MODELLING URBAN HEAT ISLAND USING REMOTE SENSING AND CITY MORPHOLOGICAL PARAMETERS

M. P. Taheri Otaghsara, H. Arefi [*]

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
(peyman.taheri, hossein.arefi)@ut.ac.ir

KEY WORDS: Remote sensing, Urban heat island, LIDAR, Partial least square, Landsat, Sky view factor, Land surface temperature

ABSTRACT:

The aim of this study was to model the surface urban heat island (SUHI) based on remote sensing data, urban morphology parameters and partial least square (PLS) regression using Santa Rosa, California, USA as a case study. Night-time Land surface temperature (LST) was estimated for all the available Landsat 8 night-time data from august to November for the year 2013. Urban morphology parameters such as Building Volume (BV) and Sky View Factor (SVF) were calculated using available LIDAR data of the study area and Normalized Difference Vegetation (NDVI), Index based Built-up Index (IBI) were calculated using Landsat 8 data sets and were used to identify the impact of urban surface characteristics on land surface temperature. Partial least square (PLS) regression analysis was used to assess the correlation and statistically significance of the variables on LST and model the night time LST of the study area. The results of the analysis showed that the LST has a strong positive relationship with IBI and BV and negative relationship with SVF and NDVI and also at night-time, results showed that SVF has a stronger impact on LST differences than NDVI where areas with high density trees had higher temperature than obstacle free vegetated areas. The result of night time LST modelling of the study area was R-square with value of 0.81 between estimated and predicted LST and RMSE with value of 1.02°C.

1. INTRODUCTION

Urbanization is referred to the conversion process of suburban agricultural based communities to industrial based urban communities and is the most apparent aspect of landscape modification caused by human. Urbanization or urban growth has increased the level of human life quality in many ways but it has caused some problems as well. For instance urbanization is causing problems like vegetation and agricultural lands reduction, changes of physical properties of lands and air pollution. The urban heat island (UHI) phenomenon, where urban areas experience higher temperature than their rural surroundings, is the best example of anthropogenic climate modifications caused by urbanization. UHI's have impact on energy use, water consumption, air quality, human health (Brazel, 2007a) and pretty much everything that is linked to human comfort in city life. This phenomenon was first discovered by Luke Howard in 1810's in London and since then, numerous researches have been conducted in various cities in different climates around the globe. The term urban heat island is mostly associated with air temperature data's collected two meters above ground level using weather stations (Emmanuel, 2012a). However with the improvements in thermal infrared remote sensing technology, more studies have been conducted for surface urban heat island (SUHI) monitoring by estimating land surface temperature (LST) using different space-borne TIR sensors. TIR Remote sensing has a major role in monitoring SUHI intensity and its spatial distribution and Monitoring LST is of prime importance for studying SUHI's because it is directly connected to surface radiation, energy exchange, internal climate of buildings and human comfort in cities. On the other hand urban morphology can influence the local urban climates through morphological parameters (Middel and Ariane, 2014a), in particular, urban areas with high density buildings and more complex morphologies exhibit more intense heat islands and lower permeability of urban air ventilation. In 2017, Fan et al. (Fan, 2017a) aimed to measure the oasis effect and its relation

with spatial and temporal change of land covers for four different arid cities using Landsat datasets. They discovered that urban areas in arid cities experience lower surface temperature compared to rural surrounding in all four arid cities. In addition, they found that urban expansion has negative relation with this cooling effect in arid cities. In 2018, Wang et al. (Wang, 2018a) focused on temporal and spatial changes and pattern of UHI's in 32 different cities in China using MODIS thermal and land cover products. They found that physical and geographical differences between cities in north and south is one of the most important factors in SUHI differences in cities. In 2017, Estoque et al. (Estoque, 2017a) aimed to monitor UHI in a tropical city using Landsat datasets from 1987 to 2015. They showed that spatial pattern of land covers reflect the spatial pattern of LST and the correlation between impervious surface (IS) and LST tend to increase through the years.

