



IMPROVING LOCAL CLIMATE ZONES AUTOMATIC CLASSIFICATION BASED ON PHYSIC-MORPHOLOGICAL URBAN FEATURES

Eldesoky, Ahmed Hazem Mahmoud ^{1*}; Colaninno, Nicola ² & Morello, Eugenio ³

Initial submission: 2019-06-16; **Definitive submission:** 2019-10-24; **Publicación:** 2019-12-21

Citación: Eldesoky, Ahmed H. M. *et al.* (2019). Improving Local Climate Zones Automatic Classification Based on Physic-Morphological Urban Features. In *XIII CTV 2019 Proceedings: XIII International Conference on Virtual City and Territory: "Challenges and paradigms of the contemporary city"*: UPC, Barcelona, October 2-4, 2019. Barcelona: CPSV, 2019, p. 8663. E-ISSN 2604-6512. DOI <http://dx.doi.org/10.5821/ctv.8663>

Abstract

The Local Climate Zone (LCZ) classification scheme, introduced by Stewart and Oke (2012), offers promising opportunities for better studying the urban climate phenomena at the micro- and local scale (e.g. the urban heat island effect). However, although several methods have been introduced to apply the concept of LCZs to cities, only few utilize publicly available data, like, for instance, the World Urban Database and Access Portal Tools (WUDAPT). However, to date, results are relatively rough, and frequent quality assessments demonstrate moderate overall accuracy. This paper proposes an approach for improving the quality of LCZ automatic classification, combining freely available multispectral satellite imagery together with morphological features of the urban environment. An overall accuracy of 67% was achieved for the Metropolitan City of Milan with an improvement of 12% with respect to using only Landsat 8 multispectral and thermal data. This ascertains the physic-morphological nature of the LCZs and opens the possibility for mapping more accurate LCZs without the need for additional thermal information.

Key words: Local Climate Zones; WUDAPT; Urban Heat Island; Remote Sensing

1. Introduction

Climate change is among the top global environmental concerns in the recent years. In urban areas, extreme heatwaves and the urban heat island (UHI) effect are clear manifestations of such a concern, with serious implications on human comfort and health, energy demand, and the resilience of cities in general. The concept of Local Climate Zones (LCZs), firstly introduced by Stewart and Oke (2012) and defined as "regions of uniform surface cover, structure, material, and human activity that span hundreds of meters to several kilometers in horizontal scale" (p. 1884), opens promising opportunities for better assessing the UHI phenomenon at the micro- and local scale, rather than relying only on the classical urban–rural classification. In particular, the LCZ classification scheme relates the physical and morphological features of the urban environment to micro- and local climate conditions.

Several methods have been introduced to apply the concept of LCZs to cities. These vary between in-situ measurements, to record relevant geometric and surface cover parameters; manual sampling based on user's knowledge and experience; and GIS or remote sensing-based methods (Ren et al., 2016; Wang, Ren, Xu, Lau, & Shi, 2018). For instance, the World Urban Database and Access Portal Tools (WUDAPT), introduced by Mills et al. (2015), is among the most ambitious initiatives to build a worldwide urban climate database and generate

¹ Università IUAV di Venezia, ORCID: 0000-0002-6381-8956, ² Laboratorio di Simulazione Urbana Fausto Curti, Dept. Architecture and Urban Studies, Politecnico di Milano, ORCID: 0000-0003-4428-639, ³ Laboratorio di Simulazione Urbana Fausto Curti, Dept. Architecture and Urban Studies, Politecnico di Milano, ORCID: 0000-0003-2382-6123. *Contact e-mail: ahmeldesoky@iuav.it



effective, automatic, remote sensing-based classifications of LCZs based on publicly available satellite data. In particular, the WUDAPT project presents three level products of urban climate data: level 0, at regional and city scale; level 1, at neighborhood scale; and level 2, at building scale (Mills et al., 2015). In WUDAPT level 0 product, each LCZ is described and classified depending on the characteristics of some key urban parameters consistent with different urban micro- and local climates (e.g. surface cover, materials, building geometry). This is done by conducting a supervised classification, using freely available multispectral and thermal satellite imagery and training samples, defined through very high-resolution (VHR) aerial or satellite imagery. However, although, results provide a multi-categorical, comprehensive classification, LCZ maps are relatively rough (Ren et al., 2016) and frequent quality assessments demonstrate moderate overall accuracy, i.e. 50 to 60% (Bechtel et al., 2019). Besides, only few freely available satellite data have thermal information.

