

AN INVESTIGATION OF LOCAL EFFECTS ON SURFACE WARMING WITH GEOGRAPHICALLY WEIGHTED REGRESSION (GWR)

Y. Xue ^{a,b,*}, T.Fung^b, J. Tsou^c

^a Satellite Surveying and Mapping Application Center, National Administration of Surveying, Mapping and Geoinformation, P.R.China- xueyc@sasmac.cn

^bDepartment of Geography and Resource Management, the Chinese University of Hong Kong, Shatin, Hongkong SAR, P.R.China - tungfung@cuhk.edu.hk

^c School of Architecture, the Chinese University of Hong Kong, Shatin, Hongkong SAR, P.R.China - jinyeutsou@cuhk.edu.hk

KEY WORDS: Urban Warming, Thermal Landscape, Geographically Weighted Regression (GWR), Thermal Infrared Remote Sensing, ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer)

ABSTRACT:

Urban warming is sensitive to the nature (thermal properties, including albedo, water content, heat capacity and thermal conductivity) and the placement (surface geometry or urban topography) of urban surface. In the literature the spatial dependence and heterogeneity of urban thermal landscape is widely observed based on thermal infrared remote sensing within the urban environment. Urban surface warming is conceived as a big contribution to urban warming, the study of urban surface warming possesses significant meaning for probing into the problem of urban warming. The urban thermal landscape study takes advantage of the continuous surface derived from thermal infrared remote sensing at the landscape scale, the detailed variation of local surface temperature can be measured and analyzed through the systematic investigation. At the same time urban environmental factors can be quantified with remote sensing and GIS techniques. This enables a systematic investigation of urban thermal landscape with a link to be established between local environmental setting and surface temperature variation. The goal of this research is utilizing Geographically Weighted Regression (GWR) to analyze the spatial relationship between urban form and surface temperature variation in order to clarify the local effects on surface warming, moreover to reveal the possible dynamics in the local influences of environmental indicators on the variation of local surface temperature across space and time. In this research, GWR analysis proved that the spatial variation in relationships between environmental setting and surface temperature was significant with Monte Carlo significance test and distinctive in day-night change. Comparatively, GWR facilitated the site specific investigation based on local statistical technique. The inference based on GWR model provided enriched information regarding the spatial variation of local environment effect on surface temperature variation which global model cannot approach.

1. INTRODUCTION

During the process of worldwide urbanization along with high rise and high density housing development in large cities, urban warming has received growing concern among the environmental issues corresponding with urban landscape change. Due to the complicate interplay between urban environment and local climate, it is far from being certain about the urban effects on local warming, few attention has been paid to the variations of local effects on urban thermal performance at a scale over whole city area during urban development, even endeavors is quite scarce to characterize and analyze the spatial heterogeneity and dependency of urban thermal landscape aiming at delimitation of local effects on surface warming at city scale combining the measurement of urban surface and geometry via a viewpoint of local variation in space and time.

The urban thermal landscape study takes advantage of the continuous surface derived from thermal infrared remote sensing at the landscape scale, the detailed variation of local surface temperature can be measured and analyzed through

the systematic investigation. At the same time urban environmental factors can be quantified with remote sensing and GIS techniques. This enables a systematic investigation of urban thermal landscape with a link to be established between local environmental setting and surface temperature variation. Then the mechanism which made the heterogeneity and complexity of urban surface thermal landscape can be studied in depth based on local statistical technique. Geographically Weighted Regression (GWR) is part of a growing trend in GIS towards local spatial analysis which intends to understand the spatial data in more detail through local statistics evaluation. In the literature a number of recent publications have demonstrated the analytical utility of GWR for various urban studies (Huang and Leung, 2002), while it has been scarcely utilized in urban climatic studies especially urban thermal landscape analysis. The interpretation of the local mechanism that affects and determines urban surface thermal abnormalities with the use of GWR presents significant potential for the in-depth investigation of local effects on urban surface temperature variation. The resulting outcome of GWR may help to examine the existence of potential correlation pattern in spatial inconstancy within

* Corresponding author.

whole study area in which way the selected environmental variables can have affected local surface temperature at different relational levels. With this endeavor it hopes to devote to a better understanding of urban effect on local surface thermal environment by taking GWR into the analysis process.

