Issues and challenges of remote sensing–based local climate zone mapping for high-density cities

Yong XU, Chao REN, Meng CAI and Ran WANG
Institute of Future Cities (IOFC),
The Chinese University of Hong Kong
Hong Kong, China
xuyong@cuhk.edu.hk

***Abstract*—The local climate zone (LCZ) classification system provides a standard organizing principle for urban climate studies, such as urban heat island analysis, in which urban structures are classified into 17 standard LCZ classes according to urban land surface properties, such as land use and land coverage. Based on the standard definitions of the LCZ, freely available Landsat satellite data have been used to generate LCZ maps for cities worldwide. This research aims to evaluate the performance of the existing methods for high-density cities and to discuss the main issues, challenges, and possible solutions for further applications. Our experimental results indicate that three main factors, including training samples, input features, and the classifiers used, had a decisive effect on the final mapping result. In particular, low-quality training samples coupled with complex urban scenarios tended to yield low-quality LCZ mapping results. To improve the LCZ mapping results, some practical solutions and suggestions were given for future applications.**

Keywords—Local Climate Zones, High-density, WUDAPT

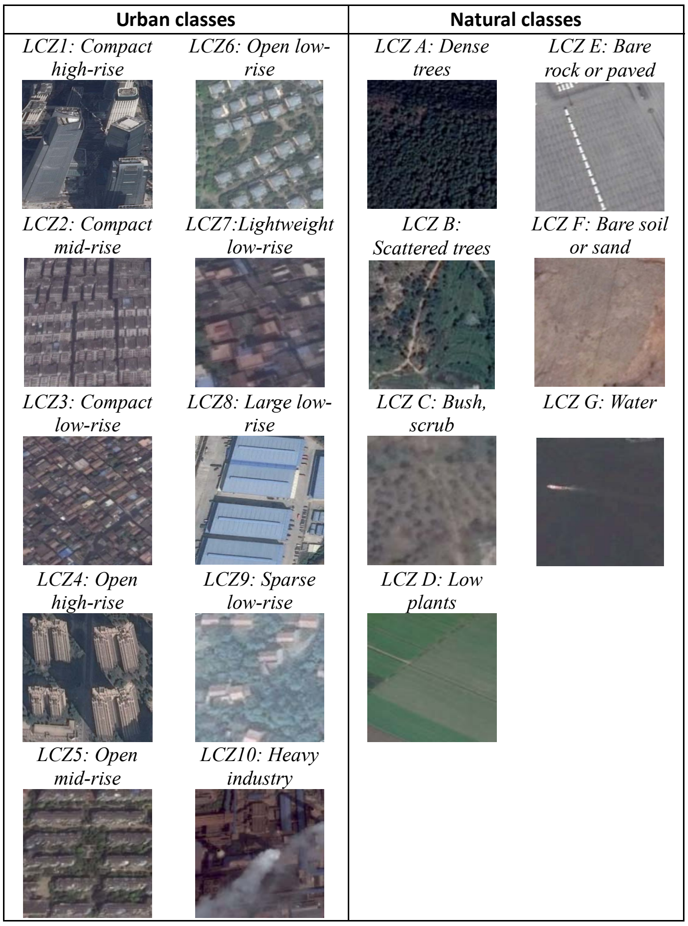
I. INTRODUCTION

The urban heat island (UHI) effect has become a perennial concern for various Asian cities. High UHI intensity has been detrimental to the quality of life in urban areas, even increasing the rate of heat-related deaths, and thus poses considerable challenges for the sustainable development of those cities [1], [2].

A novel classification system for urban land use, the local climate zone (LCZ) system, has been proposed, in which local urban structures are classified according to their corresponding UHI effect [3]. Urban structures are sorted into 17 standard classes according to land surface properties such as land use and coverage, building morphology, and building materials.

Based on the LCZ concept, freely available satellite data have been used to obtain promising LCZ mapping results for various European cities [4]–[6]. Indeed, previous LCZ studies have mainly focused on European cities, where the urban morphology tends to be amenable to the standard LCZ definitions. Thus, in this study, we aim to evaluate the performance of the conventional remote sensing–based LCZ mapping method for high-density cities and discuss the issues and challenges of the existing method and the corresponding solutions.

