A COMPARATIVE STUDY OF CITY ENVIRONMENT IN TIANJIN AREA, CHINA AND THE GREATER TORONTO AREA, CANADA BASED ON MULTI-FACTORS OF THE URBAN ECOSYSTEM

Qingni HUANG
Huadong GUO, Director-general
Xinwu LI, Associate Director
Zhongchang SUN
YiXing Ding
Laboratory of Digital Earth Science
Center for Earth Observation and Digital Earth
Chinese Academy of Sciences
No. 9 Dengzhuang South Road, Haiding District
Beijing, China 100094
huangqingni@126.com

ABSTRACT

With the rapid urbanization, more than half of the human population of the world are concentrated on cities and the land use and land cover are undergoing great changes. This has great significant impact on the ecosystem on a local to global scale. Therefore, it has profound significance to understand clearly different urban land surface characteristic spatial and temporal dynamics and their corresponding responds to the regional and even global environment change. In this paper, Tianjin of the developing country China, and the Greater Toronto area of the developed country Canada are selected as study areas. The land surface characteristic parameters including vegetation, impervious surface, soil, water are retrieved from long time series Landsat TM/ETM+ imagery (for Tianjin area, acquired in 1987, 1999 and 2006 respectively; for the Greater Toronto area, acquired in 1985, 2001 and 2006 respectively). The spatial density and dynamics of these parameters, as different indicated factors of the urban ecosystem, are analyzed and compared. This research is of meaningful for scientific urban planning and ecosystem restoring.

KEYWORDS: urban ecosystem, comparative study, multi-factors, support vector machine, V-I-S, LULC

INTRODUCTION

Humanity today is experiencing a dramatic shift to urban living. The global urban population has exceeded the rural population in 2008 and it is estimated that 70% of the world population will live in urban areas by 2050, with more than half of them concentrated in Asia (Seto and Shepherd 2009). Urban areas are hot spots that drive environmental change at multiple scales. Recent research points to mounting evidence that urban sprawl and the concomitant urban land-use and land-cover (LULC) have considerate impacts on climate system (carbon, aerosol and nitrogen), hydro-systems, biogeochemical cycles locally to regionally (Grimm etc. 2008). Therefore, it has profound significance to understand clearly different urban land surface characteristic spatial and temporal dynamics and their corresponding responds to the regional and even global environment change.

Growing interest in urban ecosystem provides rich opportunities for remote sensing research and application. In order to quantify dynamic changes of environmental patterns within cities and establish environmental comparisons between cities of the world, V-I-S (Vegetation-Impervious surface-Soil) model (Ridd 1995) is widely used to quantifying the urban environment. Chen (1996) used merged Landsat TM and SPOT-P data to examine the effect of V-I-S cover types as input to storm runoff models for runoff prediction of storms of different intensities. Gluch et al. (2006) utilized ATLAS (Advanced Thermal Land Applications Sensor) data to determine urban heat island (UHI) effect from V-I-S cover types. Rashed et al. (2001) employed four spectral bands Indian IRS-1C data to study the anatomy of Greater Cairo in terms of end-member fractions(vegetation, impervious surface, soil, and shade) through spectral mixture analysis (SMA). Artificial neural networks (ANNs) algorithms (Hu 2009, Weng 2008, Le 2006) such as Multi-layer perceptron (MLP) feed-forward network and back-propagation (BP) network were applied for estimating end-member fractions within each pixel. Classification and regression tree (CART) algorithm another approach widely used to get the impervious surface end-member(Yang 2003, Jiang 2009, Han 2007). Recently, some researchers have attempted a modern machine learning method -support vector machine (SVM) to map the urban characteristics and get a good result(Esch etc. 2009, Huang 2002, Zhu etc. 2002, Sun etc. 2011).