Due to the complexity of urban thermal environment induced by the intensive variation in spatial and temporal dimension, a systematic investigation of the relationship between urban form and urban thermal landscape in a subtropical city, Hong Kong based on statistical regression analysis can assist in a holistic understanding of local surface warming. The goal of this research is utilizing Geographically Weighted Regression (GWR) to analyze the relationship between urban form and surface temperature variation in order to clarify the local effects on surface warming, moreover to reveal the possible dynamics in the local influences of environmental indicators on the variation of local surface temperature across space and time.

2. METHODOLOGY

GWR takes the advantage of the flexibility of local statistical technique to analyze spatially varying relationships. Among the attempts of taking geographical location into its analysis of relationships between variables, Geographically Weighted Regression (GWR) is quite useful which allows complex spatial variations in parameters to be identified, mapped and modeled (Huang, 2000; Brunson et al., 1998). GWR provides the utility to scrutinize the potential discrepancy of local relationship patterns within each parameter through spatially located parameter coefficient estimation of each parameter across the whole study area, which was impossible using the Ordinary Least Squares (OLS) approach. The inference based on the varying relationships provides more in-depth examination of local effects. In this research context, the localized exploratory data analysis of GWR enables the local patterns of surface warming to be revealed and investigated in-depth. The application of GWR in the studies of local effects on urban surface thermal anomalies provides a potential avenue for a comprehensive understanding of local surface warming process.

In the literature, urban thermal environment, in particular UHIs is the integrated output of urban geography, regional climate and ecology. From the urban landscape point of view, “each component surface in urban landscapes (e.g., lawn, parking lot, road, building, cemetery, and garden) exhibits a unique radiative, thermal, moisture, and aerodynamic properties, and relates to their surrounding site environment” (Oke,1982). The surface composition and configuration of urban fabrics which is fragmented and intensively variable in the spatial distribution made the heterogeneity and complexity of urban surface thermal landscape in spatial-temporal dimension. From this perspective landscape composition and configuration are hypothesized to influence urban surface thermal landscape represented by urban surface temperature distribution.

The framework of the study is illustrated in Figure 1. Besides urban surface temperature map derived from remote sensing data, a host of biophysical indicators listed below can be generated from satellite images as well. In conjunction with geographic information systems (GIS) geospatial analysis, other measures of urban form, such as location and elevation,

site openness to the sky, etc., which is listed in the box below, can be calculated to quantify site specific physical and spatial characteristics. All these measures would be employed in statistical models to analyze the interaction between urban form and urban surface temperature variation together with climate ancillary observation.

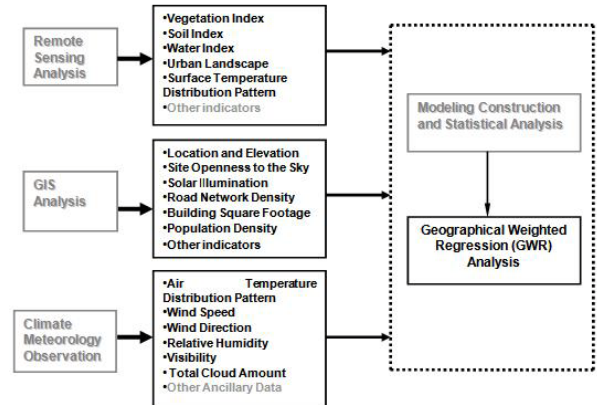


Figure 1. Conceptual design of thermal landscape correlation analysis

The study of local effect of urban environment on surface thermal landscape process is undertaken by statistical regression analysis. The first step for the statistical analysis is the determination of variables. In this analysis urban surface temperature would be the dependent variable which can be derived from surface temperature image of ASTER. Table 1 cites all variables employed in the regression models. Under the research assumptions aforementioned, the important factors that have correlation with the urban surface temperature are listed.

