International Journal of Engineering and Geosciences (IJEG),
Vol;4, Issue; 1, pp. 001-007, February, 2019, ISSN 2548-0960, Turkey,
DOI: 10.26833/ijeg.404426

OUTLIER DETECTION OF LAND SURFACE TEMPERATURE: KÜÇÜKÇEKMECE REGION

Lütfiye Kuşak [1]* Ufuk Fatih Küçükali[2]

[1]Istanbul Aydin University, Architecture and Design Faculty, Department of Interior Architecture, Istanbul, Turkey (lutfiyekusak@aydin.edu.tr); **ORCID ID 0000-0002-7265-245X**

[2]Istanbul Aydin University, Architecture and Design Faculty, Department of Architecture, Istanbul, Turkey (ufkucukali@aydin.edu.tr); **ORCID ID 0000-0002-2715-7046**

*Corresponding Author, Received: 12/03/2018, Accepted: 25/05/2018

ABSTRACT: Unplanned and rapid urbanization is one of the reasons for the rising surface temperature values in urban areas. There is a considerable amount of literature demonstrating the association of urbanization with surface temperatures. Küçükçekmece Lake, an important lake which has been meeting utility water needs of Istanbul, and unplanned and rapid urbanization around it have been affected by this inevitable change for years. Although surface temperatures generally correlate strongly with each other, very high and very low temperature values should not be disregarded and need to be investigated. The current study was conducted with the assumption that these values could be outlier values and thus they were analyzed using the Box Plot method for the selected region. Correlations between Land Surface Temperature (LST) values obtained for Küçükçekmece and its vicinity was examined using Landsat OLI images of June 20, 2016 and June 23, 2017, and LST outliers and regions with common outliers of/on both days were determined. In the study, 310 LST outliers were identified for June 20, 2016 and 34 LST outliers for June 23, 2017, and in both images, 33 outliers were found to be common and they clustered in two different buildings. The reasons for the outliers outside the standard surface temperature values and recommended solutions were discussed.

Keywords: *LST, Outlier, Box Plot, Urbanization, Land use/cover*

1. INTRODUCTION

According to the data for January 2018 from the World meters website (URL 1), 53.25% of the world's population, which is about 7.5 billion people, live in the cities. Rapid population growth in cities, which is caused by the fact that human population abandons rural areas and immigrates to cities, results in changes in land use as well as unplanned or poorly planned urbanization (Deng, et al. 2009), leading to many ecological (Zhao, et al., 2006; Wang, et al., 2008), climatic (Liu, et al., 2017), sociological (Vlahov and Galea 2002) and environmental problems (Liao, et al., 2017). Changes in land use/cover and the resultant problems have been analyzed using ecological (Nacef, et al., 2016), economic, sociological data and remote sensing images (Shen, et al., 2016).

One of the biggest problems brought about by rapid urbanization and consequent changes in land use/cover are changes in surface temperatures in intensely urbanized areas (Weng 2001; Jenerette, et al., 2007; Cui and Shi 2012; Tayyebi, et al., 2018). Numerous studies have been conducted to observe these changes and determine whether there is relevance between land use/cover and the change in land surface temperature (Chen, et al., 2006, Hasanlou and Mostofi 2015; Alhawitti and Mitsova 2016; Kirtiloglu, et al., 2016; Li, et al., 2017).

The developments in remote sensing technology allow land surface temperatures (LST) to be obtained using various satellite and airborne sensors that support acquisition of thermal infrared information (Kaya, et al., 2012; Chen, et al., 2013; Orhan, et al. 2014; Orhan and Yakar 2016; Gunawardena, et al., 2017; Ranagalage, et al., 2017).

Atmospheric conditions, geographical factors, and urbanization affect surface temperature values. However, as analyzed in this study, it is observed that the LST values are affected by unusual local changes and phenomena such as very low or very high values that may be called outliers.