China, as one of the developing country in Asia of eastern hemisphere, has undergone a constantly accelerating increase of urban population and intense urbanization. While Canada, as one of the developed country in North America of western hemisphere has also undergone urbanization. Yet considerable difference for urbanization process and rate would lead to difference of environmental impact. The coastal city, Tianjin, is situated in the eastern part of the North China Plain whose urban population proportion has exceeded 50% from 1980. The Greater Toronto area, lakeshore city, situated the similar latitude, has the most population of Canada. Therefore, the two areas are selected for dynamic changes of environmental patterns and environmental comparisons. The objective of this paper is to derive V-I-S classification maps and further to proceed environmental comparisons. According to V-I-S model the land surface characteristic parameters including vegetation, impervious surface, soil, water are retrieved from long time series Landsat TM/ETM+ imagery(for Tianjin area, acquired in 1987, 1999, and 2006 respectively; for the Greater Toronto area, acquired in 1985, 2001, 2006 respectively) using integrated classification approaches based on SVM. The spatial density and dynamics of these parameters, as different indicated factors of the urban ecosystem, are analyzed and compared. In the next section we present a summary of the study area the datasets used. This is followed by methodology. And the next is the results and analysis. The conclusions would be followed lastly.

STUDY AREA AND DATASETS

Study Area

Tianjin (Fig. 1) of the developing country, China, and the Greater Toronto area (Fig. 2) of the developed country, Canada were selected as study areas. Tianjin is located within 38.57 ° N-40.25 ° N and 116.72 ° E-118.32 ° E, covering about 11760.26 km^2. Tianjin is one of the four municipalities and its urban land area is the fifth largest in China, ranked only after Beijing, Shanghai, Guangzhou and Shenzhen. In terms of urban population Tianjin is the sixth largest city of China. From 1980, the urban population proportion in Tianjin has exceeded 50% and now reached to 78.01% (9.5809 million urban population of 12.2816 million permanent population) according to statistic data from Tianjin Municipal bureau of statistics. Tianjin's urban area is located along the Hai He River, whose ports, some distance away, are located on Bohai Gulf in the Pacific Ocean. Tianjin's climate is a monsoon-influenced humid continental climate characterized by hot, humid summers, due

to the monsoon, and dry, cold winters, due to the Siberian anticyclone. Toronto is the largest city in Canada and the provincial capital of Ontario. It is located on the northwestern shore of Lake Ontario. With over 2.5 million residents, it is the fifth most populous municipality in North America. Toronto is at the heart of the Greater Toronto Area (GTA), and is part of a densely populated region in Southern Ontario known as the Golden Horseshoe, which is home to 8.1 million residents and has approximately 25% of Canada's population. The census metropolitan area (CMA) had a population of 5.113149 million, and the Greater Toronto Area had a population of 5.555912 million in the 2006 Census.