This study aims to model the night-time urban heat island using morphological parameters, remote sensing data and partial least square regression. For this purpose, all Landsat 8 OLI data's were corrected radiometrically and atmospherically and then used to derive brightness temperature to estimate land surface temperature. LIDAR data was used to derive DEM, DSM and NDSM, and then these data's were used to calculate SVF and BV. Then, NDVI and IBI and city morphological parameters were used as variables of the regression to model and predict the night time LST of the study area. In continuance, some pixels of the study area were randomly chosen to validate the model developed by PLS. Results of partial least square regression showed that LST difference of the study area at night-time could be estimated with the accuracy of 81% using remote sensing and morphological parameters

2. MATERIALS AND METHOD

To achieve the objective of this study, the following steps were taken: (1) collection of Landsat 8 and LiDAR data's of the study

area, (2) radiometric and atmospheric correction of Landsat 8 data's, (3) DSM, DEM and NDSM derivation from LiDAR data set, (4) urban morphology (SVF and BV) calculation from digital surface models, (5) LST estimation from Landsat radiometrically and atmospherically corrected datasets, (6) resampling of the derived parameters to 100 meters, (7) calculation of the relationship between the parameters and LST, (8) modelling the LST of the study area using PLS regression and (9) validation of the model produced by PLS regression. The schema in Figure 1 shows the major steps of the method used in this study.
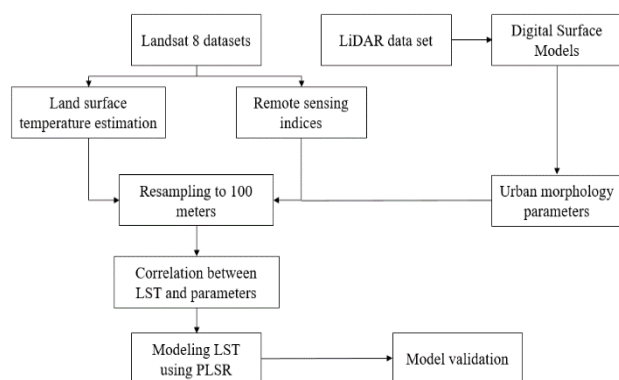

Figure 1. Flowchart of the method used in this study

2.1 Study area

In this study, Santa Rosa, CA, USA was chosen as the study area. Santa Rosa is a city in the county seat of Sonoma County, in California's Wine Country with the population more than 170,000 on 2013. This city has a warm-summer Mediterranean climate with cool, wet winters and warm, dry summers and Its Average annual rainfall is 818 mm.
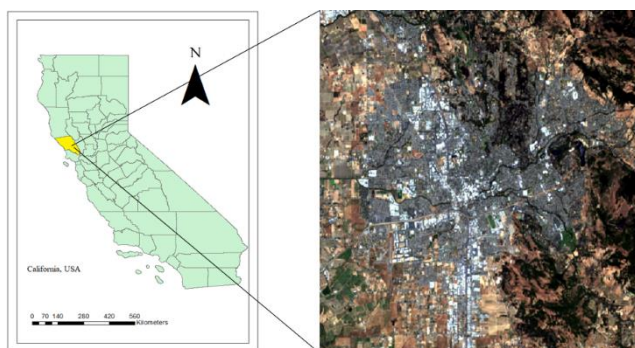

Figure 2. Study area

2.2 Data set

Based on the availability of Landsat 8 Night-Time data, four data set of the study area were collected on 14 August 2013, 15 September 2013, 17 October 2013 and 2 November 2013 from the United States Geological Survey web page (https://earthexplorer.usgs.gov/). Also, LIDAR data for the study area was downloaded from opentopoghraphy open source website (https://opentopography.org/). Table 1 shows the collected and used datasets in this study.