In this paper, our goal is to investigate the possibility of improving the quality of LCZ maps both in terms of spatial resolution and accuracy. Firstly, by using fine spatial resolution input data, to conduct the pixel-based classification; moreover, digitizing relatively small-sized training samples, to consider some small-scale urban units (Table 1), that can influence upon the micro- and local climate (Oke, Mills, Christen, & Voogt, 2017). Secondly, by combining physical features, retrieved by means of multispectral satellite imagery, together with morphological features of the urban environment. In particular, we include surface albedo (both narrow and broadband albedo), Normalized Difference Vegetation Index (NDVI), building heights, and Sky View Factor (SVF) as main variables.

Table 1. The hierarchy of urban units and their urban climate phenomena

Urban units	Built features	Urban climate phenomena	Typical horizontal scale	Climate scale
Facet	Roof, wall, road	Shadows, storage heat flux, dew and frost patterns	10 x 10 m	Micro
Element	Residential building, high-rise warehouse	Wake, stack plume	10 x 10 m	Micro
Canyon	Street, canyon	Cross-street shading, canyon vortex, pedestrian bioclimate, courtyard climate	30 x 200 m	Micro
Block	City block, factory	Climate of park, factory cumulus	0.5 x 0.5 Km	Local
Neighborhood	City center, quarter, industrial zone	Local neighborhood climates, local breezes, air pollution district	2 x 2 Km	Local
City	Built-up area	Urban heat island, smog dome, patterns of urban effects on humidity, wind	25 x 25 Km	Meso
Urban region	City plus surrounding countryside	Urban 'plume', cloud and precipitation anomalies	100 x 100 Km	Meso

Source: (Oke et al., 2017, p. 19)

2. Data and study area

Although the concept of LCZs refers only to a specific range of horizontal length scales (i.e. hundreds of meters to several kilometers), there is not an optimal resolution to conduct the pixel-based classification. Nevertheless, mapping fine-resolution LCZ maps requires utilizing relatively higher resolution input data (Bechtel et al., 2015). In this study, we used Landsat 8 satellite imagery, acquired for a summer day during daytime (15/08/2018 10:10 AM), and where the cloud cover over the region of interest (ROI) was less than 10%. In particular, Landsat 8



(launched on February 2013) carries two sensors, i.e. the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIR), which collect optical and thermal data at a spatial resolution of 30 m images, ranging from visible to near infrared (NIR), and shortwave infrared (SWIR); 100 m thermal images, which are resampled at 30 m to match the multispectral bands; and 15 m panchromatic images (USGS, 2015), as reported in Table 2.

Also, high quality digital topographic database (DTDB) of building footprints, where building heights information is available, was obtained from *Geoportale Lombardia*. The DTDB is freely available for most of the Lombardy municipalities and comes in different scales, i.e. 1:2,000 for urban centers; 1:5,000 for the extra urban areas; and 1:10,000 for mountain or non-urbanized areas.

Table 2. Landsat 8 band designations

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 – Coastal aerosol	0.43 – 0.45	30
Band 2 – Blue	0.45 – 0.51	30
Band 3 – Green	0.53 – 0.59	30
Band 4 – Red	0.64 – 0.67	30
Band 5 – Near Infrared (NIR)	0.85 – 0.88	30
Band 6 – SWIR 1	1.57 – 1.65	30
Band 7 – SWIR 2	2.11 – 2.29	30
Band 8 – Panchromatic	0.50 – 0.68	15
Band 9 – Cirrus	1.36 – 1.38	30
Band 10 – Thermal Infrared (TIRS) 1	10.60 – 11.19	100
Band 11 – Thermal Infrared (TIRS) 2	11.5 – 12.51	100