Outliers, which are often used in the field of statistics, are widely used in the field of machine learning owing to developments in software and hardware technology that allow storage and processing of much more data. Outliers, also called abnormalities, discordants, deviants, anomalies, are employed in a wide range of activities, including data security, medical diagnosis, in various fields of physics, network security, credit card and insurance fraud detection, critical systems and defense activities, and earth science. (Iglewicz, and Hoaglin 1993; Bramer, 2007; Kumar and Mathur 2014, URL 2, URL 3). Statistical or model based models, Proximity Based Models, Linear Regression Models (PCA, LMS), Information Theory Models, High Dimensional Outlier Detection Methods and many other methods were developed for the detection of outliers, in which data type and data size (univariate / multivariate) are very important in the selection of the method to be used. (Seo 2006). As it has already been emphasized above, the size of the data is very important for the method to be used. Statistical or model based models such as Standard Deviation (SD), Z-scores, and Box Plot are preferred for the detection of outliers for unvariate data (Olsonand and Delen 2008; Han, et al., 2011).

A box plot is a graphical method of displaying variation in a group of numerical data. When literature studies are examined, box plot is widely used in environmental studies (Xie et al., 1999, Zhang et al., 2009), urban studies (Cipolla and Maglionico 2014) and other scientific studies for outlier detection.

The outlier obtained in some studies may contain highly valuable information (Rousseeuw and Hubert 2017). Since non-spatial data such as temperature, pressure, and image pixel color intensity are obtained based on spatial locations, unusual local changes in these data can be seen as outliers (Aggarwal 2015).

In this study, The LST values obtained with Landsat 8 OLI satellite images were prepared considering that unusual local changes might affect them. Since the obtained LST values are univariate, the Box plot method (Tukey 1997) is used in this study. In addition, Box plot is used in this study because it is less sensitive to extreme values (Mendenhall and Sincich 2016).

The IQR (Interquartile Range) calculations used to construct the Box Plot were made and the outlier values obtained from the analysis were discussed and evaluated at the end of the study by taking into account the spatial characteristics (anthropogenic effect, slope, aspect, water cooling island (WCI (Du, et al., 2016)).

2. DATA AND METHODS

The study comprises three main stages, including acquisition of NDVI, LST images and values of indices, determination of LST values causing outliers by Box Plot method and discussion of potential reasons which might be responsible for LST values described as outliers.

2.1 Study Area

An area within a distance of 6 km from the Küçükçekmece Lake and its coastline was chosen as the study area. The Küçükçekmece District is situated on the shores of Marmara Sea on the European side of Istanbul (Fig 1). The study area, located between 40º5730 N-41º0630 N and 28º390 E-28º5030 E coordinates.

As stated in 1/100,000 environmental plan of Istanbul, in accordance with its development in the east-west direction, the study area includes Küçükçekmece Lake, residential areas, industrial areas, agricultural areas, sports facilities, educational facilities and airport functions. As is the case in the entire Istanbul, the study area has been faced with highly intense urbanization for the last 30 years (Kucukali and Kuşak 2017).

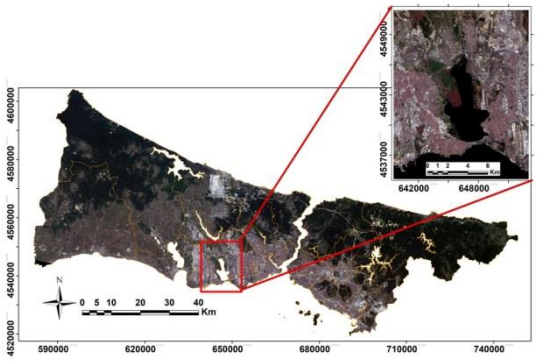

Figure. 1 Study area

2.2 Data

In the study, June 20, 2016 and June 23, 2017 Landsat 8 OLI images were used to obtain LST values and determine outlier values. Using Landsat 8 OLI image, 20 June 2016 and 23 June 2017 prepared by EROS Science Processing Architecture (ESPA) on demand interface that provides Landsat higher-level science data products, including Climate Data Records (TOA reflectance, brightness temperature, cloud masks, and surface reflectance) and spectral indices (e.g. NDVI, EVI, SAVI, and NBR). Processed bands, surface reflectance and NDVI were used for this paper. The Landsat 8 satellite payload consists of two science instruments—the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS).

