

## URBAN EFFECTS ON LAND SURFACE TEMPERATURE IN DAVAO CITY, PHILIPPINES

M. M. Tinoy<sup>1,\*</sup>, A. U. Novero<sup>1,2</sup>, K. P. Landicho<sup>1</sup>, A. B. Baloloy<sup>1</sup>, A. C. Blanco<sup>1</sup>

<sup>1</sup> Geospatial Assessment and Modelling of Urban Heat Islands in Philippine Cities (Project GUHeat),  
Training Center for Applied Geodesy and Photogrammetry, University of the Philippines-Diliman, Quezon City, Philippines 1101  
– (mmtinoy, aunovero, kclandicho1, acblanco)@up.edu.ph, alvinbbaloloy@gmail.com

<sup>2</sup> College of Science and Mathematics, University of the Philippines Mindanao, Davao City, Philippines 8000

### Commission IV

**KEY WORDS:** urban heat island, hot spots, cold spots, multiple regression, quantile regression, modelling

### ABSTRACT:

This study produced spatiotemporal hot and cold spot occurrence maps for Davao City for the period 1994-2019 using land surface temperature (LST) images. Urban heat is theorized to have been affected by some, if not all, of the following impact factors: air pollutant concentrations/particulate matter (PM10), vegetation “abundance” (using EVI), building “density” (NDBI), albedo, topography, and population density. A mobile traverse sampling was performed in the morning and afternoon of 15 April 2019 to measure PM10 in the city’s identified hot spots. The remaining factors were generated from imagery (i.e., Landsat 8, Synthetic Aperture Radar) and obtained from the Philippine Statistics Authority. These factors were analyzed against the LST which was obtained through Project GUHeat’s methodology. The relationships between the factors and LST were studied through multiple and quantile regression models (MRM & QRM). Results showed that variable PM10 does not have any significance in the MRM. Meanwhile, QRM were fitted to different quantile values, namely: 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 90<sup>th</sup>. It is only at the 90<sup>th</sup> quantile where all the independent variables are good predictors for the LST. Albedo is the most important predictor for the LST at 10<sup>th</sup> quantile whereas Elev for the 25<sup>th</sup> quantile. However, when LST is at the 50<sup>th</sup>, 75<sup>th</sup>, and 90<sup>th</sup> quantiles NDBI is the most significant variable at predicting LST. Reliable spatiotemporal assessment and modelling of surface temperature are essential for urban planning and management to formulate sustainable strategies for the welfare of people and environment.

### 1. INTRODUCTION

Urban Heat Island (UHI) is an environmental phenomenon where urban temperature is higher than its surrounding rural areas (Howard, 1818). The urban areas of developing countries, especially those with hot-humid climate like the Philippines, are vulnerable to excess heat. Urban heat is theorized to have been affected by some, if not all, of the following impact factors: air pollutant concentrations/particulate matter (PM10), vegetation “abundance” (quantified using Enhanced Vegetation Index, EVI), building “density” (measured using Normalized Difference Building Index, NDBI), surface albedo, and place-specific geography layers (i.e., topography and population density).

Various studies have identified the prominent role of urbanization in Land Surface Temperature (LST) variability. Due to the availability of advanced technology, the impact of population growth, land use and land cover (LULC) conversion, urban pollution, and other effects of global urbanization that influence LST over time have been investigated by researchers worldwide through different models. Coops et al. (2007) predicted the afternoon LST in Canada’s land cover classes (forest, shrub/grasses, and crops) in linear regression modelling with MODerate-resolution Imaging Spectroradiometer (MODIS)-derived morning LST, location and elevation data as the independent variables. Across all cover types, afternoon and morning LST values were highly correlated and significantly different. Results showed that  $R^2$  were consistently highest for all the cover classes combined while lowest for crops. A stepwise correlation analysis was performed by Xiao et al.

(2008) in Beijing, China to examine the spatial distribution of Landsat-derived LST and its relationships with biophysical and demographic parameters. Results showed that LST was significantly correlated with forest (FO), farmland (FA), water (WA), low-density built-up, high-density built-up, extremely-high buildings, low buildings by grid, and population density while roads, exposed land, and medium-high buildings were not included in the stepwise regression model. The ratios of FO, FA, and WA were found to be the most influential variables in controlling LST variation.

Hart and Sailor (2009) studied the importance of various land-use and surface parameters on the spatial distribution of the UHI across Portland metropolitan area through tree-structured regression models for both weekend and weekday daytime. The dependent variable in the analysis was mean UHI intensity for each grid cell covered by temperature traverse while the independent variables in the models were the surface and land-use characteristics (canopy cover, ground vegetation cover, impervious surface, loose surface, land cover, building floor space, and length of roads). The test of multicollinearity between the independent variables, one of the assumptions of multiple regression analysis, showed no problem based on the tolerance values (greater than 0.2). Results displayed the aerial image of canopy cover as the most important urban characteristic separating warmer from cooler regions of the study site, regardless of day of week.

In India, Pandey et al. (2012) examined MODIS-derived day and night time surface temperature distribution if it has a relationship with particulate matter concentration and LULC by performing

\* Corresponding author

multiple linear regression. Peng et al. (2016) attempted to model multiple factors (Normalized Difference Moisture Index-NDMI, Normalized Difference Vegetation Index, elevation, slope, and aspect) and LST for both sunny and shadow area of Western Sichuan Plateau, China. Prior to the stepwise regression analysis, they have performed principle component analysis because of the presence of multicollinearity among the factors. Their work revealed that NDMI and elevation have influenced LST the most for both sunny and shadow areas. The work of Zhou et al. (2014) employed quantile regression method to have an estimation of effects of land use and geographic predictor variables on temperatures during the heat wave in Greater Houston, Texas, United States in 2011. Results revealed that highly developed area and distance to the coastline have larger impacts on daily mean temperatures at higher quantiles, and open water area has greater effects on daily minimum temperatures at lower quantiles.