TABLE I. STANDARD LCZ CLASSES [3], EXAMPLES FROM GUANGZHOU



II. LOCAL CLIMATE ZONES (LCZ)

The conventional definition of UHI refers to the temperature difference between downtown and neighboring rural areas based on an urban/rural classification system [7]. However, this classification neglects quantitative metadata on site exposure and land coverage. The measurements, site descriptions, and corresponding assessments of UHI lack an objective protocol for the basis of comparison, which affects

978-1-5090-5808-2/17/$31.00 ©2017 IEEE

the value of previous urban climate studies for practical usage [8]. In this context, a new LCZ system based on dividing urban structures into 17 standard LCZ classes was proposed to guide UHI studies [3]. This system allows the cross-comparison of different UHI studies within and among cities.

Table 1 summarizes the 17 standard LCZ classes, including 10 urban classes and 7 natural classes. Each LCZ class is strictly defined on the basis of a set of standard parameters of surface properties, such as land use and coverage, as mentioned above.

III. REMOTE SENSING–BASED LCZ MAPPING

As a fast and efficient way to obtain land surface data across large areas, satellite technology has been extensively used to classify urban structures [9], [10]. Based on the concept of local climate zones, a new initiative known as the World Urban Database and Access Portal Tools (WUDAPT) (http://www.wudapt.org) was launched in 2012 to generate LCZ maps and collect urban morphology data worldwide [11]–[14]. It aims to use freely available satellite data to provide local climate zone data in support of urban climate studies, e.g., UHI assessment.

In this study, a procedure based on WUDAPT was adopted. As shown in Figure 1, it included four main stages. First, freely available Landsat satellite data were downloaded and preprocessed. Second, training areas for different LCZ classes were selected via the Google Earth platform to ensure high quality. Third, based on the acquired Landsat data with a resolution of 30m *30m and training areas, several supervised classification methods (e.g., random forest) were used to generate local climate maps. Finally, the generated LCZ map with a spatial resolution of 30 m was resampled to 120 m according to the requirement of urban climate studies [6][12], and the overall accuracy of the generated local climate maps was assessed by comparing them with an external independent test dataset.

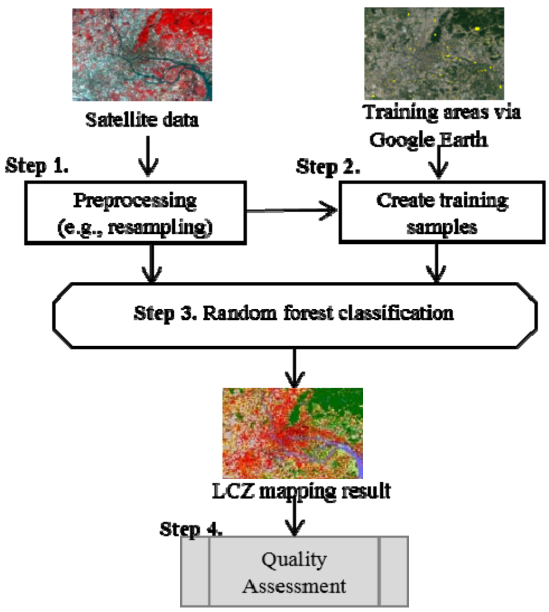


Fig.1. Procedure of remote sensing–based LCZ classification approach.

IV. CASE STUDY

A. *Study area*

Guangzhou, the largest city in southern China and the third largest in China, was selected for this study. Its location was shown in Figure 2. Guangzhou is the capital of Guangdong province in southern China. It is located in the northern part of the Pearl River Delta, facing the South China Sea, near Hong Kong and Macau. It has a permanent population of 13.5 million as of 2013 and a total area of approximately 7,434 km^2. It has a typical subtropical monsoon climate with an annual average temperature of 20°C to 22°C.

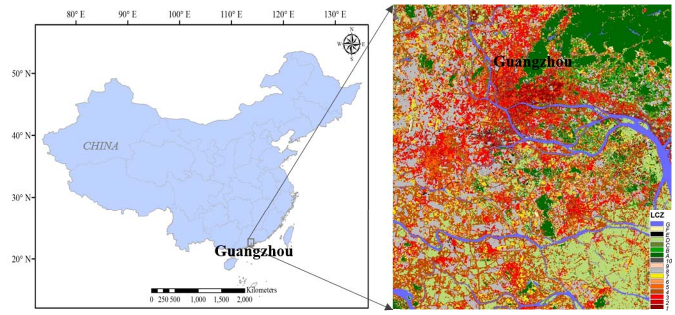


Fig. 2. Location of Guangzhou in China

Due to rapid urbanization and industrialization, this region has become one of the most densely populated area of China. Intensive and extensive economic and commercial activities have posed severe environmental challenges for this city, e.g., low wind speeds and a high thermal load [15], [16].