Source: (USGS, 2015, p. 9)

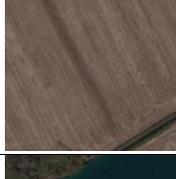
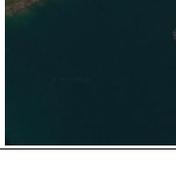
The study area is the Metropolitan City of Milan (CMM), which includes the City of Milan and other 133 municipalities. It covers a surface area of about 1,575 km² and has a population of around 3.254 million inhabitants. According to the National Plan of Adaptation to the Climate Change (PNACC), the CMM is one the most vulnerable areas in Italy in respect to the risk of extreme heatwaves.

3. Methodology

3.1 Creating the LCZ classification

The proposed methodology for improving the quality of the LCZ classification has four main steps. Firstly, Landsat 8 data were downloaded from the United States Geological Survey (USGS), pre-processed (i.e. calibrated and atmospherically corrected), and clipped to the ROI. Then, training areas were created in Google Earth for each LCZ type (Table 3), as instructed by WUDPAT. However, we have digitized an average smaller-sized training samples (down to 250 m in horizontal length) for mapping a fine-scale LCZ map as well as to consider the effect of the small-scale urban units on the micro- and local climate (e.g. extremely big buildings, elongated or wide urban canyons, and urban blocks). In fact, urban blocks or urban units of about 250 m in horizontal scale are commonly the smallest urban units where homogeneity can be found and subsequently can be regarded as climate zones (Kotharkar & Bagade, 2018; Oke et al., 2017).

Table 3. Training samples, defined using Google Earth imagery, for the LCZ classification in the CMM

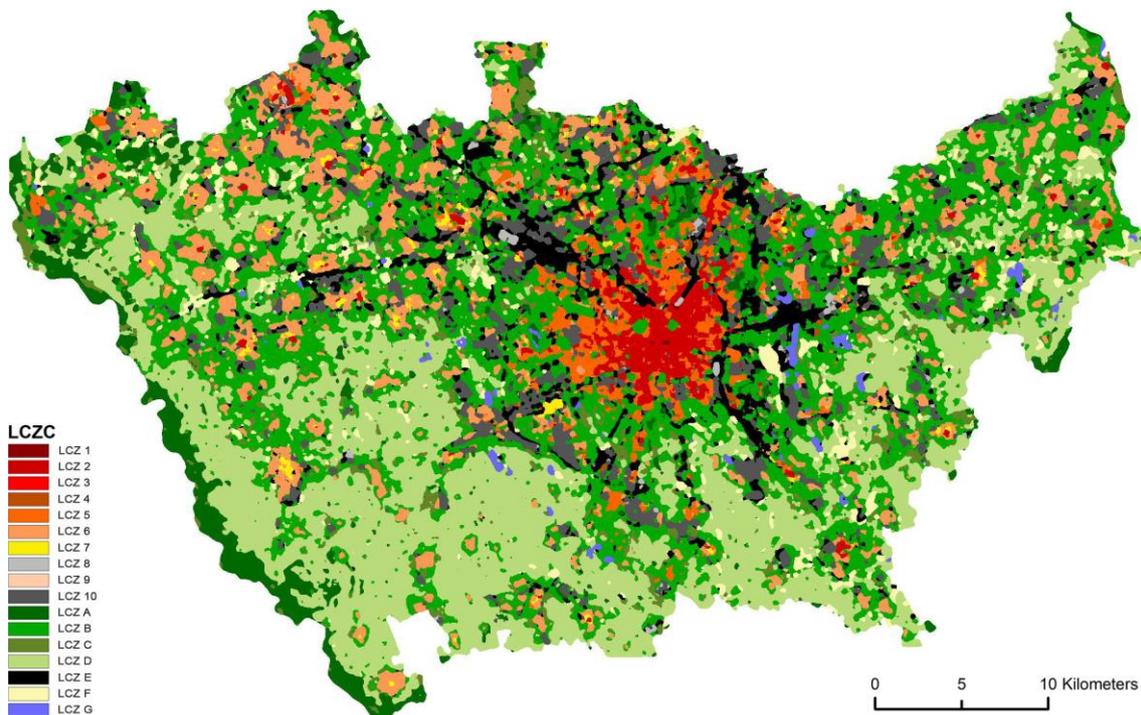
Built types		Land cover types	
LCZ 1 Compact high-rise 	LCZ 8 Large low-rise 	LCZ A Dense trees 	
LCZ 2 Compact mid-rise 	LCZ 9 Sparsely built 	LCZ B Scattered trees 	
LCZ 3 Compact low-rise 	LCZ 10 Heavy industry 	LCZ C Bush, scrub 	
LCZ 4 Open high-rise 		LCZ D Low plants 	
LCZ 5 Open mid-rise 		LCZ E Bare rock or paved 	
LCZ 6 Open low-rise 		LCZ F Bare soil or sand 	
LCZ 7 Lightweight low-rise 		LCZ G Water 	