No.	Label	Description	Note
1	ST*****	surface temperature on the day/month/year*****	dependent variable
1	SD*****	road network density in year*****	independent variables
2	TPU****	population density in year*****	
3	DisttoCoast	distance to water body	
4	hshad**	solar radiation at one time of day/month/year**	
5	footsquare-a	building foot square measurement with area	
6	elevationavg	elevation	
7	NDVI*****	vegetation NDVI at one time of day/month/year***	
8	diffuR***	diffuse radiation at one time of day/month/year***	

Table 1. Variables list for regression analysis

3. CASE STUDY

The study area mainly covers the overlapped area of all of the remote sensing images collected for this research in Hong Kong illustrated in Figure 2, with all the weather stations located within the study area where the meteorological data is recorded. The related meteorological data of these stations was purchased from Hong Kong Observatory (HKO) for this study to provide simultaneous field measurements. Within the study area, the elevation ranges from 0 rising to 957 meter of Tai Mo Shan in the New Territories, where the intensive elevation variation occurs along with the mountainous region extension. Most of the extensive urban development with high density housing located sparsely in Tuen Mun, Tsuen Wan, Yuen Long, Tin Shui Wai, Sha Tin, Tai Po, Sheung Shui, Fanling, and Kowloon, mainly focusing on New Territories and part of Kowloon region, which are detached by the mountains covering the whole study area. The northwest corner of the study area contains Hong Kong wetland park and Mai Po Ramsar Site, with fish ponds surrounded. Besides housing estates, the study area includes

a few industrial development areas located in Tuen Mun, Tsuen Wan, Tai Po and Yuen Long. Together with the mountains topography, the diverse land use in the study area provides practical meaning for the study of local effect on surface thermal anomalies under the subtropical climate.

The study period ranges from 2003 to 2006 when the ASTER data used for this site study is available. The ASTER images of these years are obtained from the achieved database in the Land Process Distributed Active Archive Center through the Earth Observing System Data Gateway (EDG), which would be demonstrated and discussed in the subsequent part of this section. Since Hong Kong is a highly developed city with mature urban infrastructure. During this study period from 2003 to 2006, Hong Kong experienced no pervasive urban development but with a few percentages of urban sprawl recognized.



Figure 2. Study area with weather stations

In this research ASTER L1B data and Land Surface Temperature (LST) product AST_08 data were both acquired for this research. ASTER L1B data is used to derive some of the ecological parameters. The surface temperature of AST_08 can be regarded as valid for cloud-free pixels, the reported accuracy of surface temperature estimation achieved proved to be within 1.5K. To facilitate further analysis, the surface temperature of AST_08 data in units of degrees Kelvin has been converted to degrees Celsius (°C). The primary descriptive statistics of surface temperature during these observation periods are given in the Table 2 to provide some background information about the surface temperature distribution within the observation periods.

DATE	TIME OF DAY	MIN	MAX	MEAN	STD
2006-04-17	11:03:00.75	18.15	44.75	26.74	2.99
2005-10-23	11:02:39.38	17.15	41.85	26.60	3.18
2005-10-01	22:34:53.95	22.15	33.85	28.27	1.50
2004-11-21	11:02:10.94	14.95	42.85	25.57	3.34
2004-10-05	22:41:13.71	11.95	29.25	19.95	2.76
2003-11-03	11:03:35.52	22.25	49.45	32.35	3.46
2003-10-28	22:35:49.75	14.85	27.25	21.03	1.82

Table 2. Descriptive statistics of surface temperature for the images used in this research

For the analysis of the relationship between local surface temperature and urban environmental measures by linear regression analysis, there were 1070 sample points with systematic sampling covering the whole study area chosen as the observations for correlation analysis which was

strategically located within the whole study area. All the values of each correlation variables corresponding to each sample point have been calculated.