B. *Input data*

The acquired satellite data include one Landsat 8 image from year 2014 (October 15, 2014). The original satellite data were downloaded from the U.S. Geological Survey (http://glovis.usgs.gov). In this study, both spectral and textural information from Landsat data were used to generate LCZ mapping result, given that textural information can slightly improve the classification accuracy [17]. The spectral features included bands 1-7 and 10-11 of the Landsat 8 data. The textural features included 8 gray-level co-occurrence textures (GLCM), representing mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation from the first principle component of Landsat data using a window size of 150m *150m. Thus, the total number of the used features from Landsat data is 17.

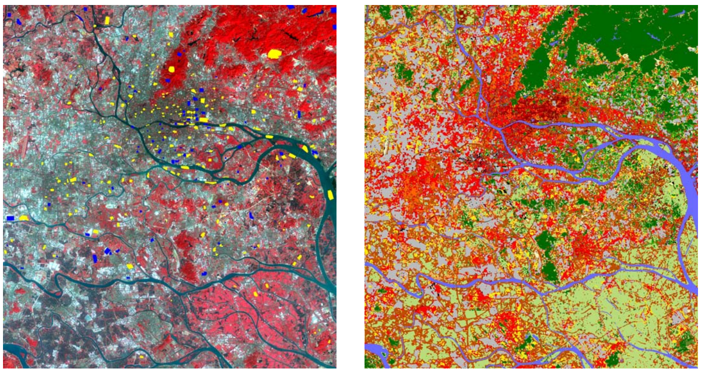
In addition to the satellite data, the input data also included training and validation data. The training and independent validation data were collected via Google Earth. To ensure that all validation data were of high quality, they were cross-checked via the Baidu 3D street map (www.baidu.com). Figures 3(a) shows the distribution of the selected training and validation areas, and Table II shows details of the training and validation samples in this study area, including the total number of pixel-based training samples (30*30m) and validation samples (120m*120m) per LCZ. Given that the actual proportion of some LCZ classes is rather small, these LCZ classes are not considered in verification (e.g., LCZ 9).

ABLE II. TRAINING AND VALIDATION SAMPLES FOR GUANGZHOU

	Number of training samples (30m*30m)	Number of validation samples (120m*120m)
LCZ1:Compacthigh-rise	1226	40
LCZ2:Compact mid-rise	1824	24
LCZ3:Compact low-rise	2609	39
LCZ4: Open high-rise	3779	40
LCZ5: Open mid-rise	1700	37
LCZ6: Open low-rise	1360	33
LCZ7:Lightweight	1209	16
LCZ8: Large low-rise	3246	40
LCZ9: Sparse low-rise	311	N/A
LCZ10: Heavy industry	459	40
LCZ A: Dense trees	4661	40
LCZ B: Scattered trees	604	11
LCZ C: Bush, scrub	477	N/A
LCZ D: Low plants	2966	40
LCZ E: Bare rock/paved	312	34
LCZ F: Bare soil/sand	267	N/A
LCZ G: Water	3256	40

C. Results

Figure 3(b) shows the LCZ mapping result based on the random forest classification method using the processed satellite data (shown in Figure 3(a)), from which it is apparent that the patterns of both the urban and natural LCZ classes are highly consistent with the actual distributions. In this image, most of the red coloration corresponds to high-density urban areas (LCZs 1–3), e.g., the centre of Guangzhou. Most of the green coloration, indicating areas of dense vegetation, are located in the north-east of this region. The areas in south and south-east of Guangzhou are lightly green-colored, corresponding to farmland.



(a) (b)

Fig. 3. LCZ mapping result for PRD region. (a) Satellite data, training areas (highlighted with yellow), and validation areas (highlighted with blue); (b) LCZ mapping result.

D. Validation

Four accuracy indices, including overall prediction accuracy, kappa index, producer accuracy, and user accuracy, were used to assess the accuracy of classification by comparing the mapping results with the independent validation result. All values fall between 0 and 1, where a larger value indicates a better prediction result.