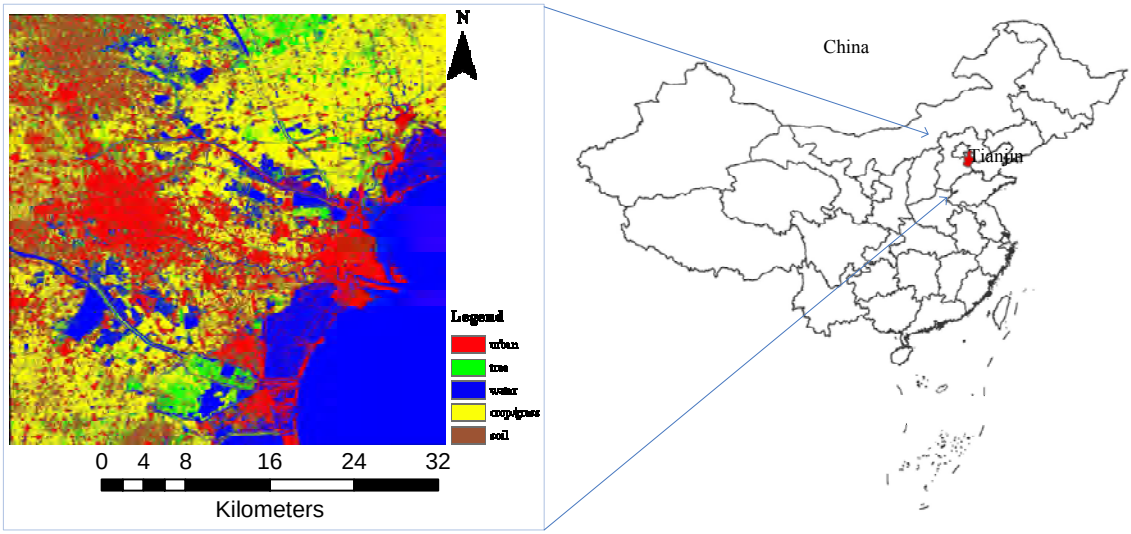

Figure 1. Location map of Tianjin, China

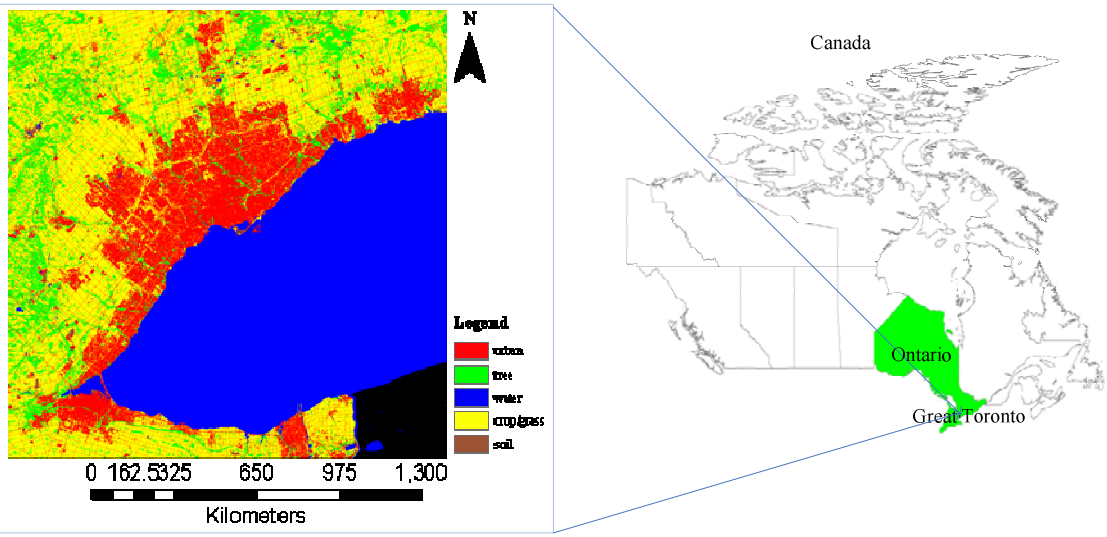

Figure 2. Location map of the Greater Toronto area, Ontario, Canada

Datasets and Preprocessing

In the research, the datasets used were presented as followed (Table 1). There were three period datasets. The preprocessing processes included general geometrical rectification, cloud removal, image matching between temporal imagery and mosaic etc.

Table 1. Datasets used for the study

Area	Data Acquired Time		
	Dataset-1	Dataset-2	Dataset-3
Greater Toronto	1985	2001	2005
Tianjin	1987	1999	2006