Source: Authors based on (Stewart & Oke, 2012)

Next, the physical and morphological indicators were calculated. In particular, the broadband albedo was calculated from Landsat 8, bands 2,4,5,6,7, using a conversion formulae (Liang, 2001), while the NDVI was obtained by using Landsat 8, NIR and Red bands (Rouse, Hass, Schell, Deering, & Harlan, 1974). On the other hand, building heights and SVF were calculated using the DTDB and resampled consistently with the spatial resolution of Landsat 8 data (30 m).

Finally, a LCZ map was generated using the Random Forest classifier (Ho, 1998) in SAGA GIS (the Local Climate Zone Classification tool) with a number of decision trees equals 64, as recommended by Oshiro, Perez, and Baranauskas (2012). In fact, we created a LCZ map twice, first, as instructed by WUDAPT, i.e. using Landsat 8 multispectral and thermal bands, and, then, by using Landsat 8 multispectral bands only (physical features) combined with the morphological indicators. A majority filter of 4-pixel radius was applied to the output maps to achieve more homogenous LCZs (of about 270 m in horizontal length scale; Figure 1).

Figure 1. LCZ post-filtered classification (majority filter, 4-pixel radius) for the CMM, using Landsat 8 multispectral bands combined with morphological indicators



Source: Authors using the Local Climate Zone Classification tool in SAGA GIS

3.2 Accuracy assessment

Assessing the level of accuracy of the LCZ classification has been always challenging, since there is not enough, independent test data (ground truth). Furthermore, using the same training data for model validation and calibration is misleading. Accordingly, for instance, in the WUDAPT level 0 method, three approaches are usually used to assess the accuracy of the LCZ maps, before they are disseminated on the online platform; these are cross-validation (rotation estimation), manual review, and cross-comparison with other data (Bechtel et al., 2019). In this paper, we used the repeated holdout cross-validation approach, using different subsamples for 25 iterations as recommended by (Bechtel et al., 2019).

In particular, in the repeated holdout method, the original sample data are separated into two portions, i.e. for training and testing (we used half of the original training data for learning and the other half for testing), where, for each iteration, a different random subset of the data is used. Subsequently, standard accuracy measures (Bechtel et al., 2017) were calculated for each run of the 25 iterations, using a confusion matrix. In particular, we considered four



measures: the overall classification accuracy (OA) for all polygons, the Kappa coefficient (Cohen, 1960), the overall accuracy on built polygons only, and the producer accuracy calculated for each LCZ type. A further certainty measure (Bechtel et al., 2019) was used, by calculating how often each pixel was assigned to the most frequent LCZ type (as obtained by calculating the majority value of the 25 iterations).

4. Results

The LCZ map of the Metropolitan City of Milan, as obtained by the developed physic-morphological model (see Figure 1), shows that the built classes cover around 24.5% of the total metropolitan area, compared to 75.5% of the non-built classes. In particular, the majority of classes are in the LCZ types B, D, 6, 10, E, and A. However, considering only the most urbanized area, i.e. the City of Milan, it is noted that the urban structure is mostly occupied by the open and the compact mid-rise LCZ built types (i.e. 5 and 2) and the land cover types of scattered trees and bare rock or paved (i.e. B and E).