Additionally, the slope and aspect properties obtained from ASTER DEM image, Istanbul environmental plan of 1/100,000 scale, and report, current Google Earth images were employed to interpret the relationship between outlier values and land use.

2.3 Methods

Digital Numbers (DN) values for OLI and TIRS bands should first be converted to spectral radiance values for Top of Atmosphere (TOA) in order to obtain LST images and spectral indices. The LST values are calculated using the brightness temperature (BT) values. When BT values are calculated, TOA data and thermal conversation constant values are utilized. Since the file prepared by ESPA contains TOA reflectance, brightness temperature, surface reflectance NDVI, these steps were omitted. The calculation of the LST was made using the data prepared by ESPA Eq. (1).

LST based on satellite brightness temperature (BT) was computed by the following equation. Therefore, emissivity values should be determined.

$$LST = \frac{BT}{1+(\lambda+\frac{BT}{\alpha})\ln\varepsilon} \quad (1)$$

BT is the effective at-satellite temperature in Kelvin, λ is the wavelength of the emitted radiance in meters, α=1.438x10-2mK; and ε is the surface emissivity.

There is a relationship between land cover and emissivity. The use of NDVI has often been the preferred method for demonstrating this relationship. (Van de Griend and Owe 1993; Valor and Caselles, 1996). Then NDVI based threshold method was developed for NDVI based studies (Sobrino, et al., 2001; Sobrino, et al., 2004). In this method, three main threshold ranges, consisting of soil, vegetation and mixed areas, were determined. Certain emissivity values are used for soil and vegetation, but they are created with the emissivity formula calculated using NDVI values for mixed areas. Today, the three main threshold ranges defined by Sobrino (2004) are increased and different emissivity values and calculation methods for these ranges are being developed. (Stathopoulou, et al., 2007; Stathopoulou and Cartalis 2007 ; Xie, et al., 2012; Tang, et al., 2015).

In this study, the method adopted by Shen, which is used in LST calculation, is applied. In this study, 0.9923 emissivity value for water areas, 0.923 emissivity value for urban impervious and bare soil areas and 0.986 emissivity value for vegetation areas are accepted. The formula Eq (2) for mixed areas is also used (Shen et al., 2016).

$$\varepsilon = 1.0094 + 0.047\ln(NDVI) \quad (2)$$

Obtaining the pixel values for the LST images is another step of the study in order to make outlier detection. The tools in ARCGIS 10.2 were exploited to obtain pixel values. With the help of spatial analysist tools used concurrently, multipoint values were obtained from the clipped images. LST, NDVI, NDBI, Slope (URL 4), Aspect (URL 5) pixel values and distance of their spatial location to the lake were obtained using spatial analysis tool for both dates. During interpretation of the values obtained, first, correlations between data were investigated by forming correlation matrix.

Box Plot was used to determine the pixel values of the obtained values, which show temperature values contrary to the data set, and to discuss the situations that could be caused by these outliers. MS Excel and RStudio software are preferred for calculating outliers. IQR is used in Box Plot method developed by Tukey. The interquartile range is the range between the first and the third quartiles (the edges of the box). Tukey took into account any data point that fell outside of either 1.5 times the IQR below the first – or 1.5 times the IQR above the third – quartile to be "outside" or "far out" (Tukey 1977).

3. RESULTS

NDVI is an index often used in detection of live vegetation in areas. NDVI values of the study area from two different days were evaluated. Accordingly, it is clear that the NDVI values were close to 1 around and in northern parts of the lake in both images (Fig 2). NDVI is also used intensively to calculate LST value.