METHODOLOGY

In our research, V-I-S (Vegetation-Impervious surface-Soil) model was utilized as the standardization for analyses of the temporal and spatial dynamics in one area and also served for environmental comparison between Tianjin area and the Greater Toronto area. According to the theory of V-I-S model, vegetation, impervious surface, surface and water are the classification result to obtain. Combined with spectral features of urban land cover types, high albedo object, low albedo object, vegetation and water were selected as end-member to join supervised classification using support vector machine (SVM). The classification result does not directly equal to the object we want that is, vegetation, impervious, surface, water. Former researches (Wu and Murray 2003) found that impervious surfaces were located on or near the line connecting the low-albedo and high-albedo end-members in the feature space. That is, the bright impervious class includes materials of high reflectance such as rooftops typically of commercial/industrial buildings. Dark impervious surfaces of low reflectance include recently placed asphalt on roads and parking lots, common in commercial and industrial areas. The class also includes shadowed areas and water. In urban area, water must be considered and it could be easily removal from the low-albedo class using the empirical value of Modified Normalized Difference Water Index (MNDWI) (Xu 2005). Thus it could get the V-I-S-W result through segmentation and combination of classification result of SVM. Fig. 3 showed the data processing framework of research. A brief description of SVM algorithm and V-I-S model were presented as followed.

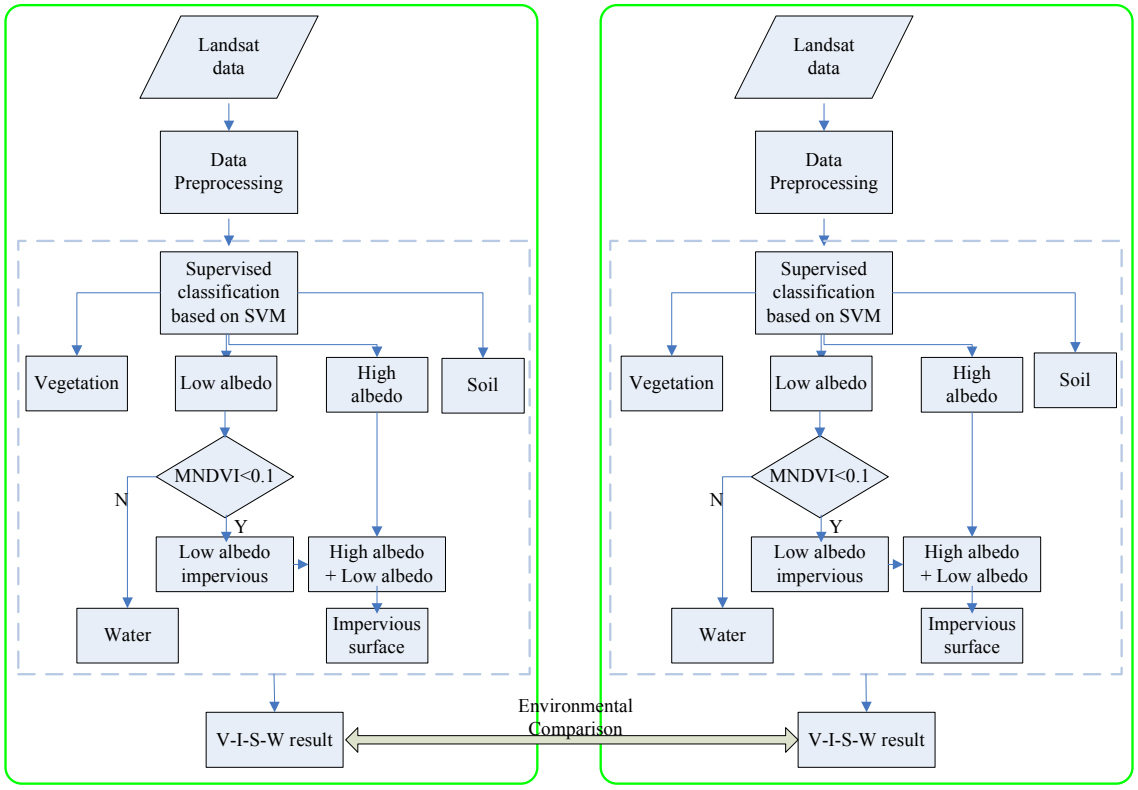

Figure 3. The data processing flowchart of research.

