the study area were randomly generated using 'create random points' tool in ArcGIS 10.8. The values of each parameter were then extracted for each point using the 'extract multi values to points' tools, and exported to the OLS and GWR models. Initially, using GeoDa 1.20.0 software, the OLS model was used to examine the relationships between LST and its influencing factors [17]. The expression for this global OLS model is Eq. (9).

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \tag{9}$$

where, y_i is the i^{th} observation of dependent variable, x_{ij} is the i^{th} observation of the j^{th} independent variable, β_j is the regression coefficient of the j^{th} variable, and ε_i is the error term.

However, the OLS regression could be biased whether spatial autocorrelation and spatial heterogeneity are present. Therefore, the Global Moran's I test was applied to assess whether there is spatial autocorrelation in the LST and the residual term. The global Moran's I Index is calculated using the following formula [18].

$$Moran's\ I = \frac{n \sum \sum w_{ij}\left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right)}{w \sum \left(x_i - \bar{x}\right)^2} \tag{10}$$

$$Local\ Moran's\ I = \frac{n\left(x_i - \bar{x}\right)\ \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})} \tag{11}$$

where, n is number of points; x_i and x_j are the LST values of sample i and j, respectively; w_{ij} is the spatial weight describing the distance between i and j; and $\bar{x}$ is the average LST value.

The local spatial autocorrelation index is often measured by the local Moran' I statistic in Eq. (11), used to explore the geographic location of such clustered/dispersed patterns of LST [19].

The local patterns of the relationship between LST and explanatory variables are better characterized by the GWR model than by the global OLS model [20]. The GWR model was developed using the GWR4 software [21]. The GWR model can be expressed mathematically as follows:

$$y_i = \beta_0(u_i\ ,\ v_i) + \sum_{k=1}^{n} \beta_k(u_i\ ,\ v_i) x_{ik} + \varepsilon_i \tag{12}$$

where, $(u_i\ ,\ v_i)$ is the geographic coordinate of the i^{th} location; $\beta_0(u_i\ ,\ v_i)$ is a constant term, $\beta_k(u_i\ ,\ v_i)$ is the regression coefficient of each variable at point i and ε_i is the random error term at point i.

In this study, to evaluate the performance of the OLS and GWR model, three parameters including adj. R^2, and the RSS were chosen, and higher adjust R2 and lower AICc and RSS values indicate better model performance [22].

Ethics statements

This declaration is not applicable.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Getahun Bekele Debele: Conceptualization, Methodology, Data curation, Software, Validation, Writing – original draft. **Kassahun Ture Beketie:** Conceptualization, Methodology, Data curation, Supervision, Writing – review & editing.

Data availability

Raw data is deposited in the Mendeley dataset repository available at: https://data.mendeley.com/datasets/ydspw5pxtd/1.

Acknowledgments

We thank the United States Geological Survey (USGS) for providing the Landsat 8 OLI/TIRS and SRTM DEM data publicly available and accessible.