Table 1. Datasets used in this study

Sensor	Date	Acquisition time
	Night	Night
Landsat 8	2013.08.14	05:58:12
	2013.09.15	05:58:10
	2013.10.17	05:58:02
	2013.11.02	05:57:56
Lidar	2013	

2.3 Data processing

2.3.1 NDVI

The NDVI value shows the presence of vegetation or vegetation condition. NDVI value can be derived by using reflectance of the red band (band 4 in OLI) and surface reflectance of the near-infrared band (band 5 in OLI). The equation is as follows (Kriegler, 1969a):

$$NDVI = \frac{NIR - red}{NIR + red} \tag{1}$$

Where   NIR band = Near Infrared band

2.3.2 Index based Built-Up Index (IBI):

IBI was created by using soil adjusted vegetation index (SAVI) (Huete, 1988a), modified normalized difference water index (MNDWI) (Gao, 1996a) and normalized difference built-up index (NDBI) (Zha, 2003a) indices. This index can distinguish the built-up land features easily because the subtraction of the SAVI band and the MNDWI band from the NDBI band (Xu, 2008a). The equation is as follows:

$$NDBI = \frac{SWIR1 - NIR}{SWIR1 + NIR} \tag{2}$$

$$SAVI = \frac{(NIR - red)(1 + l)}{(NIR + red + l)} \tag{3}$$

$$MNDWI = \frac{GREEN - NIR}{GREEN + NIR} \tag{4}$$

$$IBI = \frac{[NDBI - (SAVI + MNDWI)/2]}{[NDBI + (SAVI + MNDWI)/2]} \tag{5}$$

Where   SWIR = Short Wave Infrared band
NIR = Near Infrared band
l = Correction factor between 0 to 1, in this study = 0.5

2.3.3 Sky View Factor (SVF):

SVF consists of a dimensionless value between 0 and 1 obtained as a ratio between the radiation received by a planar surface and that from the entire hemispheric radiating environment (Watson and Johnson, 1987a). Low values of SVF are due to buildings that cover the sky and other high obstacles which intercept the outgoing radiation; on the other hand, high values of SVF indicate a flat and obstacle-free surface. For calculating SVF, DSM of the study area was first resampled to 30 meters and then used to calculate the SVF. The equation is as follows (Böhner and Antonic, 2009a):

$$SVF = \frac{1}{2\pi} \int_0^2 [\cos\beta \cos^2\varphi + \sin\beta \cos(\emptyset - a)(90 - \varphi - \sin\varphi \cos\varphi)]d\emptyset \quad (6)$$

Where
β = surface slope angle
a = surface aspect angle
φ = horizon angle
$\emptyset$ = azimuth direction

2.3.4 Building volume (BV):

BV was calculate using NDSM of the study area. NDSM only has the height of the buildings, so by multiplying the value of every pixel of the NDSM by their area, building volume will be calculated. In this study the spatial resolution of the NDSM is 100 meters.

2.3.5 Land Surface Temperature estimation:

For estimation of land surface temperature (LST) in this study, first the DN number of TIR bands (band 10 and 11 in Landsat 8) were converted to at-sensor radiance using the following equation (Chandler and Groeneveld, 2009a):

$$L_\lambda = \frac{(L_{max,\lambda} - L_{min,\lambda})}{(Q_{cal,max} - Q_{cal,min})}(Q_{cal} - Q_{cal,min}) + L_{min,\lambda} \quad (7)$$

Where
L_λ is spectral radiance ($Wm^{-2}sr^{-1}\mu m^{-1}$)
$L_{max,\lambda}$ and $L_{min,\lambda}$ are spectral at-sensor radiance
$Q_{cal,max}$ and $Q_{cal,min}$ are maximum and minimum quantized calibrated pixel values corresponding to $L_{max,\lambda} - L_{min,\lambda}$ respectively

It should be note that because Landsat 8 science team does not recommend using band 11 in split-window algorithm due to large calibration uncertainty associated with it (USGS 2014), in this study single-channel algorithm based on band 10 was used to estimate LST of study areas. In the next step, by using equation below, the radiance of band 10 was converted to brightness temperature:

$$T_b = \frac{K_2}{\ln(\frac{K_1}{l_\lambda}+1)} \quad (8)$$

Where
$K_2 = 1321.0789 \; w/(m^2 sr \; \mu m)$
$K_1 = 774.8853 \; w/(m^2 sr \; \mu m)$