Support Vector Machine (SVM)

The concept of SVM originated from a nonparametric machine learning methodology based on Vapnik's (1995) Structural Risk Minimization (SRM) principle. The basic idea of SVM is to map data into a high dimensional space and to find the hyper-plane of the different classes with the maximum margin between them. SVMs have been used for image classification (Huang 2002, Zhu 2002, Foody 2004, and Pal 2005), soil moisture estimation (Ahmad 2010), image retrieval (Wong 2006), and impervious surface estimation (Esch 2009, Sun 2011 in press) for remote sensing studies in recent years. The theory of SVM has been extensively described in the literature (Burges 1998, and Vapnik 1998). Therefore, only a brief description of the concept of SVM in the framework of classification is given here.

Suppose that training data $(\mathbf{X}_i, y_i)$ ($\mathbf{X} \in R^K$, $i = 1, \ldots, L$, L is the number of the training samples and $y_i \in \{+1, -1\}$ is the class label of the training vector $\mathbf{X}_i$) can be linearly separated by a hyperplane (Eq. (1)):

$$\mathbf{W} \cdot \mathbf{X} + b = 0 \tag{1}$$

Where, the symbol $\mathbf{W}$ is a weight vector and b is a scalar, often referred to as a bias.

To describe the separating hyperplane, let us use the following form:

$$\begin{cases} \mathbf{X}_i \cdot \mathbf{W} + b \geq +1 & \text{for } y_i = +1 \\ \mathbf{X}_i \cdot \mathbf{W} + b \leq -1 & \text{for } y_i = -1 \end{cases} \tag{2}$$

The separating hyperplane that creates the maximum margin is called the Optimal Separating Hyperplane (OSH).The goal of the learning process based on SVM is to find the OSH to separate the training data by solving the following quadratic optimization problem:

Minimize $\frac{1}{2}\|w\|^2 + C(\sum_{i=1}^{L}\xi_i)$

Subject to:

$$\begin{cases} y_i\left[\mathbf{W}\cdot\mathbf{X}_i + b\right] \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, 2, \ldots, L \end{cases} \tag{3}$$

Where the symbol C denotes a penalty parameter on the training error; ξ_i is called the slack variables. The calculations can be simplified by converting the problem with Kuhn-Tucker conditions into equivalent Lagrange dual problem.

$$W(\alpha) = \sum_{i=1}^{L}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{L}\alpha_i\alpha_j y_i y_j K(\mathbf{X}_i, \mathbf{X}_j) \tag{4}$$

Subject to:

$$\sum_{i=1}^{L}\alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \cdots, L \tag{5}$$

Where $i = 1, \ldots, L$ is the sample size and the final decision function is given by

$$f(\mathbf{X}) = \sum_{i=1}^{N}\alpha_i y_i K(\mathbf{X}_i, \mathbf{X}_j) + b \tag{6}$$

In Equations (5) and (6), α_i denotes Lagrange multipliers; the parameter N (usually $N \ll L$) stands for the number of the selected points or support vectors; and the symbol $K(\mathbf{X}, \mathbf{X}_i)$ is the kernel function that measures non-linear dependence between the two input variables $\mathbf{X}$ and $\mathbf{X}_i$. Three admissible kernel functions include: polynomial kernel function, radial basis kernel function (RBF), and sigmoid kernel function.

The steps involved in SVM modeling are: (1) selecting a suitable kernel function and kernel parameter (kernel width G) and (2) specifying the penalty parameter C. In our paper, a RBF kernel is used for constructing the classifier. In addition, specifying parameters G and C is the key step in SVM because their combined values determine the boundary complexity and thus the classification performance (Devos 2009). In our research, the kernel width G was set to 0.15 and the penalty parameter was set to 100 (Sun 2011 in press). For the implementation of the training and modeling procedure, we employed the existing SVM library LIBSVM presented by Chang and Lin (Chang 2008). The LIBSVM represents integrated software for SVM classification, regression, distribution estimation and multi-class classification. In order to keep the classification runtime low, the maximum numbers of training and testing samples for a given class do not exceed 3000 and 5000.