4. RESULTS AND DISCUSSIONS

Geographical Weighted Regression (GWR) analysis is chosen to measure the possible spatial varying relationship between the referred local environmental factors and surface temperature with the use of a software package GWR 3.0 (Fotheringham et al., 2002). In this calculation, the significance of potential spatial variation in relationship is measured and verified with a Monte Carlo statistical test. P-value for each variable is given and compared to offer a formal evaluation about the significance of spatial variation within each influential factor coefficient.

Because the adaptive kernel is more capable to reflect the scale of local variation resulting from the equal amount of sample data at each regression point as compared to the fixed kernel (Su et al., 2005), the adaptive kernel with AIC (Akaike Information Criterion) minimization bandwidth selection is adopted for GWR analysis in this study. On the other hand it would provide one valued reference indicating the grain size of local varying patterns existed and the further inference of local fragmentation in the relationship pattern within whole study area can be made.

The Monte Carlo significance test of spatial variation corresponding to each variable is summarized in Table 3, with the grey color rows showing the nighttime models, others corresponding with the daytime models. The goodness-of-fit of GWR models within each image time period can be evaluated with the adjusted R-square in this table. In general, the R-square has greatly improved to over 70%. The highest value of adjusted R-square (0.772) appears in daytime model of 11-21-2004. On the other hand, the R-square value of daytime models is consistently higher than the values of nighttime models, which implies that the daytime surface temperature can be better modeled with current environmental variables than nighttime surface temperature.

In summary the F values of GWR models at each time show that the relationships between urban environmental factors and local surface temperature are significant. When comparing the calculated bandwidth of each GWR model between daytime and nighttime, it can be found that the bandwidths of nighttime models is larger than the ones of daytime models, this also confirmed the relative stable surface heating pattern during nighttime than during daytime. It is noted that the bandwidth of GWR model during 10/01/2005 is the biggest (184) among the daytime and nighttime models. This may be due to the strong “smoothing” effect of even ‘noise’ introduced by the stripped LST image and the highest relative humidity (89.1%) at this image time. The standard residual distribution of the models confirmed that spatial association has been successfully removed with GWR model,

During daytime the Intercept and Building square footage demonstrate spatial association correlation with local surface temperature. While during nighttime these spatial variations of parameter coefficients tend to be less obvious. Except these factors referred above, no significant spatial association

was found and detected for Constant in night models, Solar Radiation and Road Density parameters within most of the models of this study during the GWR analysis. The spatial variations of the refereed parameters proved the mixed correlation between such evaluation variables and local

surface temperature across the whole study area, which is not uniformly stationary within the whole study area and may change from negative correlation to positive correlation under local circumstance.

P-value (Monte Carlo)	Constant	Vegetation NDVI	Solar Radiation	Building Square Footage	Elevation	Road Density	Distance from Coast	Openness to Sky	Population Density	Kernel Bandwidth	R ²	Adjusted R ²	Effective number of parameters	F
ST04172006= (significance) ***	0.00 ***	0.00 ***	0.12	0.00 ***	0.00 ***	0.13	0.00 ***	0.00 ***	0.00 ***	114	0.778	0.737	165.6	6.11
ST10232005= (significance) ***	0.00 ***	0.01 **	0.07	0.00 ***	0.00 ***	0.75	0.00 ***	0.00 ***	0.00 ***	162	0.778	0.730	120.4	5.89
ST11212004= (significance) ***	0.00 ***	0.00 ***	0.00 ***	0.00 ***	0.00 ***	0.42	0.00 ***	0.00 ***	0.00 ***	99	0.812	0.772	186.5	5.10
ST11032003= (significance) ***	0.00 ***	0.00 ***	0.00 ***	0.03 *	0.00 ***	0.01	0.00 ***	0.00 ***	0.00 ***	114	0.785	0.746	163.7	8.05
ST10012005= (significance) ***	0.12	0.00 ***	N/A	0.21 ***	0.00 ***	0.24	0.00 ***	0.03 *	0.00 ***	184	0.76	0.738	93.1	6.06
ST10052004= (significance) ***	0.06 ***	0.00 ***	N/A	0.18 ***	0.00 ***	0.89	0.00 ***	0.00 ***	0.00 ***	162	0.735	0.707	104.0	5.97
ST10282003= (significance) ***	0.00 ***	0.00 ***	N/A	0.28 ***	0.00 ***	0.58	0.00 ***	0.00 ***	0.00 ***	173	0.731	0.704	97.6	7.21