While few researches in the Philippines have explored the association of LST and abundance of vegetation and built-up formation through remotely-sensed indices (Estoque et al., 2016; Lu et al., 2017; Pereira and Lopez, 2004; Tiangco et al., 2008), these studies have focused only on the northern area of the country, particularly Metropolitan Manila. This paper aimed to map the hot spots (HS) and cold spots (CS) of LST in Davao City for the period 1994-2019, and then perform an analysis of the relationship between multiple impact factors (MIF) and LST. Firstly, the spatiotemporal mapping of the occurrence of statistically significant clusters of HS and CS was investigated. Secondly, regression models were used in this research to quantify the MIF that have the greatest influence on LST. Mapping and statistical analyses of LST variation using appropriate methodologies provide vital information for the improvement of thermal environment.

## 2. STUDY AREA

Davao City is a highly-urbanized city located in the southeastern part (125°13' to 125°41' E longitude, 6°58' to 7°34' N latitude) of Island of Mindanao. It is the largest city in the Philippines in terms of land area with 2,444 square kilometers. It is the third most populous city in the country after Quezon City and Manila with population of 1,632,991 (Philippine Statistics Authority, 2018). A large part of Davao is mountainous, characterized by extensive mountain ranges with uneven distribution of highlands and lowlands. The city's Southeast is surrounded by Davao Gulf spanning approximately 60-km wide. The predominant wind direction is northward from the gulf. Compared with other parts of the Philippines in which there are distinct hot and wet seasons, Davao City experiences mild tropical climate where the days are always sunny and followed by nights of rain. It is outside the typhoon belt and lacks major seasonal variations. Its annual average temperature ranges from 25.0 to 32.8 °C with the hottest occurring during the month of April, and the coldest occurring on January based on observations in 2018 (PAG-ASA, 2019).

## 3. MATERIALS AND METHODS

### 3.1 Spatiotemporal Mapping

The annual mean LST images for the period 1994-2019 in Davao City were downloaded freely from Climate Engine Application (Climate Engine, 2019) while the official boundary of the city was obtained from the City Planning and Development Office. A 100m-grid polygon and its centroid points with similar ID

spanning the extent of the city were generated in QGIS Desktop version 3.4.5 (QGIS Development Team, 2019). Mean LST values of each year were then extracted into the point shapefile. This point dataset was joined to the polygon dataset with ID as the target field layer. This resulted to a database of mean LST values per grid cell per year. The polygon shapefile was loaded in GeoDa version 1.14.0 (Anselin, 2005). Then, the weighted shapefile served as an input file to the univariate local Moran's I function to produce cluster maps and database of yearly HS and CS with significance level of  $p=0.001$ . After consolidating the yearly database, Shapes to Grid function was performed in SAGA GIS version 2.3.2 (Conrad et al., 2015) and finally the spatiotemporal HS and CS occurrence frequency map was created.

Maps of occurrence of cluster transition were generated. This means that each 100-m grid cell for the period 1994-2019 was analyzed. There were three cluster types used, namely: Cold, Normal, and Hot. A transition was determined from one year to its succeeding year ( $y_i$  to  $y_{i+1}$ ) for each grid and for each type of transition. A positive result means that a count for that grid for that type of transition was recorded. The total number of occurrences were then summed up and divided by the number of years of observation for each grid. As an example, if the temperature was Cold for year 1994 and Normal for year 1995 for a certain grid, then it means that an occurrence from Cold to Normal was observed and a count toward the transition Cold to Normal was recorded. This was continued for 1995-1996, 1996-1997, until 2018-2019 for this grid for this type of transition. The total number of Cold to Normal transitions were then summed up and divided by the total number of years to determine the percentage of the Cold to Normal type of transition for the years in observation. The same methodology was applied for all other types of transitions for every grid. The type of transition refers to the different permutations of transitions between the different types of temperatures. There were nine different transition combinations, namely: Cold to Normal (CN), Normal to Hot (NH), Cold to Hot (CH), Normal to Cold (NC), Hot to Normal (HN), Hot to Cold (HC), Hot to Hot (HH), Normal to Normal (NN), and Cold to Cold (CC). The last three types of transition (HH, NN and CC) simply means that there were no changes in conditions for adjacent years  $y_i$  to  $y_{i+1}$ . The following transitions were described as change from warmer to colder: NC, HN and HC, while the following transitions were described as change from colder to warmer: CN, NH, and CH. The transitions CH and HC were considered drastic transitions having jumped from two different temperature types in adjacent years. The years at which each transition type has occurred over the years 1994-2019 were also recorded for analysis. The total number of grids for which a transition type occurred for a particular year for each grid was then summed up.

### 3.2 Multiple and Quantile Regression Modelling

Data	Source
LST	Project GUHeat's method using Landsat 8 (L8)
NDBI	Computed from downloaded bands of L8
EVI	Computed from downloaded bands of L8
Albedo	Computed from downloaded bands of L8
PM10	Traverse measurement
Pop Dens	Philippine Statistics Authority
Elev	Synthetic Aperture Radar (SAR) imagery

Table 1. Data summary for regression modelling.

Multiple regression and quantile regression analyses were performed to generate an equation (model) that would best describe the statistical relationships between

response/dependent variable-LST, and the predictor/independent variables, namely, Normalized Difference Building Index (NDBI), Enhanced Vegetation Index (EVI), Surface Albedo (Albedo), Particulate Matter 10 (PM10), Population Density (Pop\_Dens), and Elevation (Elev). The summary of these datasets is listed in Table 1.