V-I-S Model and Basic Land Cover Mapping

The V-I-S model suggests that the great variety of urban land cover types can be grouped into the three general categories – vegetation, impervious surface, and soil – plus water (Ridd 1995). These four land cover types exhibit highly contrasting influences on the two most important factors in an ecosystem: energy and moisture flux. Variations within each category can be recognized as well by identifying sub-categories of vegetation, impervious surface, soil, and water. Table 2 shows the environmental/ecological basis of the model.

Table 2. Comparative energy and moisture exchange rates

	Healthy green vegetation	Impervious surface	Exposed soil	Water
Energy Insolation	Drives photosynthesis and Evapo-transpiration	Reflects some, but absorbs most, becomes a heat radiator	Generally absorbs and becomes a heat radiator	Absorbs and stores energy; radiates very little
Diurnal heat flux	Minimal	Great	Moderate	Miniscule
Moisture Precipitation	Intercepts and stores in plants and soil; retards runoff	Promotes rapid runoff, miniscule storage	Intercepts and stores; intermediate runoff rate	Accepts and stores
Evapo-transpiration	Significant evaporation and transpiration	Miniscule evaporation; no transpiration	Moderate evaporation; no transpiration	Considerable evaporation; no transpiration

From Table 2, we can see that t these four land cover types have different response to energy and moisture which lead to spectral difference. According to spectral features of land cover types, generally high albedo object, low albedo object, vegetation and water were selected as end-member. As mentioned above, the impervious surfaces can be estimated based this relationship by adding the fractions of high-albedo and low-albedo end-members excluded water. Water can be easily detected by the low-albedo classification result and the empirical value of MNDWI as follows.

$$MNDWI = \frac{\rho_2 - \rho_5}{\rho_2 + \rho_5} \tag{7}$$

Where ρ_2 and ρ_5 are the unitless planetary reflectances of band2 and band5 of TM, respectively. The low-albedo class pixel is regarded as water when satisfying $MNDWI > 0.1$. Vegetation and soil end-members have little or no contribution to impervious surface estimation. Vegetation has particular spectral and is easily separated from others and the left is soil excluding the others. In this way the basic land cover including vegetation, water, impervious surface and soil can be mapped.

RESULTS AND ANALYSIS

According to the flowchart as showed in Fig. 3, the basic land cover maps of Tianjin area consisted of V-I-S-W were acquired (Fig. 4). As the same way, the V-I-S-W feature map of the Greater Toronto area, was gotten as followed (Fig. 5).

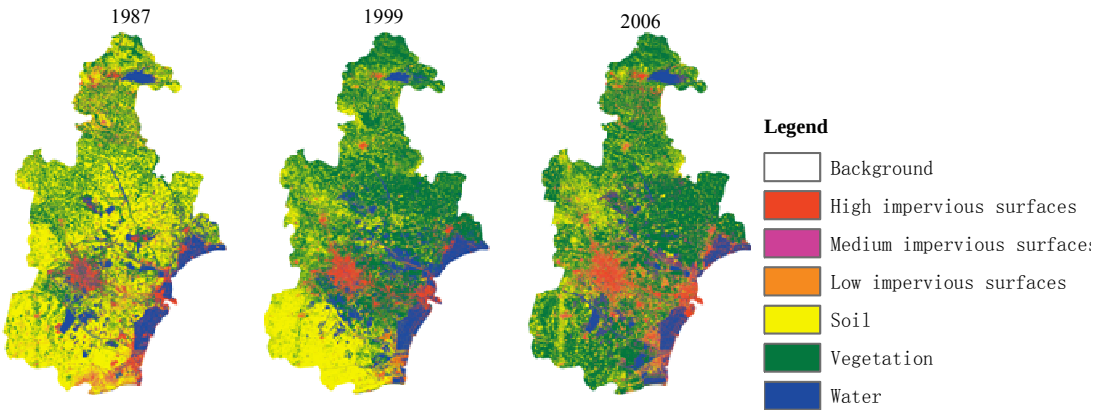

Figure 4. the Basic land cover special and temporal dynamics in Tianjin area.