Table 4 and Figure 2 show the results of the accuracy assessment, based on four standard measures. In general, both the WUDAPT level 0 method and the developed physic-morphological model achieved an accuracy above 50%. However, it is noticeable that there is an overall improvement in the classification output, using the developed model, with an OA of 67% compared to 55% using the WUDAPT method.

Similarly, the Kappa coefficient for all the testing polygons is 0.61, using the physic-morphological model, which implies a good agreement (McHugh, 2012), compared to a moderate agreement of 0.47 for the WUDAPT level 0 product. Also, considering only the quality of the built classes, where misclassification mostly occurs, the OA on all the built polygons is 48% compared to 33% using the WUDAPT method; however, this can be improved to 54% and 38% respectively, by excluding the least representative classes (i.e. LCZ 3 and 9).

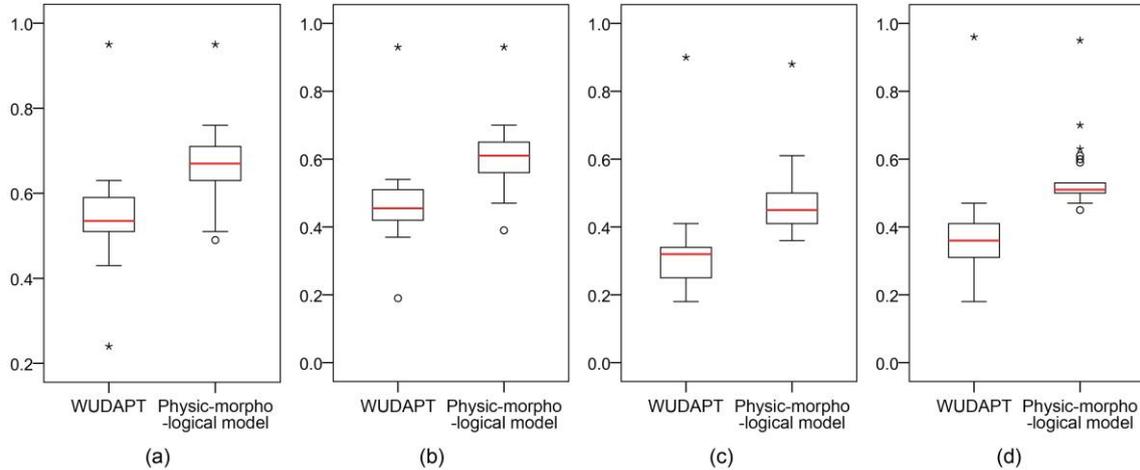
Likewise, the output classification of the new model has a higher certainty of the results over the entire built domain (0.68) compared to that in WUDAPT (0.58); however, both the output maps have almost the same certainty over the non-built classes (0.65), as shown in Figure 3. In general, according to (Bechtel et al., 2019), results are considered of acceptable quality if they achieve a minimum average accuracy of 50%.

Table 4. Accuracy measures

Measure (average of 25 iterations)	WUDAPT level 0 product	The physic-morphological model
OA for all testing polygons	55%	67%
Kappa coefficient for all testing polygons	0.47	0.61
OA of all built types	33%	48%
OA of all built types (excluding LCZ 3 and 9)	38%	54%

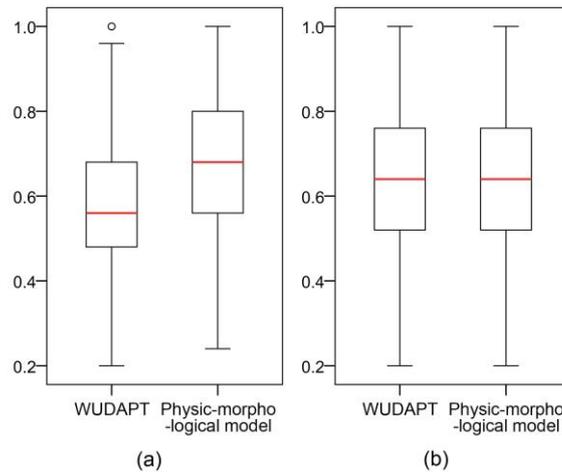
Source: Authors

Figure 2. **Boxplots of the distribution of the accuracy measures across the 25 iterations. (a) Overall accuracy for all testing polygons; (b) Kappa coefficient for all testing polygons; (c) Overall accuracy of all built polygons; (d) Overall accuracy of all built polygons (excluding LCZ 3 and 9)**