*** = significant at .1% level ** = significant at 1% level * = significant at 5% level

Table 3. Geographically weighted regression diagnosis

Since each environmental measurement has various range of value, moreover the NDVI values range from negative to positive, this made the direct comparison of parameter coefficient difficult in order to measure the importance of each parameter impact on local surface temperature variation, including the direction (the possible cooling or warming effect) and extent (the altitude of local surface temperature change induced). For easy comparison between the influences of environmental factors on local surface temperature variation across space and time, the indicating impacts of each environmental parameter can be quantified with the statistical variation of surface temperature induced by each factor, i.e. here called component contribution using the following formula:

$$\text{Component Contribution (CC)} = \text{factor value} * \text{parameter coefficient} \quad (1)$$

Then all the component contribution to surface temperature variation corresponding to each factor have been calculated in order to measure the influence of each environmental factor on local surface temperature change within each image time, with positive CC value indicating the possible warming effect by increasing local surface temperature with corresponding statistical CC value ° C, and negative CC value indicating the possible cooling effect by decreasing local surface temperature with CC value ° C.

In order to demonstrate the local variation pattern, the illustration would emphasize on parameters whose coefficients distribution show significant spatial variation along with the seasonal and day-night change. To this end, the CC values related to parameters, such as Elevation, Distance from coast, in daytime models during 04/17/2006 and 11/03/2003 together with nighttime model during 10/28/2003 were chosen as examples and mapped against the corresponding VNIR image to facilitate further interpreting the spatial varying impact of each factor on surface temperature variation which is shown in Figures 3~8.

Both the nighttime and daytime models during 10/28/2003 and 11/03/2003, 04/17/2006 demonstrate significant spatially varying relationships between Elevation and local surface

temperature with diverse parameter coefficients of Elevation within whole study area, CCs of Elevation at each image time are mapped in Figures 3, 4 and 5. Figure 3 shows that the nighttime model in 10/28/2003 demonstrates consistent negative correlation between Elevation and local surface temperature which implies a stable cooling effect of Elevation on local warming during nighttime period. The cooling effect is becoming stronger with lower negative CC values along with the altitude rising, with the lowest located in the peak of the hill indicating the strongest cooling effect in high altitude area. On the other hand the daytime models with seasonal change during 11/03/2003 and 04/17/2006 demonstrate a more diverse correlation ranging from positive to negative within whole study area in Figures 4 and 5, this may be induced by the various heating situation created with solar radiation during daytime under intensive elevation variation of urban canopy. Most of the areas with positive CC values indicating a heating effect on local surface temperature located in relatively low elevation area, under these low elevations area increasing the elevation may increase the opportunity to receive more solar radiation and then heighten the local surface temperature. At the same time it is interesting to find that the area coverage and value of positive CC is larger (with the biggest value 2.78) during dry season 11/03/2003 than during the growing season 04/17/2006 (with the biggest value 2.25). This may be due to the fact that the solar elevation angle during 04/17/2006 is larger (67.33 deg) compared with 11/03/2003 (49.39 deg), which helps decrease the area coverage and influence of solar shading.