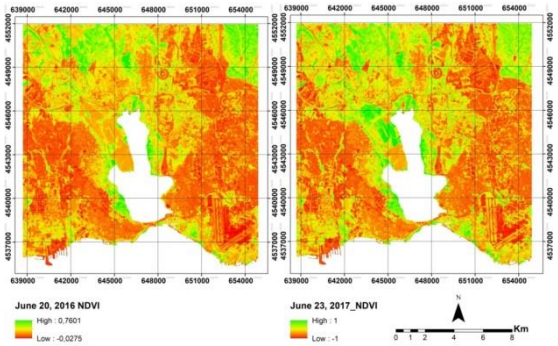

Figure. 2 NDVI images

LST images generated using Eq (1) and Eq (2) as well as other emissivity values (Fig 3) showed that the surface temperature of close vicinity of the lake was lower than those of inner parts of the study area, as expected. Similarly, the surface temperature of northern parts, where NDVI value was close to 1, was low too. On the other hand, the surface temperature was noted to be high in the eastern and western parts of the lake.

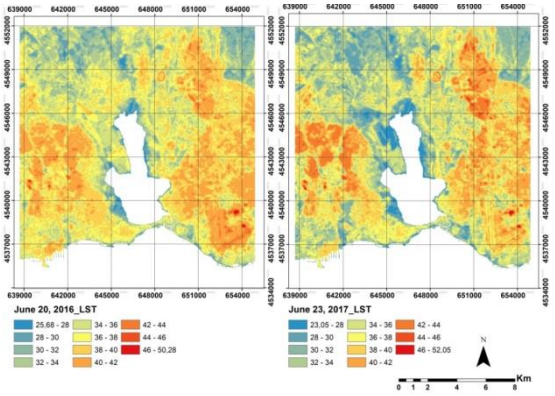

Figure. 3 LST images

Following visual interpretation, 256592 values were obtained by ARCGIS spatial analysist tool in order to determine numeric counterparts of surface temperatures and LST average, StdDev, Min, Max and R2 values of two different LST images were obtained (Table 1). Table 1 demonstrates that there is an average drop of 0.82 ^{0}C in the selected region in LST image for June 23, 2017. The LST correlation for June 20, 2016 was 0.9651, while it was 0.9879 for June 23, 2017.

Table 1 LST values and correlations

Year	Average	Stddev	Min	Max	R^2
6/20/16	36,73	3,27	25,68	50,28	0,9651
6/23/17	35,91	3,93	23,24	52,05	0,9879

The correlations between LST, NDVI, NDBI, Slope, Aspect pixel values and distance of their spatial location to the lake were investigated. As expected, the highest negative correlation was found between NDVI and LST. In addition to this, there is no significant relationship between the other parameters. Moreover, the correlation between LST values of two different years was 0.889.

Despite this high correlation, very low and high LST values were found when the LST values were listed in increasing order. A Box Plot analysis (Fig 4) was performed to determine whether these values were outlier values or not.

According to the results of the Box Plot method, the upper level of LST value was 46. 72°C while the lower level of LST value was 27.12 °C in 2016. For the year 2017, the upper level of LST value was found to be 48.42°C while the lower level of LST value was found to be 23.66 °C.

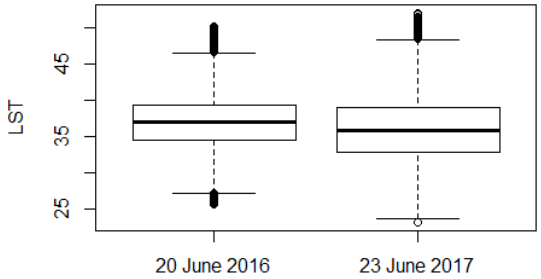

Figure. 4 The results of the Box Plot method

234 data pertaining to the year 2016 were below the lower level, while 84 were above the upper level. A total of 318 outlier data were found.