A mobile traverse sampling was performed in 15 April 2019 on a steady-state condition to measure PM10 in the city's identified HS from the first part of this study. Prior to the field activity, a 60-km route was created to weave in and out of the HS by means of a hired vehicle. A low-cost mobile PM sensor was used, which simultaneously transmits data into an android device through the CrowdSense application (CrowdSense, 2019). To track vehicle movement and determine the location of each reading, a handheld Global Position System device was also utilized. PM10 data, with a measurement unit of  $\mu\text{g}/\text{Nm}^3$ , were collected at two runs, namely: 9:15 to 11:35 in the morning and 2:56 to 5:40 in the afternoon, for the first and second run, respectively (Figure 1). Meanwhile, the values of LST, EVI, NDBI, and Albedo layers were acquired from Landsat 8 bands. This satellite passed through the city in 14 April 2019 at 9:54AM, hence a morning observation only. Moreover, the Elev layer was gathered in 2015 from SAR imagery with a spatial resolution of 10 meters while Pop\_Dens layer was obtained from the population density of each barangay of the city in the year 2015. All values of LST, EVI, NDBI, Albedo, Elev, and Pop\_Dens layers of each surveyed PM10 point were extracted and added as fields into the point shapefile of the mobile traverse sampling. This joined dataset served as an input file to RStudio (RStudio Team, 2019). The statistical packages *cars*, *lmtest*, *alr3*, and *stats* were used to fit and check the aptness of the model for the Multiple Linear

Regression and *quantreg* to fit a model using the Quantile Regression method.

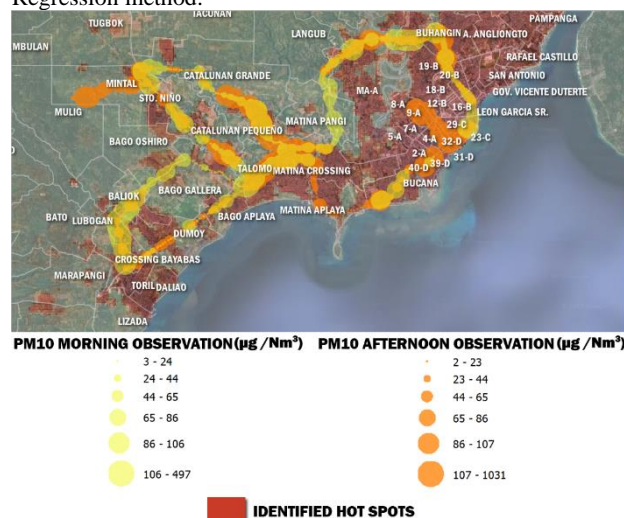


Figure 1. Track of Particulate Matter 10 traverse survey.

As shown in the figure, there are points along the route that have low and high values of PM10. This can be attributed to longer period of stay in those points due to light to moderate traffic and the presence of road intersections. The variation of PM10 readings can also be associated to the volume of vehicles at the time of the data collection since traffic volume differs depending on the time of the day. The emitted air pollutants from fuel exhausts of vehicles and establishments, wind speed and direction, land use or land cover, and various other reasons should also be valuably considered.

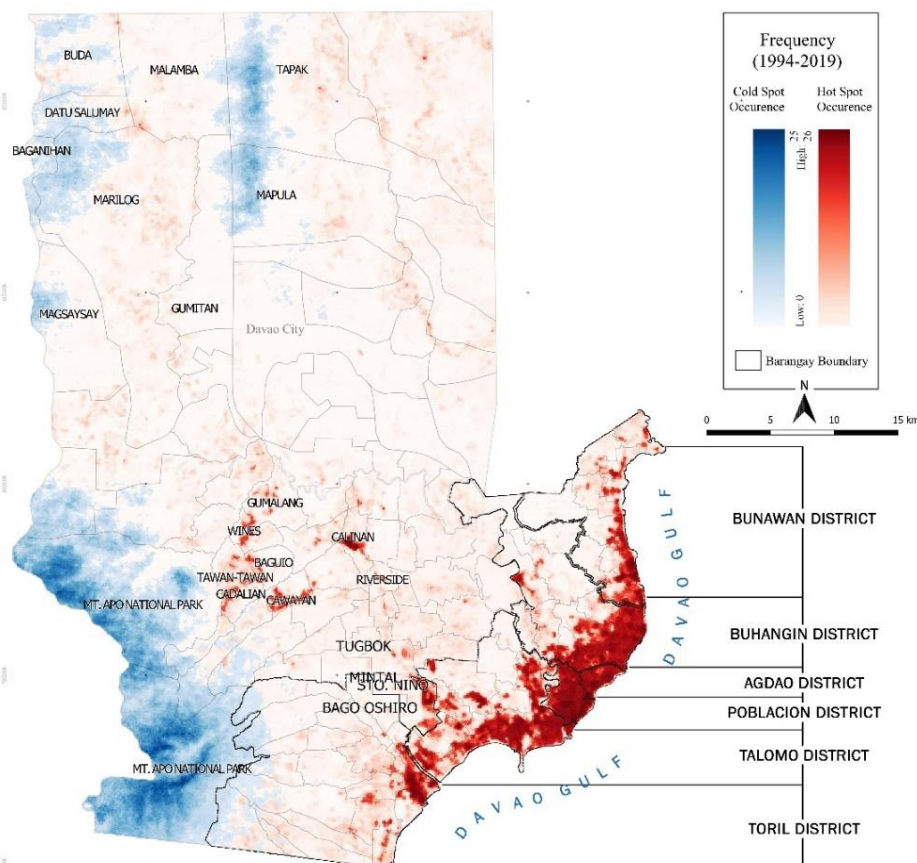


Figure 2. Spatiotemporal hot spots and cold spots frequency map in Davao City for the period 1994-2019.

## 4. RESULTS AND DISCUSSION

### 4.1 Spatiotemporal Mapping

Figure 2 shows the spatiotemporal HS and CS occurrence frequency map in Davao City. The maximum occurrence scores for the HS and CS were 26 and 25, respectively. The recorded minimum occurrence scores for both spots were 0. Mt. Apo Natural Park and its adjacent barangays, situated at Southwest of Davao, were identified as CS. Also with colder temperature are areas in North and Northwest: Buda, Datu Salumay, Baganihan, Marilog, Magsaysay, Malamba, Gumitan, Tapak and Mapula. Moreover, at the Southeast of Davao, HS were dominated by the main urban center (whole coverage of Districts of Poblacion and

Agdao; and few areas of Districts of Talomo and Buhangin) and coastal barangays of Bunawan and Toril Districts. Furthermore, part of neighboring barangays composed of Tugbok, Mintal, Sto. Niño, and Bago Oshiro were marked as HS. Lastly, small portion of Calinan, Cawayan, Wines, Baguio, Cadalian, Tawan-Tawan, Gumalang, and Riverside geographically located in the city center had warmer temperatures. This agrees with the distribution of land cover in these areas which consist of dense low-rise to lightweight low-rise structures, compact open mid-rise buildings, few scattered trees, and heavily paved roads. Meanwhile, shown in Table 2 is the annual count of all 100-m grids of HS and CS as well as the percentage of the clustered grids over all grids spanning the extent of Davao City.