TABLE III. CONFUSION MATRIX FOR GUANGZHOU LCZ MAPPING RESULTS USING LANDSAT DATA

	1	2	3	4	5	6	7	8	10	A	B	D	E	G	∑ref	PA
1	27	0	0	13	0	0	0	0	0	0	0	0	0	0	40	.68
2	0	12	3	0	8	0	0	0	1	0	0	0	0	0	24	.50
3	0	0	33	0	0	1	5	0	0	0	0	0	0	0	39	.85
4	16	0	1	17	6	0	0	0	0	0	0	0	0	0	40	.43
5	1	7	0	7	19	0	1	0	0	0	2	0	0	0	37	.51
6	0	0	2	8	3	5	13	0	0	0	0	2	0	0	33	.15
7	0	0	0	0	0	5	11	0	0	0	0	0	0	0	16	.69
8	0	0	0	1	0	0	0	39	0	0	0	0	0	0	40	.98
10	0	4	5	0	0	0	1	3	19	0	8	0	0	0	40	.48
A	0	0	0	0	0	0	0	0	0	40	0	0	0	0	40	1.0
B	0	0	0	0	0	0	0	0	0	0	5	6	0	0	11	.45
D	1	0	0	4	0	9	0	0	0	2	0	24	0	0	40	.60
E	0	0	0	1	15	0	0	9	4	3	0	0	2	0	34	.06
G	0	0	0	0	0	0	0	0	0	0	0	0	0	40	40	1.0
∑class	45	23	44	51	51	20	31	51	24	45	15	32	2	40	Kappa:	0.59
UA	.60	.52	.75	.33	.37	.25	.35	.76	.79	.89	.33	.75	1.0	1.0	OA (%):	62

Table 3 shows the confusion matrix, in which the overall prediction accuracy for the final LCZ mapping result of the study area is 62%, with a standard kappa index of 0.59. This result is even worse than that obtained in a similar study for Kyiv, Ukraine, in which the overall accuracy using Landsat data was 64%. These results illustrate that obtaining overall accuracy for complex urban scenarios remains challenging.

V. DISCUSSIONS

Although the WUDAPT method provides a fast and efficient way to obtain urban morphology data, our testing results indicated that the quality of the LCZ mapping result was still not satisfactory, especially for high-density and highly compact urban areas. The main reason might be due to the limitation of Landsat data itself in differentiating certain LCZ classes, given that some classes, e.g., factories and non-industrial buildings, present very similar spectra based on Landsat data [12][13]. Based on the spectral features of different LCZ classes obtained from the Landsat data, it was found that some LCZ classes were liable to be misclassified: for example, LCZ 2 (compact mid-rise buildings), LCZ 7 (lightweight low-rise), LCZ 10 (heavy industry, i.e., factories), and LCZ E (bare rock or paved) were frequently conflated with each other. This confusion among classes is also one of the main reasons why some classes were always predicted with slightly lower accuracy than the others, regardless of location.

Other than the Landsat data itself, some reasons, like the quality of training samples provided by the helpers, might also lead to low-quality mapping results. An experiment conducted at a WUDAPT workshop in Hong Kong on Dec. 13, 2015, showed that the overall accuracy of the selected training samples provided to the participants without prior knowledge was around 70%, which means that one third of the training samples were inaccurate. Moreover, different classifiers sometimes generate slightly different results [12]. By adopting four different classifiers, including maximum likelihood (ML), neural network (NN), support vector machine (SVM), and random forest (RF). Our experimental results showed that the accuracies using these four classifiers ranged from 59% to 64%. This finding confirms that different classifiers generated different prediction results in the present case, with the overall accuracy based on different classifiers varying by around 5%.

Based on the above-mentioned issues, several solutions are provided as follows. First, due to the spectral similarity of some LCZ classes, further improvements might include the integration of other satellite data (e.g., ASTER and SAR data) to provide additional information to differentiate similar LCZ classes. Second, given that the training samples provided by non-qualified operators might contain intrinsic errors that affected the final LCZ mapping result, we recommend the creation of a standard training database at the continent or country level to eliminate the effects of the training samples on LCZ applications for world-wide cities. Third, the LCZ classification methods in use operate at the pixel level, in which spatial contextual information about the surrounding area has not been fully utilized. However, the original LCZ system was actually defined on the basis of local areas, where spatial contextual information should be considered. Thus, development of object-based LCZ classification method to include additional spatial contextual information would merit further study.

VI. Conclusions

The WUDAPT project is a new initiative to provide LCZ maps for cities worldwide in a fast and efficient way. In this study, we evaluated the performance of the conventional WUDAPT method for high-density cities. The experimental results obtained from the analysis of high-density Chinese city of Guangzhou showed an overall accuracy of 62%, lower than those of similar studies conducted in other study areas (e.g., for Kyiv, Ukraine, with an overall accuracy of 64%). This finding indicates that the conventional WUDAPT method might generate low-quality LCZ mapping results for high-density cities.