Source: Authors

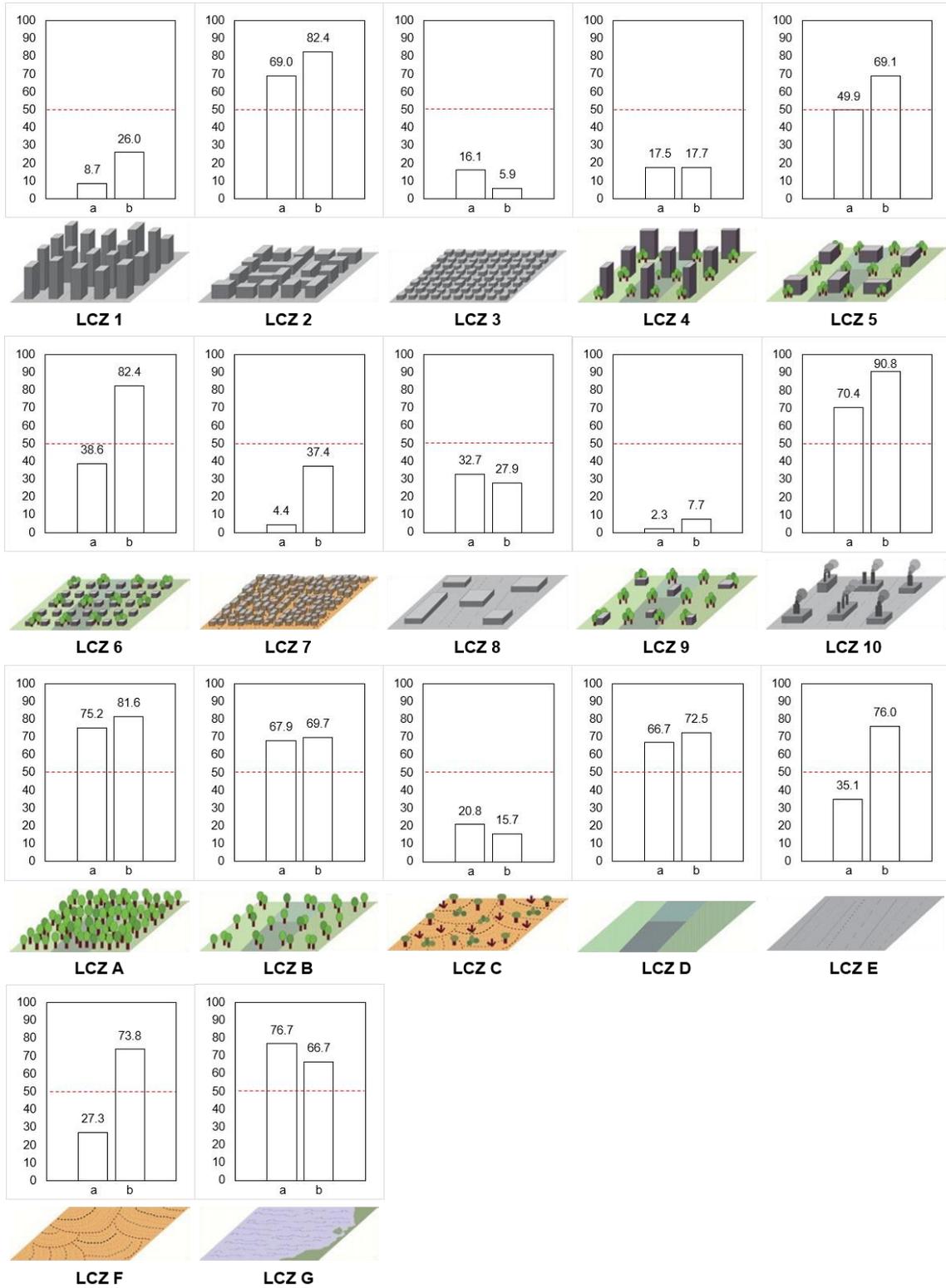
Figure 3. **Boxplots for the certainty of the results over the entire domain. (a) Built local climate zones; (b) Non-built local climate zones**