In the next step, NDVI threshold method [9] was used to calculate emissivity for each pixel. In literature, it has been shown that NDVI threshold method achieves root mean deviation with value of less than 0.01 on emissivity estimation (Sobrino, 2004a). In NDVI-based method it is assumed that a pixel is bare soil if its NDVI is less than 0.2 and it is considered vegetation if its NDVI value is bigger than 0.5 and it obtains emissivity for each presumed land use type. For pixels with NDVI between 0.2 and 0.5, it is assumed that those pixels are mixture of bare soil and vegetation and proportion of vegetation (P_v) is calculated to estimate the emissivity for those pixels. The equation is as follows (Carlson, 1997a):

$$P_v = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right)^2 \quad (9)$$

By using ASTER spectral library which result in an emissivity with values of 0.97 and 0.99 for bare soil and vegetation respectively and following (Yu, 2014a), the emissivity for each pixel in Landsat 8 is calculated using the following equation:

$$E_v = 0.00149P_v + 0.98481 \quad (10)$$

And in the last step, by using brightness temperature and land surface emissivity, LST for each pixel was calculated using the equation below (Jiménez-Muñoz, 2006a):

$$T = \frac{T_b}{1 + \left(\lambda * \frac{T_b}{C_2}\right)\ln\varepsilon} \quad (11)$$

Where
T_b : brightness temperature
λ : Wavelength of band 10 (10.8 µm)
C_2 = h*c/s
h : Planck's constant ($6.624 * 10^{-34}$ J s)
c : Velocity of light ($2.998 * 10^8$ m/s)
s : Boltzmann constant ($1.38 * 10^{-23}$ J/K)
ε : land surface emissivity

2.3.6 Regression Analysis for model development:

Partial least square regression is a statistical method that bears some relation to principle component regression but instead of finding hyper planes of maximum variance between response and independent variables, it finds a linear regression by projecting independent and response variables to a new space and subsequently reduces the collinearity and noise within dataset. This two-step approach moderates the predictors of a dataset to a smaller set of uncorrelated components, and performs least square regression on these components instead of the original data (Ryan, Kimberly and Khalid Ali, 2016a). Partial least square regression has been applied successfully in numerous remote sensing study such as estimation of grass/herb biomass (Cho MA, 2007a) and estimation of lunar soil composition (Li L, 2006a). These studies discovered that use of PLSR method is preferred over principle component regression for analysis. In this study, for modelling day-time and night-time LST, partial least square regression was used. The parameters such as NDVI, SVF, IBI and BV were used as independent variables and LST was considered the response variable. Partial Least square (PLS) regression was carried out to determine how well the estimated LST could be explained by the combination of the parameters and indices that were calculated in this study. The regression results give insights about the effect of different variables on LST at different points of the study area. The coefficient of determination (R^2) indicates the proportion of variation in LST that could be explained by the regression models; the standardized coefficients (Beta coefficients) of predictive models indicate the contributions of different variables to the LST difference. The RMSE can be calculated by using reference and predicted values. RMSE calculation equation is as follows (Laurent, 1998a):

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(e_i - v_i)^2\right]^{\frac{1}{2}} \quad (8)$$

Where
v_i = Reference values
e_i = predicted values

In statistics, RMSE is a frequently used measure of the differences between values predicted by a model or the values

actually observed from the thing being modelled or estimated (Laurent, 1998a).