Year	Count of HS	Count of CS	Perc. of HS	Perc. of CS	Year	Count of HS	Count of CS	Perc. of HS	Perc. of CS
1994	25,765	16,523	10.62	6.81	2007	19,133	12,207	7.89	5.03
1995	25,656	8,689	10.58	3.58	2008	27,371	31,196	11.29	12.86
1996	21,749	6,743	8.97	2.78	2009	21,795	30,825	8.99	12.71
1997	28,177	37,531	11.62	15.48	2010	21,888	7,502	9.03	3.09
1998	32,209	0	13.28	0	2011	21,702	15,286	8.95	6.30
1999	33,545	26,069	13.83	10.75	2012	18,511	8,494	7.63	3.50
2000	20,062	11,693	8.27	4.82	2013	17,622	9,907	7.27	4.08
2001	17,312	10,562	7.14	4.36	2014	17,279	8,365	7.12	3.45
2002	19,024	6,479	7.84	2.67	2015	16,710	9,876	6.89	4.07
2003	23,891	19,021	9.85	7.84	2016	21,105	26,648	8.70	10.99
2004	17,119	9,652	7.06	3.98	2017	15,075	4,324	6.22	1.78
2005	17,247	15,693	7.11	6.47	2018	19,544	10,658	8.06	4.39
2006	15,463	7,697	6.38	3.17	2019	23,051	6,586	9.50	2.72

Table 2. Annual count and percentage of grids of HS and CS.

Based on Table 2, the area coverage of HS decreased from 10.62% in 1994 to 9.50% in 2019. Similarly, the percentage of areas classified as CS decreased by 4.09%, from 6.81% in 1994 to 2.72% in 2019. The highest percentage of HS was observed

in 1999 followed by its preceding year, 1998 with 13.83% and 13.28%, respectively. For the CS, 15.48% in 1997 was the highest recorded area coverage. In 1998, there were no observed CS and the explanation for this is unknown.

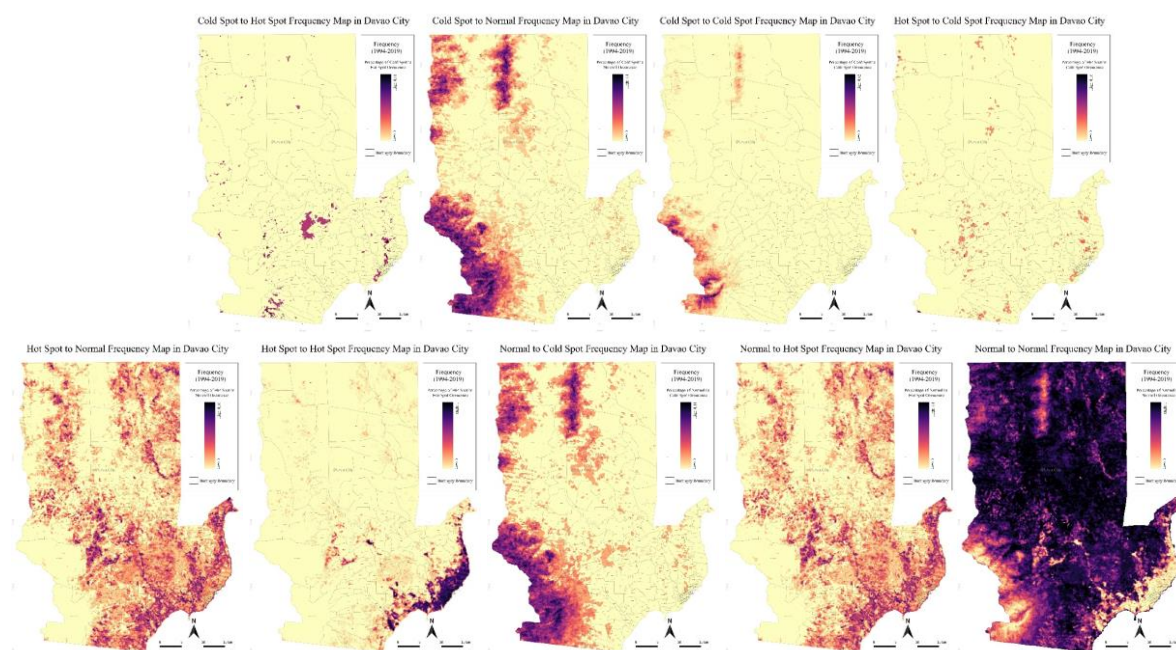


Figure 3. Frequency maps of transitions of cluster types. 1st row L-R: Cold to Hot, Cold to Normal, Cold to Cold, and Hot to Cold; 2nd row L-R: Hot to Normal, Hot to Hot, Normal to Cold, Normal to Hot, and Normal to Normal.



The maps of nine different transition combinations (CH, CN, CC, HC, HN, HH, NC, NH, and NN) are shown in Figure 3. The lowest occurrence was 0% for all transition types while the highest transition occurrence range from 8%-100%. The transition combinations were ranked as follows: CH-8%, HC-16%, HN-36%, CN-40%, NH-40%, NC-44%, CC-92%, HH-100% and NN-100%. The last two transitions implied no changes from being Hot and Normal throughout the period were recorded.

Figure 4 presents the count of grids occurring per year per transition type. It can be noted that there is a big difference on the trend of the NN transition with all the other transition types. NN refers to no transition from temperature type Normal on succeeding years. It is noticeable that the drastic transition types (CH and HC) generated the lowest records for almost all of the years compared to the rest of the one-temperature transition types (CN, NH, NC, HN). The drastic transition types were only observed for the years 1996 to 2010. This means that, for Davao City and years in consideration, it was not likely that there were two temperature transition types in succeeding years.