With the comparison of the CC values distribution of distance from coast for nighttime 10/28/2003, and daytime 11/03/2003 and 04/17/2006 shown in Figures 6, 7 and 8, it is easy to find that the CC values ranges from negative to positive when being far away from coast. Within a certain distance to the coast which may vary with local wind speed and direction and relative humidity, proximity to coast shows a significant negative correlation with the local surface temperature which indicates the obvious cooling effect on local surface within this extent. This confirmed that GWR model revealed the cooling effect of proximity with coast, when being further far away from coast the cooling effect vanished, instead of warming effect emerging. Comparing

the pattern of CC variation in Figures 7 and 8 in daytime model 11/03/2003 and 04/17/2006, when approaching coast (the blue sample points area), the daytime model 11/03/2003 tends to show a more obvious negative correlation by a lower negative CC value with wider covering than the model 04/17/2006. This may be due to the relatively higher wind speed (5.67 m/s) at image time 11/03/2003 than at image time 04/17/2006 (2.81m/s) which favored the cooling effect of coast. Moreover with the wind direction change from blowing from coast to land (at image time of 11/03/2003) to blowing from land to coast (at image time of 04/17/2006) in yellow circle area, the CC values appeared from negative to positive indicating that the component contribution (CC) of proximity with coast to local surface temperature variation from temperature decrease to increase (referred as from cooling effect to warming effect) within this area. This may be due to the fact that when the wind blows from land to coast, the coast in low surface temperature during daytime may have no obvious cooling effect on local surface, even slightly contributes to the heightening of the local surface temperature with low positive CC values.