3. RESULTS

3.1 Remote sensing indices and morphological parameters

As mentioned above, remote sensing indices and morphological parameters were calculated by the equations mentioned and are shown in Figure 3. As shown in Figure 3, in this city, low sky view factor is mostly due to buildings in the urban areas. Also in some parts of the suburban area low SVF is due to tall and high density trees which have the most values of NDVI as well.
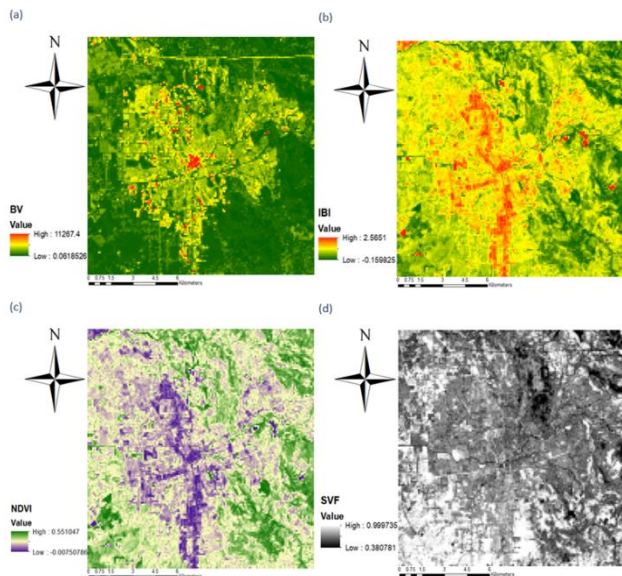

Figure 3. (a) BV; (b) IBI; (c) NDVI; (d) SVF

3.2 Estimation of land surface temperature

For estimating the land surface temperature, first all data sets were corrected atmospherically and radiometrically. Land surface temperature of the study area was calculated using single-channel method explained in the previous section. It should be note that for minimizing the influence of precipitation and seasonal effect on LST, the average LST of all months were used to investigate the LST differences in all the study areas. Figure 4 shows the average of estimated land surface temperature of the observed city.

3.1 Correlation between LST and parameters

As shown in Figure 5, LST had positive relationship with BV and IBI which both these parameters show the urbanized parts of the study area indicating that at night, these parts of the study area which are urban areas have higher LST comparing to the suburban parts.
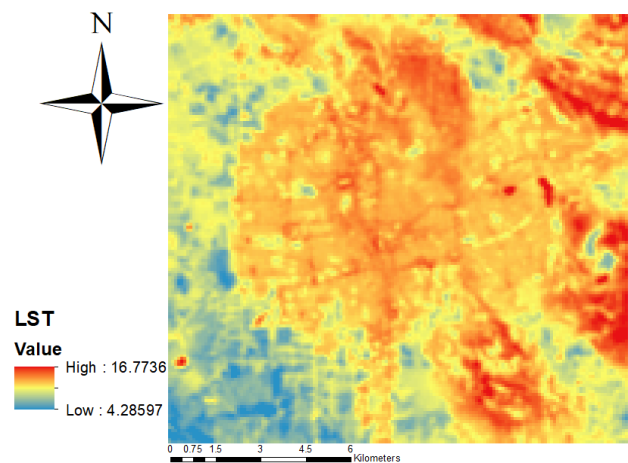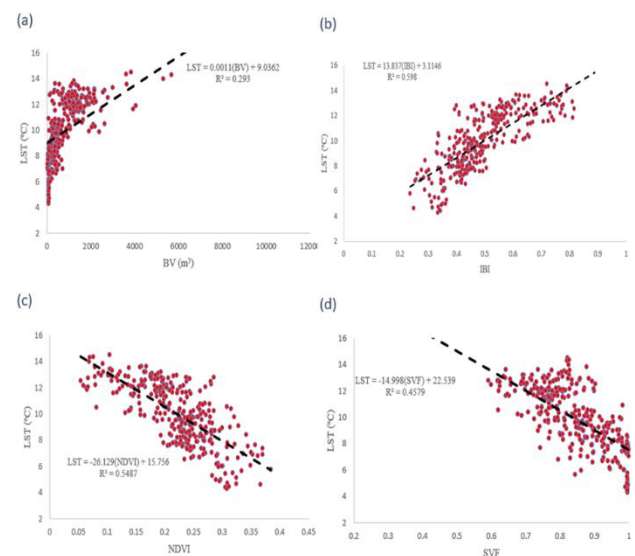

Figure 4. Average LST of the study area

Figure 5. Correlation between LST and (a) BV; (b) IBI; (c) NDVI and (d) SVF

Also LST had a negative relationship with SVF and NDVI indicating that wherever SVF was high which indicates an obstacle free environment and in fact suburban areas, LST was low. Moreover wherever NDVI was high LST was low too. Although in some places of the city where there were high vegetation like dens trees, SVF were low, and the LST of those area were high, which indicates that at night, SVF have larger effect on LST than NDVI. Table 2 shows the relation between the four parameters and LST during night time. Also t-test and p.value are in a meaningful ranges of values which means that modelling the LST with these parameters is not irrelevant.