Based on the analysis of the results, we propose that three main factors, including limitation of Landsat data itself, the quality of training samples, and the classifiers, might have impact on the final mapping result. Low-quality training samples tended to generate low-quality prediction results. Due to the limitations of the input features of the Landsat data, some LCZ classes were still predicted with low accuracy even when using high-quality training samples. Moreover, different classification methods generated different results. Based on these findings, various practical solutions, including integration of additional satellite data, the establishment of a standard training sample database were proposed to aid the generation of high-quality LCZ mapping results for complex urban scenarios.

Acknowledgment

This work was supported by the Hong Kong Research Grants Council General Research Fund (CUHK 14643816).

References

[1] C. Ren, E. Ng, and L. Katzschner, "Urban climatic map studies: a review. *International Journal of Climatology*, vol. 31, no.15, pp.2213-2233, 2011.

[2] M. Roth, "Review of urban climate research in (sub) tropical regions," *International Journal of Climatology*, vol. 27, no. 14, pp. 1859-1873, 2007.

[3] I. D. Stewart, and T. R. Oke, "Local Climate Zones for Urban Temperature Studies," *Bulletin of the American Meteorological Society*., vol. 93, no. 12, pp. 1879-1990, Dec. 2012.

[4] P. Gamba, G. Lisini, P. Liu, P. J. Du, and H. Lin, "Urban climate zone detection and discrimination using object-based analysis of VHR scenes," *Proceedings of the 4th GEOBIA, Rio de Janeiro*, Brazil, 2012, pp. 70-74.

[5] B. Bechtel and C. Daneke, "Classification of Local Climate Zones Based on Multiple Earth Observation Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*., vol. 5, no. 4, pp. 1191–1202, Aug. 2012.

[6] B. Bechtel et al., "Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities," *International Journal of Geographic Information*, vol. 4, no. 1, pp. 199-219, Feb. 2015.

[7] T. Oke, "City size and the urban heat island," *Atmospheric Environment*, vol. 7, pp.769-779, 1973.

[8] I. D. Stewart, "A systematic review and scientific critique of methodology in modern urban heat island literature," *International Journal of Climatology*, vol. 31, no.2, pp. 200-217, Feb. 2011.

[9] P. Gamba, M. Aldrighi, and M. Stasolla, "Robust extraction of urban area extents in HR and VHR SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*., vol. 4, no. 1, pp. 27-34, Mar. 2011.

[10] M. Voltersen, C. Berger, S. Hese, and C. Schmullius," Object-based land cover mapping and comprehensive feature calculation for an automated derivation of urban structure types at block level," *Remote Sensing of Environment*., vol. 154, pp. 192-201, Nov. 2014.

[11] G. Mills, J. Ching, L. See, B. Bechtel, and M. Foley, "An Introduction to the WUDAPT project," in 9th International Conference on Urban Climate, Toulouse, 2015.

[12] B. Bechtel, L. See, G. Mills, and M. Foley, "Classification of Local Climate Zones Using SAR and Multispectral Data in an Arid Environment," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9. no. 7, pp. 3097-3105, July. 2016.

[13] O. Danylo, L. See, B. Bechtel, D. Schepaschenko, and S. Fritz, "Contributing to WUDAPT: A Local Climate Zone Classification of Two Cities in Ukraine," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9. no. 5, pp. 1841-1853, May. 2016.

[14] C. Ren, and M. Cai, "Local Climate Zone (LCZ) Classification Using the World Urban Database and Access Portal Tools (WUDAPT) Method: A Case Study in Wuhan and Hangzhou," *presented at the Fourth International Conference on Countermeasure to Urban Heat Islan*ds (4th IC2UHI), Singapore, 2016.

[15] Q. H, Weng and S. Yang, "Urban air pollution patterns, land use, and thermal landscape: an examination of the linkage using GIS," *Environmental monitoring and assessment*, vol. 117, no. 1, pp. 463-489, June. 2006.

[16] E. Ng, C. Yuan, L. Chen, C. Ren, and J. C. Fung, "Improving the wind environment in high-density cities by understanding urban morphology and surface roughness: a study in Hong Kong," *Landscape and Urban Planning*, vol. 101, no. 1, pp. 59-74, May. 2011.

[17] A. P. Montanges, G. Moser, H. Taubenbock, M. Wurm, and D. Tuia, "Classification of urban structural types with multisource data and structured models," *in Urban Remote Sensing Event (JURSE)*, 2015 Joint, 2015, pp. 1–4.