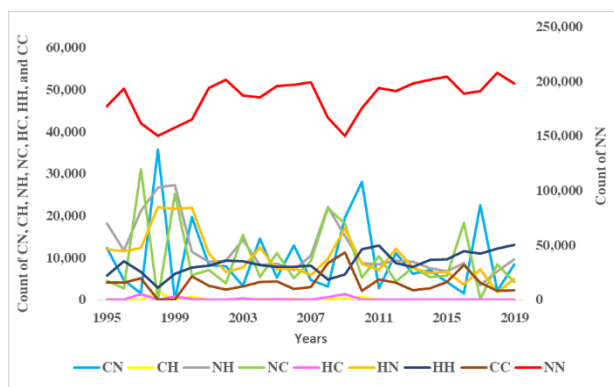


Figure 4. Count of grids for Cold to Normal, Cold to Hot, Normal to Hot, Normal to Cold, Hot to Cold, Hot to Normal, Hot to Hot, and Cold to Cold transition types (at first vertical axis: left) and Normal to Normal transition type (at secondary vertical axis: right) in Davao City from 1994-2019.

In almost all of the years in consideration, there were more than 5,000 grids (2.06%) that transitioned from a warmer to colder temperature (NC and HN) except for HC. This count peaked in 1997 when the number of grids reached around 31,000 (12.78%) for the transition type NC. The transition type HN only reached a maximum of around 22,000 (9.07%) grids recorded between 1998-2000. There was an obvious alternating increase and decrease of number of grids every year for the NC transition type except for the years 2006-2008, 2008-2010 and 2014-2016. This trend was not generally noticeable in the HN and HC transition types. And while further assessing the transition type occurrence in the city, it is alarming to note that the average count of 100-m grids that transitioned from colder to warmer temperature (CN and NH) except for CH is around 11,000 grids in all of the years in consideration. The former and latter peaked at 1998 and 1999 with approximately 36,000 and 27,000 grids, respectively.

## 4.2 Multiple Regression Modelling

A linear model was fitted to the given point dataset. The response variable (dependent), LST, was modelled using PM10, EVI, NDBI, Albedo, Elev, and Pop\_Dens as its predictors (independent variables). As can be observed from Table 3, the

only t-statistic with a large p-value is the variable PM10. This means that PM10 can be dropped as it does not have any significance in the model. The r-square and adjusted r-square of the model are 0.4457 and 0.4438, respectively. Further from the results, the full model for the LST is given below:

$$\text{LST} = 32.9100 - 906.4000\text{EVI} + 6.2350\text{NDBI} - 53.6900\text{Albedo} - 0.0139\text{Elev} - 0.0024\text{Pop\_Dens} \quad (1)$$

The regression regular coefficient estimates embody the mean change in the response variable given a one unit of change in the predictor variable while holding other predictors in the model constant. Interpreting the model is as follows: 0.01 unit increase in EVI decreases LST by 9.064 °C; 0.01 unit increase in NDBI increases LST by 0.06235 °C, and; 0.01 unit increase in Albedo decreases LST by 0.5369 °C. As for the Elev and Pop\_Dens: the regular coefficient estimates indicate that for every additional 1 meter in Elev and additional 1 person per hectare in Pop\_Dens, a decrease in LST is expected by an average of 0.0139 °C and 0.0024 °C, respectively.

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	32.9100	287.827	0.0000 <sup>1</sup>
PM10	0.0039	0.227	0.8203
EVI	-0.0594	-3.195	0.0014
NDBI	0.2914	13.217	0.0000 <sup>1</sup>
Albedo	-0.4299	-23.498	0.0000 <sup>1</sup>
Elev	-0.2640	-12.014	0.0000 <sup>1</sup>
Pop_Dens	-0.1292	-6.519	0.0000 <sup>1</sup>

Table 3. Multiple linear regression results.

To determine predictor variable's importance, the standardized coefficient estimates were generated as shown in Table 3. These represent the mean change in the response variable for one standard deviation change in the predictor variable. Looking at the standardized estimate, Albedo showed the largest absolute value among the impact factors, followed by NDBI, Elev, Pop\_Dens, and EVI. This suggests that Albedo is the most important variable in the MRM.

Model diagnostics were performed to check the aptness of the given model. Variance Inflation Factor (VIF) was used to measure the multicollinearity between the independent variables. A VIF value of 5 or more suggests that a strong multicollinearity exists in the model. Results showed that VIF of PM10, EVI, NDBI, Albedo, Elev, and Pop\_Dens were 1.02, 1.07, 1.51, 1.04, 1.50, and 1.26, respectively. Since there are no VIF values exceeding 5, the model is safe from strong multicollinearity.

Shapiro-Wilk Test was employed to check the normality of the residuals. The test statistic is 0.9297, with a p-value of 0.0000<sup>1</sup>. This implies that the residuals are not normal. Meanwhile, in using the Breusch-Pagan Test to check the assumption of constant variance, the resulting test statistic is 591.94 with a p-value of less than 0.001. Since both tests for normality and constancy/homogeneity have failed, this indicates that the estimate of the parameters in the multiple regression model is biased.

## 4.3 Quantile Regression Modelling

Quantile regression models the relation between a set of predictor variables and specific percentiles (or quantiles) of the

<sup>1</sup> Exact p-values are lower than 0.001

response variable. It specifies changes in the quantiles of the response. Quantile regression overcomes the problem of heterogeneity of variance by fitting linear regressions on different conditional quantiles of the range of a response variable (Cade and Noon, 2003; Koenker and Bassett, 1978). Other advantages of the Quantile Regression over the Ordinary Least Squares Regression is that it is robust to non-normal errors and to outliers.

Models were fitted to different quantile values. These values are 0.1 (10<sup>th</sup>), 0.25 (25<sup>th</sup>), 0.50 (50<sup>th</sup>), 0.75 (75<sup>th</sup>), and 0.90 (90<sup>th</sup>), dividing the LST into 5 levels, namely: very low (18.30 °C – 27.32 °C), low (27.33 °C – 29.13 °C), moderate (29.14 °C – 30.61 °C), high (30.62 °C – 31.86 °C), and very high (31.87 °C – 33.42 °C).