Table 2. Statistical results of the parameters used in this study

Variables	R^2	t-test	p.value
BV	0.29	1.506	0.133e-8
IBI	0.59	8.228	2.665e-14
SVF	0.45	-18.73	0.003e-11
NDVI	0.54	-2.954	7.442e-69

3.2 Partial least square regression results

As mentioned before, for modeling the LST of the study area and assessing the effect of calculated parameters on it, PLS regression was used. Table 2 shows the values PLS regression coefficients.

Table 3. PLS regression coefficient

Variables	BV	IBI	SVF	NDVI
Regression coefficient	6.654	4.157	-12.241	-4.401

Equation 9 shows the prediction model for LST by four variables used to develop the model. For validation of the model developed, some pixels of the study area were chosen randomly for testing. As shown in Table 3, SVF's beta coefficient value was bigger than NDVI's which indicates that at night-time for this city, SVF has a stronger impact on night0time LST differences.

$$LST = -4.401 * NDVI - 12.241 * SVF + 6.654 * BV + 4.157 * IBI \qquad (9)$$
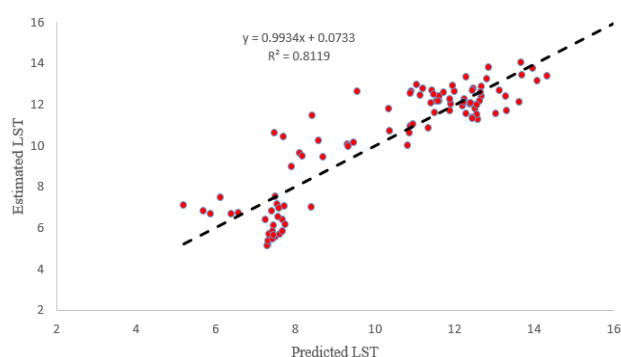

Figure 6. Scatterplot of the predicted and estimated LST

For testing the model developed by PLSR in this study, some pixels were randomly chosen as test pixels and then their parameters were put in the PLSR model and LST predicted by the model was compared to the estimated LST using single-channel method. Figure 6 shows the scatterplot of the predicted and estimated LST of the study area. It can be understand from Table 3 that 81% of the night time LST can be accurately predicted with the RMSE value of 1.02 ºC in radius of 100 m using the model developed by PLS regression. Table 4 shows the accuracy of the predicted LST using the PLS regression model.

Table 4. The accuracy of the predicted LST using PLSR model

Predicted LST	R^2	RMSE	Relative RMSE
	0.81	1.02 (ºC)	0.092

4. CONCLUSION

The aim of this study was to model the urban heat island (UHI) based on remote sensing data, urban morphology parameters and partial least square (PLS) regression using Santa Rosa, California, USA, as a case study. In order to reach the goal of this study some remote sensing parameters like NDVI and IBI and some city morphological parameters were calculated and their relation with night-time LST of the study area were investigated.

The result showed that both SVF and NDVI had negative relation with night-time LST. Also the regression model showed that at night-time for the observed city, SVF have a bigger role compared to NDVI where areas with high density trees were experiencing higher LST compared to other rural areas with both high NDVI and SVF. In conclusion, the parameters and indices which were used as variable of the regression have a good performance for modelling the night-time LST differences in the study area (Table 3). Also some more morphological parameters could improve the performance of PLS regression. IBI and BV, which both indicate the urbanized areas, had positive relationship with night time LST that indicates urban areas have higher LST comparing to suburban areas at night which means that SUHI is occurring at night for Santa Rosa city. Furthermore, night-time remote sensing datasets in every month of every season of cities in different climates can lead to a better understanding of the correlation between the morphological parameters and remote sensing indices and urban heat islands in different conditions.