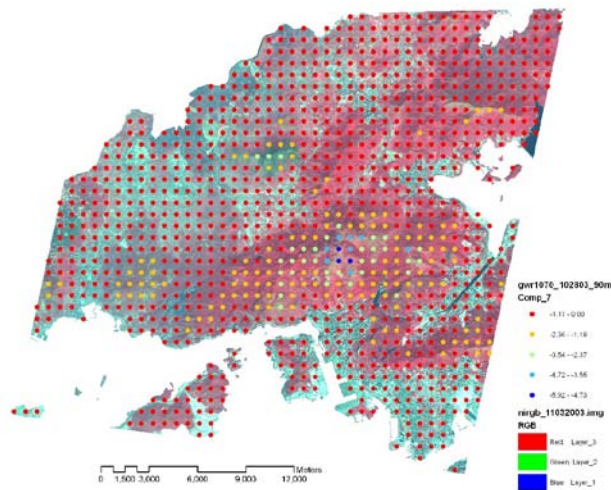


Figure 3. Coefficients contribution of Elevation in nighttime model 10/28/03

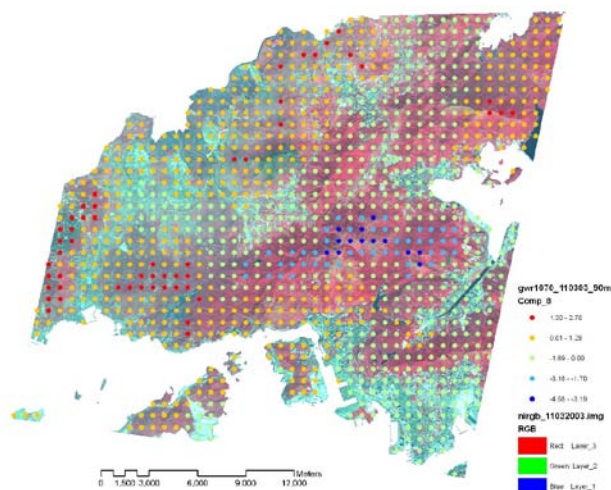


Figure 4. Coefficients contribution of Elevation in daytime model 11/03/03

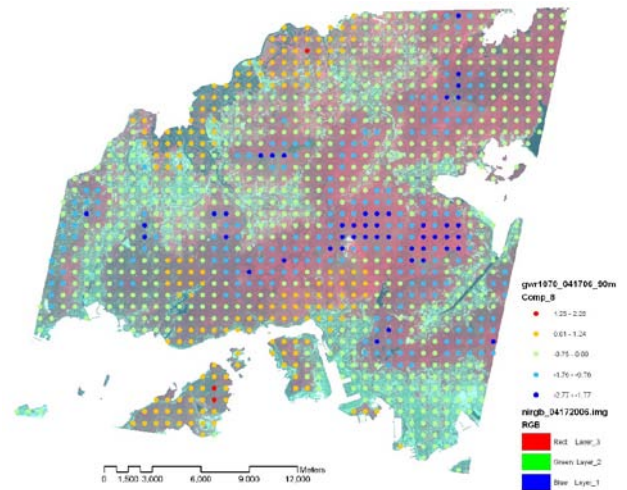


Figure 5. Coefficients contribution of Elevation in daytime model 04/17/06

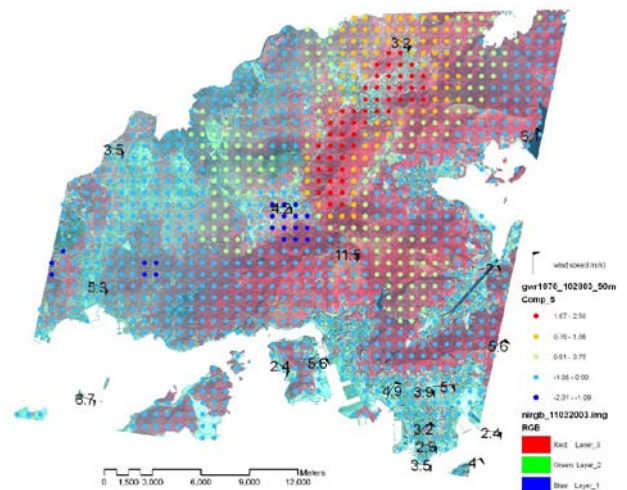


Figure 6. Coefficients contribution of Distance from coast in nighttime model 10/28/03

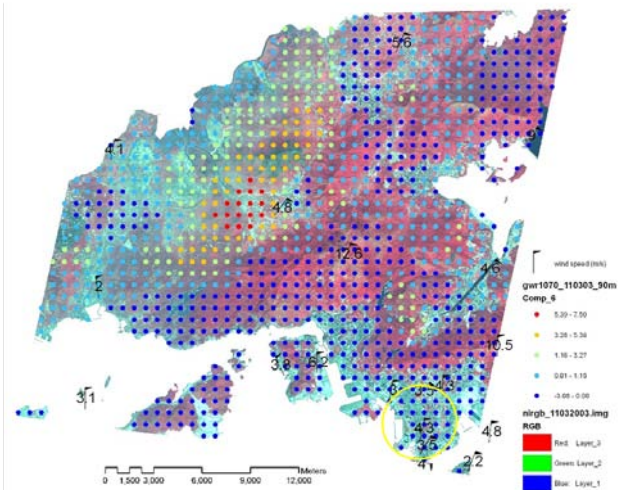


Figure 7. Coefficients contribution of Distance from coast in daytime model 11/03/03

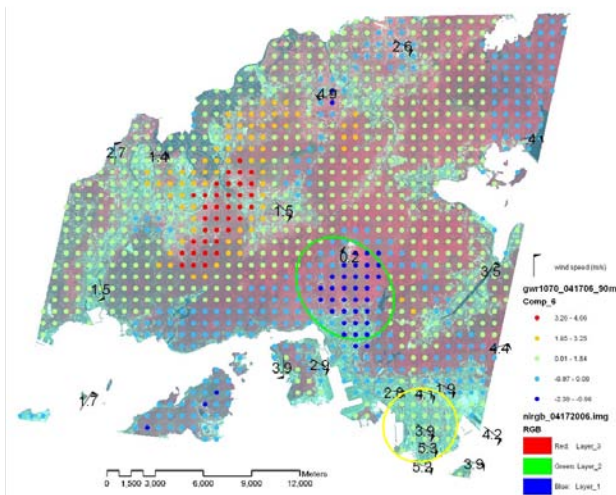


Figure 8. Coefficients contribution of Distance from coast in daytime model 04/17/06