**Models at the 10<sup>th</sup>, 25<sup>th</sup>, and 50<sup>th</sup> Quantiles.** From Table 4-6, the p-values for the PM10 and EVI are greater than 0.05. This implies that at the 10<sup>th</sup>, 25<sup>th</sup>, and 50<sup>th</sup> quantiles of the data, PM10 and EVI are not good predictors for the LST. Furthermore, the p-value for Pop\_Dens may still be considered significant at the 10<sup>th</sup> and 25<sup>th</sup> quantiles, however, caution should be exercised when including this variable in the model when considering the 10<sup>th</sup> and 25<sup>th</sup> quantiles of the data. The models at the 10<sup>th</sup>, 25<sup>th</sup>, and 50<sup>th</sup> quantiles could be written as:

$$LST^{[10]} = 31.3239 + 4.8620NDBI - 91.7007Albedo - 0.0127Elev - 0.0033Pop\_Dens \quad (2)$$

$$LST^{[25]} = 31.6792 + 5.5209NDBI - 25.7099Albedo - 0.0189Elev - 0.0025Pop\_Dens \quad (3)$$

$$LST^{[50]} = 32.1851 + 6.8533NDBI - 9.8706Albedo - 0.0177Elev - 0.0022Pop\_Dens \quad (4)$$

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	31.3239	113.4306	0.0000 <sup>1</sup>
PM10	-0.0214	-0.3171	0.7512
EVI	0.0064	0.2414	0.8092
NDBI	0.2412	5.0924	0.0000 <sup>1</sup>
Albedo	-1.0403	-12.5137	0.0000 <sup>1</sup>
Elev	-0.2848	-4.00004	0.0001
Pop_Dens	-0.1023	-1.8951	0.0583

Table 4. 10<sup>th</sup> quantile regression results.

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	31.6792	375.8977	0.0000 <sup>1</sup>
PM10	0.0098	0.1445	0.8852
EVI	-0.1840	-1.0316	0.3024
NDBI	0.9105	16.4334	0.0000 <sup>1</sup>
Albedo	-0.6383	-6.6337	0.0000 <sup>1</sup>
Elev	-1.4782	-26.1608	0.0001
Pop_Dens	-0.3738	-15.8466	0.0583

Table 5. 25<sup>th</sup> quantile regression results.

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	32.1851	307.4929	0.0000
PM10	0.0810	0.8014	0.4230
EVI	-0.1519	-1.3038	0.1925
NDBI	1.6196	16.2320	0.0000 <sup>1</sup>
Albedo	-0.3040	-4.9394	0.0000 <sup>1</sup>
Elev	-1.5806	-15.9620	0.0001
Pop_Dens	-0.7045	-4.5410	0.0000 <sup>1</sup>

Table 6. 50<sup>th</sup> quantile regression results.

**Models at the 75<sup>th</sup> and 90<sup>th</sup> Quantiles.** Albedo may still be considered significant with a p-value of 0.0556 (Table 7). Meanwhile, all the variables at 90<sup>th</sup> quantile (Table 8) are good predictors for the LST since all were included in the model. The models at the 75<sup>th</sup> and 90<sup>th</sup> quantiles could be written as:

$$LST^{[75]} = 32.6970 + 0.0026PM10 - 881.803EVI + 7.0760NDBI - 4.0043Albedo - 0.0160Elev - 0.0010Pop\_Dens \quad (5)$$

$$LST^{[90]} = 33.0096 + 0.0029PM10 - 1030.40EVI + 6.4796NDBI + 3.9496Albedo - 0.0128Elev - 0.0010Pop\_Dens \quad (6)$$

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	32.6970	409.6543	0.0000 <sup>1</sup>
PM10	0.2630	5.4551	0.0000 <sup>1</sup>
EVI	-0.1733	-3.0307	0.0025
NDBI	1.7037	19.9719	0.0000 <sup>1</sup>
Albedo	-0.1760	-1.9159	0.0556
Elev	-1.3927	-16.6647	0.0001
Pop_Dens	-0.4296	-3.3537	0.0008

Table 7. 75<sup>th</sup> quantile regression results.

Coefficients	Standardized Estimate	t-statistic	p-value
Intercept	33.0096	390.3824	0.0000 <sup>1</sup>
PM10	0.4770	4.5111	0.0000 <sup>1</sup>
EVI	-0.1744	-2.4206	0.0156
NDBI	1.2128	14.1869	0.0000 <sup>1</sup>
Albedo	0.1868	2.1605	0.0309
Elev	-0.9643	-8.1089	0.0001
Pop_Dens	-0.3700	-3.6968	0.0002

Table 8. 90<sup>th</sup> quantile regression results.

Based on the QRM results, NDBI and Elev are statistically significant variables at all five quantiles having consistent p-values of 0.0000<sup>1</sup> and 0.0001, respectively. Thus, these two are consistently good predictors for the surface temperature at 18.30 °C – 33.42 °C. In the same manner, Albedo is a good predictor at all quantiles except that caution should be made when including this variable at the 75<sup>th</sup> quantile. Meanwhile, PM10 and EVI are only good predictors for the LST at 75<sup>th</sup> quantile and 90<sup>th</sup> quantile. It is at the 50<sup>th</sup> quantile where Pop\_Dens started to become a good predictor for the LST and this continued at the 75<sup>th</sup> and 90<sup>th</sup> quantiles. Furthermore, it is only at the 90<sup>th</sup> quantile where all the independent variables are statistically significant. This implies that PM10, EVI, NDBI, Albedo, Elev, and Pop\_Dens are all good predictors for the LST when the temperature is at 31.87 °C – 33.42 °C. Focusing on the 90<sup>th</sup> quantile results, if PM10 increases by 1 µg/Nm<sup>3</sup>, it is projected that the surface temperature will get warmer by 0.0029 °C because coefficient for PM10 is positive. This means that when areas in the city will have high readings of PM10 coming from various sources (e.g. vehicle exhausts), there will also be a significant increase in LST. However, if the vegetation “abundance” (using EVI) grows by 0.01, the temperature will reduce by 10.304 °C because EVI is a negative coefficient; conversely, if values of building “density” (using NDBI) and surface albedo rise by 0.01, then LST will increase by 0.064796 °C and 0.039496 °C, respectively since both are positive coefficients. On the other hand, if Elev increases by 1 meter and Pop\_Dens increases by 1 person per hectare, it is estimated that LST becomes less warm by 0.0128 °C and 0.0010 °C, respectively since both are negative coefficients.