Source: Authors

Further, in order to assess the impact of the proposed methodology on the accuracy of the individual LCZs, we have calculated the producer accuracy (%) for each LCZ type (Figure 4). The producer accuracy is the likelihood that a pixel in a certain class was classified correctly. In sum, the results show that the physic-morphological model can achieve remarkable improvements in seven LCZ types (i.e. LCZ 1, 5, 6, 7, 10, E, and F) and slight improvement in other four types (i.e. LCZ 2, 9, A, and D). However, almost no or very little improvement was achieved in the LCZs 4 and B. It was also reported that the WUDAPT level 0 product is performing slightly better in four LCZ types (i.e. LCZ 3, 8, C, and G); this can be returned to the presence of the thermal information that can provide further differentiation between the classes, since the surface permeability is negatively well correlated with land surface temperature (Bechtel et al., 2019; Weng, 2009; see also Weng, Lu, & Schubring, 2004).

Figure 4. The producer accuracy (%) for each LCZ type. (a) The WUDAPT level 0 product; (b) The physic-morphological model



Source: Authors



5. Discussion and conclusions

In this paper, we have investigated a methodology for improving the Local Climate Zone (LCZ) automatic classification, in the framework of the World Urban Database and Access Portal Tools (WUDAPT) initiative. In particular, we have utilized medium-resolution multispectral satellite imagery (30 m), combining physical and morphological features of the urban environment, with the aim of conducting a pixel-based image classification. Further, we have digitized relatively small-sized training samples, using very high-resolution (VHR) aerial imagery, to generate a fine-scale LCZ classification, suitable for better studying the urban climate phenomena at the micro- and local scale (e.g. the urban heat island effect).

The results of the accuracy assessment showed a noticeable improvement of the classification output with respect to the WUDAPT level 0 product, for both the overall accuracy (OA) and the Kappa coefficient by 12% and 0.14 respectively. It was also found that the OA, considering only the built classes, is higher by 15%, which demonstrates the effectiveness of the proposed approach in mapping fine-scale LCZs without additional thermal information. This also ascertains the physic-morphological nature of the LCZs which are intrinsically related to certain temperature regimes and comfort levels (Stewart & Oke, 2012). Hence, one may conclude that the WUDAPT level 0 method is best suitable for mapping coarse-scale LCZs (> 1 km²); however, mapping finer resolution LCZ maps, while maintaining a minimum average accuracy, requires better description of the physical and morphological features of the urban environment.

On the other hand, there are still limitations in identifying some individual LCZ types, where the OA is less than 50% (e.g. LCZ 9, 4, C). This is mainly, either because of the size and the quality of the training samples, that can vary from a user to another and require further expert knowledge (Bechtel et al., 2015; Ren et al., 2016) or because of the non-representativeness of some LCZ types, considering the generalization of the original LCZ scheme. Therefore, it would be of interest, in future work, to evaluate the effect of using different sets of training samples (e.g. in terms of area, perimeter, quantity), on the OA of the output maps.

Also, multi-seasonal satellite imagery and higher quality elevation models (e.g. LIDAR derived digital surface model) could be utilized for better reliability of the results, especially when classifying land cover types. Moreover, considering the universality of the LCZ classification scheme (Stewart & Oke, 2012; Wang et al., 2018; Wicki & Parlow, 2017) with respect to the heterogeneity of the internal urban structure among cities of different size and location (Bechtel et al., 2015), the possibility of readapting the standard LCZ scheme by describing site-specific LCZ types or subtypes that are related to certain micro- and local climate conditions, should be investigated.