References

[1] Z. Lin, H. Xu, X. Yao, C. Yang, L. Yang, Exploring the relationship between thermal environmental factors and land surface temperature of a "furnace city" based on local climate zones, Build. Environ. 243 (2023) 110732, doi:10.1016/j.buildenv.2023.110732.
[2] H. Liu, Q. Zhan, C. Yang, J. Wang, The multi-timescale temporal patterns and dynamics of land surface temperature using Ensemble Empirical Mode Decomposition, Sci. Total Environ. 652 (2019) 243–255, doi:10.1016/j.scitotenv.2018.10.252.
[3] G.B. Debele, K.T. Beketie, Studying the spatial non-stationary relationships of some physical parameters on the Earth's surface temperature using GWR in Upper Awash basin, Ethiopia. Scientific African 23 (2024) (2023) e02052, doi:10.1016/j.sciaf.2023.e02052.
[4] D.P. Roy, V. Kovalskyy, H.K. Zhang, E.F. Vermote, L. Yan, S.S. Kumar, A. Egorov, Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity, Remote Sens. Environ. 185 (2016) 57–70, doi:10.1016/j.rse.2015.12.024.
[5] C.J. Tucker, Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ. 8 (2) (1979) 127–150, doi:10.1016/0034-4257(79)90013-0.
[6] Z. Xu, Y. Li, Y. Qin, E. Bach, A global assessment of the effects of solar farms on albedo, vegetation, and land surface temperature using remote sensing, Sol. Energy 268 (2024) 112198, doi:10.1016/j.solener.2023.112198.
[7] H. Xu, Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, Int. J. Remote Sens. 27 (14) (2006) 3025–3033, doi:10.1080/01431160600589179.
[8] Y. Zha, J. Gao, S. Ni, Use of normalized difference built-up index in automatically mapping urban areas from TM imagery, Int. J. Remote Sens. 24 (3) (2003) 583–594, doi:10.1080/01431160304987.
[9] X.L. Chen, H.M. Zhao, P.X. Li, Z.Y. Yin, Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes, Remote Sens. Environ. 104 (2) (2006) 133–146, doi:10.1016/j.rse.2005.11.016.
[10] S. Liang, Narrowband to broadband conversions of land surface albedo I: algorithms, Remote Sens. Environ. 76 (2) (2001) 213–238, doi:10.1016/S0034-4257(00)00205-4.
[11] A.R. Huete, H.Q. Liu, K.V. Batchily, W.J.D.A. Van Leeuwen, A comparison of vegetation indices over a global set of TM images for EOS-MODIS, Remote Sens. Environ. 59 (3) (1997) 440–451, doi:10.1016/S0034-4257(96)00112-5.
[12] V. Ihlen, K. Zanter, Landsat 8 Data Users Handbook, US Geological Survey 8 (2019) 114 https://landsat.usgs.gov/documents/Landsat8DataUsersHandbook.pdf.
[13] F.A. Abir, S. Ahmmed, S.H. Sarker, A.U. Fahim, Thermal and ecological assessment based on land surface temperature and quantifying multivariate controlling factors in Bogura, Bangladesh. Heliyon 7 (9) (2021) e08012, doi:10.1016/j.heliyon.2021.e08012.
[14] Z.L. Li, B.H. Tang, H. Wu, H. Ren, G. Yan, Z. Wan, I.F. Trigo, J.A. Sobrino, Satellite-derived land surface temperature: current status and perspectives, Remote Sens. Environ. 131 (2013) 14–37, doi:10.1016/j.rse.2012.12.008.

[15] J.A. Sobrino, J.C. Jiménez-Muñoz, L. Paolini, Land surface temperature retrieval from LANDSAT TM 5, Remote Sens. Environ. 90 (4) (2004) 434–440, doi:10.1016/j.rse.2004.02.003.
[16] T.N. Carlson, D.A. Ripley, On the relation between NDVI, fractional vegetation cover, and leaf area index, Remote Sens. Environ. 62 (3) (1997) 241–252, doi:10.1016/S0034-4257(97)00104-1.
[17] L. Anselin, I. Syabri, Y. Kho, GeoDa: an introduction to spatial data analysis, in: M. Fischer, A. Getis (Eds.), Handbook of Applied Spatial Analysis, Springer, Berlin, Heidelberg, 2010, doi:10.1007/978-3-642-03647-7_5.
[18] A. Guo, J. Yang, W. Sun, X. Xiao, J.X. Cecilia, C. Jin, X. Li, Impact of urban morphology and landscape characteristics on spatiotemporal heterogeneity of land surface temperature, Sustainable Cities and Society 63 (2020) 102443, doi:10.1016/j.scs.2020.102443.
[19] S. Das, D.P. Angadi, Land use-land cover (LULC) transformation and its relation with land surface temperature changes: a case study of Barrackpore Subdivision, West Bengal, India, Remote Sens. Applicat.: Society and Environ. 19 (2020) 100322, doi:10.1016/j.rsase.2020.100322.
[20] C. Zhao, J. Jensen, Q. Weng, R. Weaver, A geographically weighted regression analysis of the underlying factors related to the surface urban heat island phenomenon, Remote Sens (Basel) 10 (9) (2018) 1428, doi:10.3390/rs10091428.
[21] Nakaya, T., Fotheringham, S., Charlton, M. and Brunsdon, C., 2009. Semiparametric geographically weighted generalised linear modelling in GWR 4.0.
[22] Y. Gao, J. Zhao, L. Han, Exploring the spatial heterogeneity of urban heat island effect and its relationship to block morphology with the geographically weighted regression model, Sustainable Cities and Society 76 (2022) 103431, doi:10.1016/j.scs.2021.103431.