One data of 2017 was below the lower level, whereas 33 were above the upper level. A total of 34 data were determined as outliers.
After the outlier points determined for the year 2016 were excluded, the correlation value increased to 0.9675, while the correlation value for the year 2017 increased to 0.9883. LST correlations for both days were unchanged and remained at 0.889, compared to the correlations before exclusion of the outlier data. However, when outlier data are focused on (Fig 5, Fig 6), this may lead to results which are too important to neglect, especially in micro scale studies, although there is not any noticeable change.

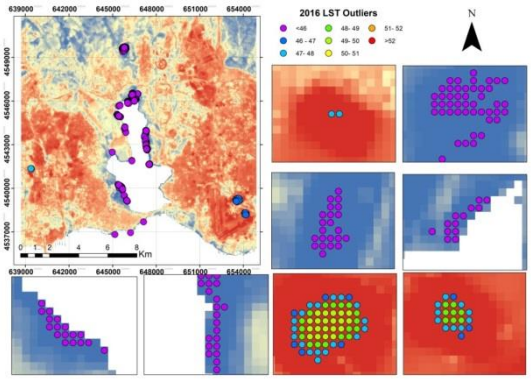

Figure. 5 2016 LST outliers images

When 318 outlier points determined by the box-plot analysis of the LST image from June 20, 2016 were examined, 234 points at the lower level were generally distributed in the lake shoreline, seaside and wetlands,

whereas 84 outlier points were in the Atatürk Airport area, located in the southeastern part of the image.

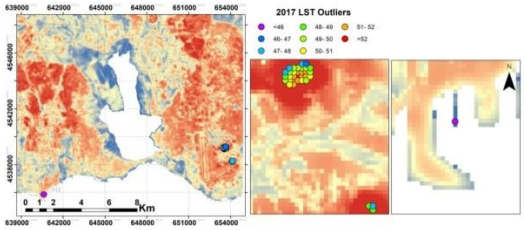

Figure. 6 2017 LST outliers images

When 34 outlier points detected by the Box Plot analysis of the LST image from June 23, 2017 were examined, 1 point at the lower level was in the seaside, whereas 33 outlier points were highly distributed in the Atatürk Airport area, located in the southwestern part of the image.

When the points where the outlier points determined by the Box Plot analysis of the LST images from June 20, 2016 and June 23, 2017 overlapped were examined, it is noteworthy that the points at the lower level did not overlap, whereas 33 points at the upper level overlapped. When position checks of these points were made, these points corresponded to CNR EXPO (First Building), Atatürk Airport Warehouse (Second Building), the reasons for which were then investigated (Fig7).

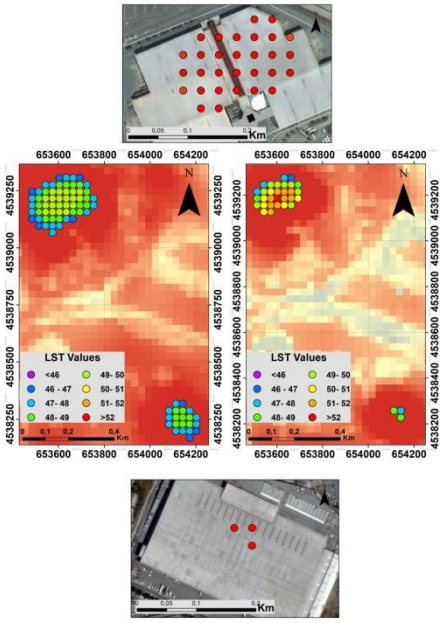

Figure. 7 CNR EXPO (First Building), Atatürk Airport Depo (Second Building) outlier images

4. DISCUSSION

In this study, the Box Plot method developed by Tukey was used since the data used in this study was univariate considering that unusual local changes might have affected the LST values obtained from Landsat 8 OLI satellite images prepared to include the area within a distance of 6 km from Küçükçekmece Lake and its vicinity. (Tukey 1997). The IQR calculations used to construct the Box Plot were made and the outlier values obtained from the analysis were discussed and evaluated at the end of the study by taking into account the spatial characteristics (anthropogenic effect, slope, aspect, water cooling island (WCI)).