Acknowledgements: The authors would like to thank the WUDAPT team for their effort in making climate-relevant urban data more accessible worldwide. Also, for having set a universal LCZ mapping protocol for cities globally. We would like also to thank the United States Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) of the free Landsat data. Most of the data used in this project have been produced as part of the research project “Verso paesaggi dell’abitare e del lavorare a prova di clima”, which was funded by Fondazione Cariplo (rif. 2018-2411) and coordinated by the Metropolitan City of Milan; hence, we are indebted to these two latter institutions.



Authors' contribution: All authors equally contributed to the idea of the paper and the discussion of its content. Specifically, Ahmed Eldesoky and Nicola Colaninno defined the scientific methodology for the technical and GIS analysis and for the validation procedures; Eugenio Morello supervised the scientific research work. The paper was revised several times by all authors.

Conflict of interest: The authors declare no conflict of interests.

Bibliography

Bechtel, B., Alexander, P., Böhner, J., Ching, J., Conrad, O., Feddema, J., ... Stewart, I. (2015). Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. *ISPRS International Journal of Geo-Information*. <https://doi.org/10.3390/ijgi4010199>

Bechtel, B., Alexander, P. J., Beck, C., Böhner, J., Brousse, O., Ching, J., ... Xu, Y. (2019). Generating WUDAPT Level 0 data – Current status of production and evaluation. *Urban Climate*. <https://doi.org/10.1016/j.uclim.2018.10.001>

Bechtel, B., Demuzere, M., Sismanidis, P., Fenner, D., Brousse, O., Beck, C., ... Verdonck, M.-L. (2017). Quality of Crowdsourced Data on Urban Morphology—The Human Influence Experiment (HUMINEX). *Urban Science*. <https://doi.org/10.3390/urbansci1020015>

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*. <https://doi.org/10.1177/001316446002000104>

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/34.709601>

Kotharkar, R., & Bagade, A. (2018). Local Climate Zone classification for Indian cities: A case study of Nagpur. *Urban Climate*. <https://doi.org/10.1016/j.uclim.2017.03.003>

Liang, S. (2001). Narrowband to broadband conversions of land surface albedo I algorithms. *Remote Sensing of Environment*. [https://doi.org/10.1016/S0034-4257\(00\)00205-4](https://doi.org/10.1016/S0034-4257(00)00205-4)

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*.

Mills, G., Ching, J., See, L., Bechtel, B., Feddema, J., Masson, V., & Steeneveld, G. J. (2015). An Introduction to the WUDAPT project. *Proceedings of the ICUC9. Meteo France*, (February 2016), 6.

Oke, T., Mills, G., Christen, A., & Voogt, J. (2017). *Urban Climates*. Cambridge: Cambridge University Press. doi:10.1017/9781139016476

Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012). How many trees in a random forest? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-31537-4_13



Ren, C., Wang, R., Cai, M., Xu, Y., Zheng, Y., & Ng, E. (2016). The accuracy of LCZ maps generated by the World Urban Database and Access Portal Tools (WUDAPT) Method: A Case Study of Hong Kong. *The Fourth International Conference on Countermeasure to Urban Heat Islands*. <https://doi.org/10.1093/mnras/stv403>

Rouse, J. W., Hass, R. H., Schell, J. A., Deering, D. W., & Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *Final Report, RSC 1978-4, Texas A & M University, College Station, Texas*.

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*. <https://doi.org/10.1175/BAMS-D-11-00019.1>

USGS. (2015). Landsat 8 (L8) Data Users Handbook. *Earth Resources Observation and Science (EROS) Center*.

Wang, R., Ren, C., Xu, Y., Lau, K. K. L., & Shi, Y. (2018). Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: A case study of Hong Kong. *Urban Climate*. <https://doi.org/10.1016/j.uclim.2017.10.001>

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing*. <https://doi.org/10.1016/j.isprsjprs.2009.03.007>

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2003.11.005>

Wicki, A., & Parlow, E. (2017). Attribution of local climate zones using a multitemporal land use/land cover classification scheme. *Journal of Applied Remote Sensing*. <https://doi.org/10.1117/1.jrs.11.026001>