Studying the trend of various regular coefficient estimates at different quantiles, it was observed that at median quantile (50<sup>th</sup>)

and higher quantiles (75<sup>th</sup> and 90<sup>th</sup>) PM10 values increased, whereas of EVI values increased throughout the five quantiles. For NDBI coefficient, its values increased as well but only from first up to the fourth quantile levels. In the case of Albedo, values decreased while the LST quantile levels increased. Elev values decreased from 25<sup>th</sup> to 90<sup>th</sup> quantiles. Similarly, Pop\_Dens values decreased but from 10<sup>th</sup> to 75<sup>th</sup> quantiles and the same value was noticed at the 75<sup>th</sup> and 90<sup>th</sup> quantiles. Likewise, the signs of each regular coefficient were checked. PM10 coefficient was only negative at the 10<sup>th</sup> quantile. In contrast, EVI was a positive coefficient at the 10<sup>th</sup> quantile and became negative at the succeeding quantiles. NDBI was positive at all quantiles. Albedo turned positive at the 90<sup>th</sup> quantile from being negative at the preceding four quantiles. Finally, Elev and Pop\_Dens were both negative coefficients at all quantiles.

It should be noted that the regular regression estimates and p-values cannot determine as to which variable/s is/are the most important predictor/s or have influenced LST the most. As discussed in MRM, it is the standardized regression estimates that determine the variable importance wherein larger absolute estimates values represent better or best predictor variables. In summary, Albedo is the most important predictor for the LST at 10<sup>th</sup> quantile (very low: 18.30 °C – 27.32 °C) followed by Elev. Then at the 25<sup>th</sup> quantile (low: 27.33 °C – 29.13 °C) is the Elev variable whereas NDBI is the next most important. However, when LST is at the 50<sup>th</sup> quantile (moderate: 29.14 °C – 30.61 °C), 75<sup>th</sup> quantile (high: 30.62 °C – 31.86 °C) and 90<sup>th</sup> quantile, NDBI showed the largest standardized coefficient estimate among other predictors with 1.6196, 1.7037, and 1.2128, respectively. Thus, NDBI is the most significant variable at predicting LST at median and higher quantiles. It is also interesting to note that as the LST quantile increases the standardized estimate value of NDBI also increases.

**Residuals Vs Fitted.** The graph of residuals vs fitted values could tell if the model violates the assumptions such as constant variance, independent error terms, and linearity of the model in a regression. Erratic or not having any patterns in the graph is a good indication that no major violations in the assumptions mentioned above are violated.

Figure 5 displays the plot between the residuals and the fitted values. Most of the residuals were noted to fluctuate around zero when the values of x is around 28-34. However, for small and large values of x (lower than around 28 and greater than 34), the residuals are largely negative. A systematic pattern may exist in the graph. It is suggested to use formal statistical tests to check whether or not violations in the model exists.

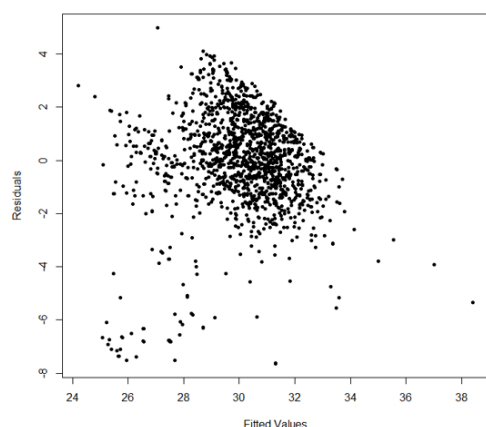


Figure 5. Residuals vs Fitted Values.

**Normal Quantile-Quantile Plot.** It can be observed from Figure 6 that serious deviations from the diagonal line is obvious. This indicates that the residuals are assumed to be not normal and that the assumption of normality is violated. This result is also evident when Shapiro-Wilk Test was employed to test formally the normality of the residuals.

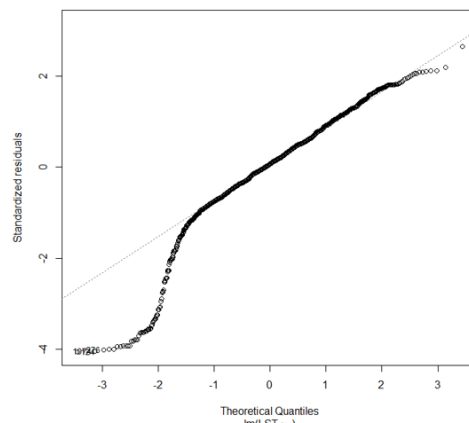


Figure 6. Normal Quantile-Quantile Plot.

## 5. CONCLUSIONS AND RECOMMENDATIONS

This study investigated the spatiotemporal distribution of LST and its quantitative relationships with atmospheric, demographic, and remote sensing-derived parameters in Davao City. The results revealed that occupants of the main urban center (Districts of Poblacion, Talomo, Agdao, and Buhangin) and residents from coastal barangays of Bunawan and Toril Districts experienced the warmest temperatures in Davao City for the period 1994-2019. Multiple regression analysis showed that the air pollutant variable (PM10) exhibited no significance, hence was dropped in the model. Albedo has greater impact in the MRM followed by NDBI, Elev, Pop\_Dens, and EVI. Quantile regression analysis further revealed that at the 90<sup>th</sup> quantile (31.87 °C – 33.42 °C) where all impact factors are good predictors, the presence of buildings in the city (NDBI) showed the biggest impact for the LST. The result of these two regression methods indicated that the models are reliable enough to express LST of Davao City, but also suggested that other impact factors are more significant in order to effectively understand the variations of LST. This study recommends that higher spatial and temporal resolution of satellite imageries should be employed to continuously improve the HS and CS analyses. Moreover, further research should be done by replacing other impact factors in the model and try to include surface moisture, distance to water, solar insolation, and sky view factor. Lastly, to evaluate and compare the urban expansion of other developed areas in Mindanao such as Digos City, Panabo City, and Tagum City performing the same procedure of this study is suggested.