According to the results obtained by Box Plot method of LST data of 2016, outlier temperature values lower than 27.12°C and higher than 46.72°C were determined. Visual assessment of these outlier data revealed that particularly the outlier values lower than 27.12 ^{o}C were located in lake shoreline and seaside. This result was ascribed to similar atmospheric conditions, similar geomorphologic properties of the locations of these outliers and cooling effect of water surface. On the other hand, the outliers higher than 46. 72 ^{o}C were densely clustered in southeastern region of the study area (Fig 7). These outliers were located in CNR EXPO (First Building), Atatürk Airport Warehouse (Second Building), which are in the Atatürk Airport region.

Additionally, according to the results obtained by Box Plot method of LST data of 2017, outlier temperature values higher than 48.42 ^{o}C were determined. They matched the locations of 2016 LST outliers by 100%, which were CNR EXPO (First Building), Atatürk Airport Depot (Second Building) buildings, which are in the Atatürk Airport region. This result suggests that the massive size of these buildings, determined according to the results of both analyses, technical properties of their roofing material and positioning of the building's air-conditioning elements have an impact on the formation of outliers.

5. CONCLUSION

In this type of studies, temperature, pressure, wind, and pixel values obtained from satellite images should not be evaluated independently of spatial location characteristics, which would be more significant in interpretation of the analyses.

Therefore, the correlations between slopes, aspect, lake distance, NDVI, NDBI, LST values of the study area were identified in the preliminary phase of the study. However, except for high correlation between generally accepted NDVI, NDBI and LST indices for the selected study area, there was no relationship between slope, aspect, distance to lake and LST so they were excluded from the evaluation, but they were included as secondary data in the interpretation of the analyses.

The structures identified as a result of the examination of the outlier locations determined by Box Plot analysis of the data obtained from the LST images were found to have a very large mass in terms of their architectural design. In such large structures with flat roofs, the reflection and solar orientation design of the roof, planting by appropriate selection of species in the vicinity of the building, selection of proper roofing material and color, green roof applications, correct positioning of air-conditioning elements, and restoration and improvement of structures may help to reduce the LST value.

It is recommended that further research subjects should include continuous observation and control of locations with outliers on a seasonal and annual basis and taking precautions, where necessary, by expanding methodology, scale, spatial and temporal characteristics of the current study.

REFERENCES

Aggarwal, C. C. (2015). Outlier analysis. In Data mining (pp. 237-263). Springer, Cham.

Alhawitti, R. H., and Mitsova, D. (2016). Using Landsat-8 data to explore the correlation between urban heart island and urban land uses. IJRET: International Journal of Research in Engineering and Technology, 5(3), 457-466.

Bramer, M. (2007). Principles of data mining (Vol. 180). London: Springer.

Chen, L., Li, M., Huang, F., and Xu, S. (2013, December). Relationships of LST to NDBI and NDVI in Wuhan City based on Landsat ETM+ image. In Image and Signal Processing (CISP), 2013 6th International Congress on (Vol. 2, pp. 840-845). IEEE.

Chen, X. L., Zhao, H. M., Li, P. X., and Yin, Z. Y. (2006). Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote sensing of environment, 104(2), 133-146.

Cipolla, S. S., and Maglionico, M. (2014). Heat recovery from urban wastewater: Analysis of the variability of flow rate and temperature. Energy and Buildings, 69, 122-130.

Cui, L., and Shi, J. (2012). Urbanization and its environmental effects in Shanghai, China. Urban Climate, 2, 1-15.

Deng, J. S., Wang, K., Hong, Y., and Qi, J. G. (2009). Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. Landscape and urban planning, 92(3-4), 187-198.