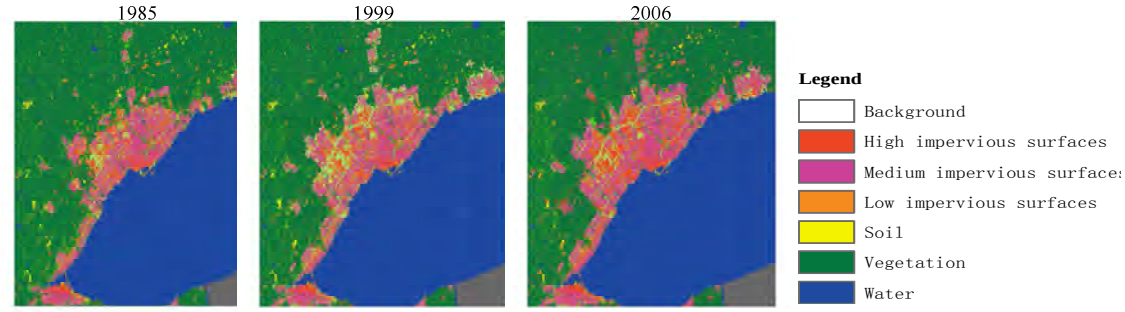

Figure 5. the Basic land cover special and temporal dynamics in of the Greater Toronto area.

From Fig.4 it can be seen that it is very obvious that the impervious surface of Tianjin area has a significantly increase in the past twenty years which demonstrates that the rapid urbanization has taken place. The Fig. 4 also shows that water area has decreased sharply. The hot indicator impervious surface increased and the cold indicator water decreased would explain the warmer trend in Tianjin areas. In Fig. 5, a clearly impervious increase also could be seen. Comparatively, Tianjin has great rate of urbanization.

CONCLUSIONS

This paper presented a case study to get the space and temporal dynamics change in urban area and the environmental difference between two areas in the world based on the V-I-S-W model. It is very effective and easily utilized. The further work is to use V-I-S-W value to get a ternary V-I-S diagram. And a profound analysis would be presented combining with the population data and weather data.

ACKNOWLEDGEMENTS

This research in this paper was sponsored by 973 Program “Earth Observation for Sensitive Factors of Global Change: Mechanisms and Methodologies” (No. 2009CB723906). The authors would like to thank Prof Yingkui

Li (Department of Geography, University of Tennessee) for providing the LULC data of Greater Toronto and giving helpful discussion on improving the presentation of this paper.

REFERENCES

Ahmad, S., A. Kalra and H. Stephen, 2010. Estimating soil moisture using remote sensing data: A machine learning approach, *Advances in Water Resources, 33*(1): 69–80.

Burges, C. J. C., 1998. A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery, 2*(2):121-167.

Chang, C. C., and C. J. Lin, 2008. LIBSVM: *A Library for Support Vector Machines*. Department of Computer Science and Information Engineering, National Taiwan University, Taiwan. Available online at: http://www.csie.ntu.edu.tw/~cjlin/libsvm (accessed 15 October 2008).

Chen, J, 1996. *Satellite image processing for urban land cover composition analysis and runoff estimation*, Doctoral dissertation, The Department of Geography, University of Utah, Salt Lake City, Utah.

Devos, O., C. Ruckebusch, A. Durand, L. Duponchel and J.-P. Huvenne, 2009. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation, *Chemometrics and Intelligent Laboratory Systems, 96*(1): 27-33.