REFRENCES

Böhner, J., and O. Antonic. 2009. "Land-Surface Parameters Specific to Topo-Climatology." Developments in Soil Science 33: 195–226. doi:10.1016/S0166-2481(08)00008-1.

Brazel, A.; Gober, P.; Lee, S.; Grossman-Clarke, S.; Zehnder, J.; Hedquist, B.; Comparri, E. Determinants of changes in the regional urban heat island in metropolitan phoenix (Arizona, USA) between 1990 and 2004. Clim. Res. 2007, 33, 171–182.

Carlson, T.N.; Ripley, D.A. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. 1997, 62, 241–252.

Chander, G., and D. P. Groeneveld. 2009. "Intra-Annual NDVI Validation of the Landsat 5 TM Radiometric Calibration." International Journal of Remote Sensing 30 (6): 16211628. doi:10.1080/01431160802524545.

Cho MA, Skidmore A, Corsi F, Wieren, SE, Sobhan I (2007) Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. Int J Appl Earth Obs 9:414–424.

Emmanuel, R., Kruger, E., 2012. Urban heat island and its impact on climate changeresilience in a shrinking city: the case of Glasgow UK. Build. Environ. 53,137–149, http://dx.doi.org/10.1016/j.buildenv.2012.01.020.

Estoque, Ronald C., and Yuji Murayama. "Monitoring surface urban heat island formation in a tropical mountain city using Landsat data (1987–2015)." ISPRS journal of photogrammetry and remote sensing 133 (2017): 18-29.

Fan, Chao, et al. "Understanding the impact of urbanization on surface urban heat islands—a longitudinal analysis of the oasis effect in subtropical desert cities." Remote Sensing 9.7 (2017): 672.

Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 1996, 58, 257–266.

Huete, A. R. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 1988, 25, 295–309.

Jiménez-Muñoz, J.C.; Sobrino, J. Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. Int. J. Remote Sens. 2006, 27, 999–1014.

Kriegler, F. J.; Malila, W. A.; Nalepka, R. F.; Richardson, W. Preprocessing transformations and their effects on multispectral recognition. In Remote Sensing of Environment, VI; 1969; Vol. 1, p. 97.

Laurent, H., I. Jobard, and A. Toma. "Validation of satellite and ground-based estimates of precipitation over the Sahel." Atmospheric Research 47 (1998): 651-670.

Li L (2006) Partial least squares modeling to quantify lunar soil composition with hyperspectral reflectance measurements. J Geophys Res 111(E4):E04002. doi:10.1029/2005JE00259.

Ryan, Kimberly, and Khalid Ali. "Application of a partial least-squares regression model to retrieve chlorophyll-a concentrations in coastal waters using hyper-spectral data." Ocean Science Journal 51, no. 2 (2016): 209-221.

Sobrino, J.A.; Jiménez-Muñoz, J.C.; Paolini, L. Land surface temperature retrieval from landsat TM 5. Remote Sens. Environ. 2004, 90, 434–440.

USGS, 2014. (U.S. Geological Survey) (accessed 9. 09. 14) http://landsat.usgs.gov/calibrationnotices.php.

Wang, Juan, et al. "Mapping Spatiotemporal Patterns and Multi-Perspective Analysis of the Surface Urban Heat Islands across 32 Major Cities in China." ISPRS International Journal of Geo-Information 7.6 (2018): 207.

Watson, I. D., and G. T. Johnson. 1987. "Graphical Estimation of Sky View-Factors in Urban Environments." Journal of Climatology 7: 193–197. doi:10.1002/joc.v7:2.

Xu, H. A new index for delineating built-up land features in satellite imagery. Int. J. Remote Sens. 2008, 29, 4269–4276.

Yu, X.; Guo, X.; Wu, Z. Land surface temperature retrieval from landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method. Remote Sens. 2014, 6, 9829–9852.

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