Du, H., Song, X., Jiang, H., Kan, Z., Wang, Z., and Cai, Y. (2016). Research on the cooling island effects of water body: A case study of Shanghai, China. Ecological indicators, 67, 31-38.

Gunawardena, K. R., Wells, M. J., and Kershaw, T. (2017). Utilising green and bluespace to mitigate urban heat island intensity. Science of the Total Environment, 584, 1040-1055.

Han, J., Pei, J., and Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

Hasanlou, M., and Mostofi, N. (2015). Investigating urban heat island estimation and relation between various land cover indices in tehran city using landsat 8 imagery. In Proceeding.

Iglewicz, Boris and David Hoaglin (1993), How to Detect and Handle Outliers. American Society for Quality Control, Milwaukee WI.

Jenerette, G. D., Harlan, S. L., Brazel, A., Jones, N., Larsen, L., and Stefanov, W. L. (2007). Regional relationships between surface temperature, vegetation, and human settlement in a rapidly urbanizing ecosystem. Landscape ecology, 22(3), 353-365.

Kaya, S., Basar, U. G., Karaca, M., and Seker, D. Z. (2012). Assessment of urban heat islands using remotely sensed data. Ekoloji, 21(84), 107-113.

Kirtiloglu, O. S., Orhan, O., & Ekercin, S. (2016). A Map Mash-Up Application: Investigation The Temporal Effects Of Climate Change On Salt Lake Basin. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 41.

Kucukali, U. F., and Kuşak, L. (2017). Environmental, Social, and Economic Indicators of Urban Land Use Conflicts. Urbanization and Its Impact on Socio-Economic Growth in Developing Regions, pp. 285, IGI Global,USA.

Kumar, M. and Mathur, R., (2014). April. Outlier detection based fault-detection algorithm for cloud computing. In Convergence of Technology (I2CT), 2014 International Conference for (pp. 1-4). IEEE.

Li, W., Cao, Q., Lang, K., and Wu, J. (2017). Linking potential heat source and sink to urban heat island: Heterogeneous effects of landscape pattern on land surface temperature. Science of the Total Environment, 586, 457-465.

Liao, J., Jia, Y., Tang, L., Huang, Q., Wang, Y., Huang, N., and Hua, L. (2017). Assessment of urbanization-induced ecological risks in an area with significant ecosystem services based on land use/cover change scenarios. International Journal of Sustainable Development and World Ecology, 1-10.

Liu, Y., Chen, Z. M., Xiao, H., Yang, W., Liu, D., and Chen, B. (2017). Driving factors of carbon dioxide emissions in China: an empirical study using 2006-2010 provincial data. Frontiers of Earth Science, 11(1), 156-161.

Mendenhall, W. M., and Sincich, T. L. (2016). Statistics for Engineering and the Sciences. Chapman and Hall/CRC.

Nacef, L., Bachari, N. E. I., Bouda, A., & Boubnia, R. (2016). "Variability and decadal evolution of temperature and salinity in the mediterranean sea surface". International Journal of Engineering and Geosciences, 1(1), 20-29.

Olson, D. L., and Delen, D. (2008). Advanced data mining techniques. Springer Science and Business Media.

Orhan, O., Ekercin, S., & Dadaser-Celik, F. (2014). Use of landsat land surface temperature and vegetation indices

for monitoring drought in the Salt Lake Basin Area, Turkey. The Scientific World Journal, 2014.

Orhan, O., & Yakar, M. (2016). Investigating Land Surface Temperature Changes Using Landsat Data in Konya, Turkey. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 41, B8.

Ranagalage, M., Estoque, R. C., and Murayama, Y. (2017). An urban heat island study of the Colombo metropolitan area, Sri Lanka, based on Landsat data (1997–2017). ISPRS International Journal of Geo-Information, 6(7), 189.

Rousseeuw, P. J., and Hubert, M. (2017). Anomaly detection by robust statistics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.

Seo, S. (2006). A review and comparison of methods for detecting outliers in univariate data sets (Doctoral dissertation, University of Pittsburgh).