## ACKNOWLEDGEMENTS

Authors would like to acknowledge the Department of Science and Technology (DOST) – Philippine Council for Industry, Energy, and Emerging Technology Research and Development (PCIEERD), Republic of the Philippines for funding and monitoring the Project Number 4028, a nationwide research entitled "Geospatial Assessment and Modelling of Urban Heat Islands in Philippine Cities (Project GUHeat)". The technical support in project implementation was provided by the Training Center for Applied Geodesy and Photogrammetry (TCAGP) of

the University of the Philippines Diliman, Quezon City. The work was performed at the Phil-LiDAR2 Computing Facility, College of Science and Mathematics, University of the Philippines Mindanao, Davao City.

## REFERENCES

- Anselin, L., 2005. *Exploring spatial data with Geoda: A workbook*. Spatial Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign, Illinois.
- Cade, B. S., Noon, B. R., 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, 1(8), 412-420. [https://doi.org/10.1890/1540-9295\(2003\)001\[0412:AGITQR\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2)
- Climate Engine, 2019. Climate Engine: cloud computing and visualization of climate and remote sensing data. Desert Research Institute and University of Idaho. [climateengine.org](http://climateengine.org) (27 August 2019)
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Boehner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.*, 8(7), 1991-2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Coops, N. C., Duro, D. C., Wulder, M. A., Han, T., 2007. Estimating afternoon MODIS land surface temperatures (LST) based on morning MODIS overpass, location and elevation information. *International Journal of Remote Sensing*, 28(10), 2391-2396. <https://doi.org/10.1080/01431160701294653>
- CrowdSense, 2019. SensingKit: A multi-platform mobile sensing framework. [sensingkit.org/projects](http://sensingkit.org/projects) (27 August 2019)
- Estoque, R.C., Murayama, Y., Myint, S.W., 2016. Effects of landscape composition and pattern on land surface temperature: An urban heat island study in the megacities of Southeast Asia. *Science of the Total Environment*, 577, 349-359. <https://doi.org/10.1016/j.scitotenv.2016.10.195>
- Hart, M. A., Sailor, D. J., 2009. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theoretical and Applied Climatology*, 95(3-4), 397-406. <https://doi.org/10.1007/s00704-008-0017-5>
- Howard, L., 1818. The Climate of London: Deduced from meteorological observations, made at different places in the neighbourhood of the metropolis. Volume 1. W. Phillips, George Yard, Lombard Street, London.
- Koenker, R., Bassett Jr, G., 1978. Regression quantiles. *Econometrica*, 46(1), 33-50. <https://doi.org/10.2307/1913643>
- Lu, Y., Coops, N. C., Hermosilla, T., 2017. Chronicling urbanization and vegetation changes using annual gap free landsat composites from 1984 to 2012. *2017 Joint Urban Remote Sensing Event (JURSE)*, 1-4. <https://doi.org/10.1109/JURSE.2017.7924610>
- PAG-ASA, 2019. Philippine Atmospheric, Geophysical and Astronomical Services Administration. Philippine Freedom of Information Portal. [foi.gov.ph/requests?agency=PAGASA](http://foi.gov.ph/requests?agency=PAGASA) (18 February 2019).
- Pandey, P., Kumar, D., Prakash, A., Masih, J., Singh, M., Kumar, S., Kumar, K., 2012. A study of urban heat island and its association with particulate matter during winter months over Delhi. *Science of the Total Environment*, 414, 494-507. <https://doi.org/10.1016/j.scitotenv.2011.10.043>
- Peng, W., Zhou, J., Wen, L., Xue, S., Dong, L., 2016. Land surface temperature and its impact factors in Western Sichuan Plateau, China. *Geocarto International*, 32(8), 919-934. <https://doi.org/10.1080/10106049.2016.1188167>
- Pereira, R. A., Lopez, E. D., 2004. Characterizing the spatial pattern changes of urban heat islands in Metro Manila using remote sensing techniques. *Philippine Engineering Journal*, 25(1), 15-34.
- Philippine Statistics Authority, 2018. 2018 Philippine Statistical Yearbook. Philippine Statistics Authority, CVEA Bldg., East Avenue, Quezon City 1101, Philippines.
- QGIS Development Team, 2019. Quantum Geographic Information System (QGIS) Software. Open Source Geospatial Foundation Project. [qgis.osgeo.org](http://qgis.osgeo.org) (27 August 2019).
- RStudio Team, 2019. RStudio: Integrated development for R. RStudio, Inc., Boston, Massachusetts. [rstudio.com](http://rstudio.com) (27 August 2019)
- Tiangco, M., Lagmay, A., Argete, J., 2008. ASTER-based study of the night-time urban heat island effect in Metro Manila. *International Journal of Remote Sensing*, 29(10), 2799-2818. <https://doi.org/10.1080/01431160701408360>
- Xiao, R., Weng, Q., Ouyang, Z., Li, W., Schienke, E. W., Zhang, Z., 2008. Land surface temperature variation and major factors in Beijing, China. *Photogrammetric Engineering & Remote Sensing*, 74(4), 451-461. <https://doi.org/10.14358/PERS.74.4.451>
- Zhou, W., Ji, S., Chen, T. H., Hou, Y., Zhang, K., 2014. The 2011 heat wave in Greater Houston: Effects of land use on temperature. *Environmental Research*, 135, 81-87. <https://doi.org/10.1016/j.envres.2014.08.025>