Esch, T., V. Himmler, G. Schorcht, M. Thiel, T. Wehrmann, F. Bachofer, C. Conrad, M. Schmidt and S. Dech 2009. Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data, *Remote Sens. Environ, **113**(8):*1678–1690.

Foody, G. M., and A. Mathur, 2004. A relative evaluation of multiclass image classification by support vector machines, *IEEE Transactions on Geoscience and Remote Sensing, 42*(6): 1335-1343.

Grimm, Nancy B, Stanley H. Faeth, Nancy E. Golubiewski, Charles L. Redman, Jianguo Wu, Xuemei Bai and John M. Briggs, 2008. Global Change and the Ecology of Cities, *Science, 319 (5864):* 756-760.

Gluch, R, DA Quattrochi, JC Luvall, 2006. A multi-scale approach to urban thermal analysis, *Remote Sens Environ, 104 (2):*123–132.

Huang, C., L.S. Davis and J.R.G. Townshend, 2002. An assessment of support vector machines for land cover classification, *Int. J. Remote Sens, **23**(4):*725-749.

Han, J and M. Kamber, 2007. *Data mining concepts and techniques*, China Machine Press ,Beijing.

Hu, X and Q. Weng, 2009. Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks, *Remote Sens. Environ, 113(10):*2089–2102.

Jiang,vL., M. Liao, H. Lin and L. Yang, 2009. Synergistic use of optical and InSAR data for urban impervious surface mapping: a case study in Hong Kong, *Int. J. Remote Sens, 30(11):*2781–2796.

Lee,S and R. G. Lathrop, 2006. Subpixel analysis of Landsat ETM+ using self-organizing map (SOM) neural networks for urban land cover characterization, *IEEE Trans. Geosci. Remote Sens, 44(6):*1642–1654.

Pal, M., and P. M. Mather, 2005. Support vector machines for classification in remote sensing, *International Journal of Remote Sensing, 26*(5):1007-1011.

Rashed, T, JR Weeks, MA Gadalla, AG Hill, 2001. Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: a case study of the greater Cairo region, Egypt, *Geocarto Int, 16(4)*:5–15.

Ridd, MK, 1995. Exploring a V-I-S (Vegetation-Impervious Surface-Soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities, *Int J Remote Sens, 16:*2165–2185.

Seto, Karen C and J Marshall Shepherd, 2009. Global Urban Land-Use Trends And Climate Impacts, *Current*

*Opinion in Environmental Sustainability, 1:*89-95.

Sun, Z., H. Guo, X. Li, L. Lu, and X. Du, (in press). Estimating urban impervious surfaces from Landsat-5 TM imagery using multi-layer perceptron neural network and support vector machine, *Journal Applied of Remote Sensing.*

Vapnik, V. N., 1995. *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.

Vapnik, V. N., 1998. *Statistical Learning Theory*, Wiley, New York.

Weng, Q and X. Hu, 2008. Medium spatial resolution satellite imagery for estimating and mapping urban impervious surfaces using LSMA and ANN, *IEEE Trans. Geosci. Remote Sens, 46(8):*2397–2406.

Wong, W. -H., and S. -H. Hsu, 2006. Application of SVM and ANN for image retrieval, *European Journal of Operational Research, 173*(3):938–950.

Wu, C. and A.T. Murray, 2003. Estimating impervious surface distribution by spectral mixture analysis, *Remote Sens. Environ, 84(4):*493-505.

Xu, H., 2005. A study on information extraction of water body with the modified normalized difference water index (MNDWI), *Journal of Remote Sensing, 9(5):*589-595.

Yang, L, C. Huang, C.G. Homer, B.K. Wylie and M.J. Coan, 2003. An approach for mapping large-scale impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery, *Canadian Journal of Remote Sensing, 29(2):*230-240.

Zhu, G., and D.G. Blumberg, 2002. Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel, Remote Sens. *Environ, 80(2):*233-240.