Shen, H., Huang, L., Zhang, L., Wu, P., and Zeng, C. (2016). Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of Wuhan in China. Remote Sensing of Environment, 172, 109-125.

Sobrino, J. A., Jiménez-Muñoz, J. C., and Paolini, L. (2004). Land surface temperature retrieval from LANDSAT TM 5. Remote Sensing of environment, 90(4), 434-440.

Sobrino, J. A., Raissouni, N., and Li, Z. L. (2001). A comparative study of land surface emissivity retrieval from NOAA data. Remote Sensing of Environment, 75(2), 256-266.

Stathopoulou, M., and Cartalis, C. (2007). Daytime urban heat islands from Landsat ETM+ and Corine land cover data: An application to major cities in Greece. Solar Energy, 81(3), 358-368.

Stathopoulou, M., Cartalis, C., and Petrakis, M. (2007). Integrating Corine Land Cover data and Landsat TM for surface emissivity definition: application to the urban area of Athens, Greece. International Journal of Remote Sensing, 28(15), 3291-3304.

Tang, B. H., Shao, K., Li, Z. L., Wu, H., and Tang, R. (2015). An improved NDVI-based threshold method for estimating land surface emissivity using MODIS satellite data. International Journal of Remote Sensing, 36(19-20), 4864-4878.

Tayyebi, A., Shafizadeh-Moghadam, H., and Tayyebi, A. H. (2018). Analyzing long-term spatio-temporal patterns of land surface temperature in response to rapid urbanization in the mega-city of Tehran. Land Use Policy, 71, 459-469.

Tukey, J. W. (1977). Box-and-whisker plots. Exploratory data analysis, 39-43.

Valor, E., and Caselles, V. (1996). Mapping land surface emissivity from NDVI: Application to European, African, and South American areas. Remote sensing of Environment, 57(3), 167-184.

Van de Griend, A. A., and Owe, M. (1993). On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. International Journal of remote sensing, 14(6), 1119-1131.

Vlahov, D. and Galea, S. (2002). Urbanization, urbanicity, and health. Journal of Urban Health, 79(1), pp.S1-S12.

Wang, J., Da, L., Song, K., and Li, B. L. (2008). Temporal variations of surface water quality in urban, suburban and rural areas during rapid urbanization in Shanghai, China. Environmental Pollution, 152(2), 387-393.

Weng, Q. (2001). A remote sensing? GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. International journal of remote sensing, 22(10), 1999-2014.

Xie, S., Dearing, J. A., Bloemendal, J., & Boyle, J. F. (1999). Association between the organic matter content and magnetic properties in street dust, Liverpool, UK. Science of the Total Environment, 241(1-3), 205-214.

Xie, Q., Zhou, Z., Teng, M., and Wang, P. (2012). A multi-temporal Landsat TM data analysis of the impact of land use and land cover changes on the urban heat island effect. J. Food Agric. Environ, 10(2), 803-809.

Zhang, C., Tang, Y., Luo, L., & Xu, W. (2009). Outlier identification and visualization for Pb concentrations in urban soils and its implications for identification of potential contaminated land. Environmental Pollution, 157(11), 3083-3090.

Zhao, S., Da, L., Tang, Z., Fang, H., Song, K., and Fang, J. (2006). Ecological consequences of rapid urban expansion: Shanghai, China. Frontiers in Ecology and the Environment, 4(7), 341-346.

URL 1. Web Map Tile Service, http://www.worldometers.info [Accessed 1 Jan 2018]

URL 2. Web Map Tile Service, http://www.kdd.org [Accessed 1 Jan 2018]

URL 3. Web Map Tile Service, https://towardsdatascience.com [Accessed 12 Dec 2017]

URL 4. Web Map Tile Service, http://esdac.jrc.ec.europa.eu [Accessed 30 Jun 2017]

URL 5. Web Map Tile Service, http://eusoils.jrc.ec.europa.eu [Accessed 30 Jun 2017]