The local comparisons of the relative magnitudes of CC among environmental variables are carried out to investigate the relative contributions of the independent variables to local surface temperature variation. The local dominant factors which contribute most to location specific surface temperature variation represented with the factor possessing the site specific biggest CC can be mapped and compared to differentiate the relative contribution of environmental factors. It was found that the local dominant factors to surface temperature variation are varying in space and time. This uncovered the local disparities of surface heating mechanism, which also indicates the dynamics of local effect of environmental factors on local surface heating in space and time, at one specific location, the order in relative magnitudes of environmental indicators contributing to local surface temperature variation is varied with geographical location and time. Especially in urban area it is difficult to generate a local pattern regarding the relative importance of environmental factor on site specific surface temperature due to the highly variation of the local dominant factors, which may in part due to the relative low resolution in sampling and measurement of some factors like population density and road density comparing with the intensive variation of urban landscape.

The varying relationships between urban surface temperature and local environmental setting in space and time represented with the referred environmental factors made the emerged pattern of local dominant factors in terms of component contribution (CC) challenging to be interpreted and explained which may be due to limitations in the definition and measure of environmental indicators (Holt and Lo, 2008). It also made the clear delineation in the relative importance of environmental indicators on surface temperature variation, in particular surface warming difficult due to the complexity induced by the local varying effect in space and time.

5. CONCLUSIONS

In this research, GWR analysis proved that the spatial variation in relationships between environmental setting and surface temperature was significant with Monte Carlo significance test and distinctive in day-night change. The site specific relation patterns related to each environmental parameter are mapped and compared in order to analyze the

local impact of each factor on surface temperature variation. Through analysis, we found that the local dominant factor accounted for most to the site specific surface temperature variation was highly varied in space and time which prevented a general delineation of the relative association among environment factors to surface temperature disparities. This implied that the effective adaptive measures should be devised locally with reference to day-night needs in the identification of this feature. Comparatively, GWR facilitated the site specific investigation based on local statistical technique. The inference based on GWR model provided enriched information regarding the spatial variation of local environment effect on surface temperature variation which global model cannot approach.

REFERENCES

- Brunsdon, C., Fotheringham, S. and Charlton, M., 1998. Geographically Weighted Regression-Modelling spatial non-stationarity. *The Statistician*, 47(3), pp. 431-443.
- Fotheringham, A.S., Brunsdon, C. and Charlton, M., 2002. *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley & Sons Ltd, England.
- Holt, J. B. and Lo, C. P., 2008. The geography of mortality in the Atlanta metropolitan area. *Computers, Environment and Urban Systems*, 32(2), pp.149-164.
- Huang, Y., 2000. Regional economic development in yangtze river delta since 1978: Jiangsu province as a particular case. PHD thesis, CUHK.
- Huang, Y. and Leung, Y., 2002. Analyzing regional industrialization in Jiangsu province using geographically weighted regression. *Journal of Geographical Systems*, 4, pp.233-249.
- Miller, P. C., Stoner, W. A. and Tieszen, L. L., 1976. A model of stand photosynthesis for the wet meadow tundra at Barrow, Alaska. *Ecology*, 57(3), pp.411-430.
- Oke, T.R., 1982. The energetic basis of the urban heat island. *Quarterly Journal of Royal Meteorology Society*, 108, pp.1-24.
- Su, B. S., Chong, A. K., Moore, A., 2005. Geostatistical Modeling, Analysis and Mapping of Epidemiology of Dengue Fever in Johor State, Malaysia. Paper Presented at SIRC 2005 – The 17th Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, November 24th-25th 2005.
- Voogt, J.A. How Researchers Measure Urban Heat Islands. http://www.epa.gov/heatisland/resources/pdf/EPA_How_to_measure_a_UHI.pdf, EPA presentation.