Contributing to WUDAPT: A Local Climate Zone Classification of Two Cities in Ukraine

Olha Danylo, Linda See, Benjamin Bechtel, Dmitry Schepaschenko, and Steffen Fritz

***Abstract*—Local climate zones (LCZs) divide the urban landscape into homogeneous types based on urban structure (i.e., morphology of streets and buildings), urban cover (i.e., permeability of surfaces), construction materials, and human activities (i.e., anthropogenic heat). This classification scheme represents a standardized way of capturing the basic urban form of cities and is currently being applied globally as part of the world urban database and portal tools (WUDAPT) initiative. This paper assesses the transferability of the LCZ concept to two Ukrainian cities, i.e., Kyiv and Lviv, which differ in urban form and topography, and considers three ways to validate and verify this classification scheme. An accuracy of 64% was achieved for Kyiv using an independent validation dataset while a comparison of the LCZ maps with the GlobeLand30 land cover map resulted in a match that was greater than 75% for both cities. There was also good correspondence between the urban classes in the LCZ maps and the urban points of interest in OpenStreetMap (OSM). However, further research is still required to produce a standardized validation protocol that could be used on a regular basis by contributors to WUDAPT to help produce more accurate LCZ maps in the future.**

***Index Terms*—GlobeLand30, Landsat, local climate zones (LCZs), OpenStreetMap (OSM), remote sensing, Ukraine, urban areas.**

I. Introduction

Local climate zones (LCZs) were developed by [1] as a way of dividing cities into different homogenous thermal regimes for the purpose of sitting weather stations, making representative temperature measurements and for providing urban climate models with a range of possible values for different types of model parameters, e.g., sky view factor and building surface fraction. LCZs are also useful for studying the urban heat island (UHI) effect, where increased temperatures are experienced relative to more rural areas [2]. More recently, the LCZ classification scheme has moved beyond its original purpose and is now recognized as a valuable way of characterizing the urban form and function of cities in a standardized way. The LCZ classification system consists of 10 urban classes, which can be characterized by urban structure (i.e., the morphology of the streets and buildings), urban cover (i.e., permeability and vegetation/built fraction), urban fabric (i.e., the materials), and human activity (i.e., anthropogenic heating). The other seven classes within this scheme are pure, natural land cover types such as forest and water. A list of these classes is provided in Table I and more details can be found in [1]. The LCZs are generic enough that they should capture the main types of urban form globally (although this has yet to be fully tested) while providing a culturally neutral framework for characterizing the structure of cities.

The Urban Atlas, which is produced by the European Environment Agency as part of the Copernicus land monitoring program [3], represents a detailed urban classification but it is only available for large cities in European Union member countries. The urban types in the LCZ scheme are also more detailed than the urban fabric classes of the Urban Atlas. No other detailed urban classification exists that has been applied globally. The world urban database and access portal tools (WUDAPT) initiative (http://www.wudapt.org) is working toward the goal of mapping the LCZs of all major cities globally [4], [5].

There is a considerable literature emerging on the use of remote sensing to classify cities according to urban structure types (USTs) [6]–[8], also referred to as urban morphology types [9] and urban structural units [10]. However, as pointed out in [6], most of the previous studies have analyzed only one city with little thought for transferability to other areas. Each has their own classification scheme, which renders multicity comparisons impossible. Moreover, many of the methods use imagery that is not openly available as well as additional data such as building heights and footprints that are difficult to obtain globally. The WUDAPT philosophy is based on the use of data that are freely available and can be processed in a simple workflow using free software for any city in the world. Numerous multispectral, thermal, and morphological features as well as machine learning methods have been tested for discrimination of LCZs [11] and subsequently a workflow based on Landsat imagery and random forest has been developed by [11] and [12] and implemented in SAGA. Single studies have applied the method to cities with different climatic and cultural backgrounds including Khartoum in this Special Issue [13]. However, it has yet to be further tested and validated on other cities than those previously reported, i.e., Dublin, Houston, and Hamburg. Although building heights and building densities differ between the urban classes, it is possible to use spectral differences in urban materials and cover to differentiate urban structure, negating

Manuscript received July 10, 2015; revised March 01, 2016; accepted March 04, 2016. This work was supported in part by the Ernst Mach Grant, Austrian Exchange Service (OeAD-GmbH), Centre for International Cooperation and Mobility (ICM) and in part by the EU-funded ERC CrowdLand Grant 617754.

O. Danylo, L. See, D. Schepaschenko, and S. Fritz are with the International Institute for Applied Systems Analysis, Laxenburg 2361, Austria (e-mail: danylo@iiasa.ac.at; see@iiasa.ac.at; schepd@iiasa.ac.at; fritz@iiasa.ac.at).

B. Bechtel is with the University of Hamburg, Hamburg 20148, Germany (e-mail: benjamin.bechtel@uni-hamburg.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2016.2539977

TABLE I
LCZ CLASSES [1]

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

the need for very high resolution data that are required, e.g., for USTs.

The aim of this paper is to further test this Landsat-based LCZ workflow on two large cities in Ukraine: Kyiv and Lviv, which differ in terms of urban structure and topography. These LCZ maps will provide a contribution to WUDAPT while also considering issues such as transferability of the methodology and independent validation, which has not been addressed previously in [12]. In particular, we use an independent stratified sample as well as additional datasets including OpenStreetMap (OSM) and the GlobeLand30 land cover product to validate the LCZ classification.

II. Study Area

Two cities in Ukraine were chosen: Kyiv and Lviv; their locations are shown in Fig. 1. The choice of locations was based on possessing local knowledge of the urban landscape of these two cities. Local knowledge has been identified by [12] as a critical element in developing an LCZ classification. This is primarily because urban experts know their own cities well and are, therefore, the best placed to create the training areas for the LCZ classification. Validation is also aided by good local knowledge, which is used when comparing the resulting LCZ maps with very high resolution imagery in Google Earth. A brief description of these two cities is provided below.

Kyiv is the capital of Ukraine. This city dates back to at least the ninth century and has long been a city of importance; it had a population similar to Paris by the year 1200 [14]. With a population of around 2.87 million people in 2014 [15], it is the largest city in Ukraine and the eighth largest in Europe [16]. Kyiv is located in the northern part of the country on the Dnipro (or Dnieper) River with an area of around 839 km^2 and an average elevation of 179 m [17]. The river cuts the city into two parts with the center located on the western bank of the river.

Lviv is located in the western part of Ukraine and was founded in the middle of the 13th century [18]. The city is much smaller than Kyiv, with a population of around 730K and an area of 182 km^2. It is the seventh largest city in Ukraine. The city has an average elevation of 289 m, with the highest hill (412 m) on the northern part of the city.

As the capital of Ukraine, Kyiv is six times larger in area than Lviv and is an agglomeration of surrounding satellite urban areas, reflecting a large commuter population, so has quite a different layout compared to Lviv. The street layout of Kyiv is an irregular grid like structure, probably reflecting the Roman influence, whereas Lviv has an irregular street layout, where the main streets follow the original underground water ways [19]. Despite the difference in sizes and populations, the average living area per person is similar [15], [20]. Both cities also have different topographic characteristics, which will affect the local climate. Moreover, their histories are quite different, i.e., Lviv was part of the Austro-Hungarian empire, whereas Kyiv was part of the Russian empire so the urban form, i.e., the building architecture and street layouts, differs.

Both cities have a humid continental climate with cold winters (Köppen–Geiger classification of Dfb). The average high temperature in summer is around 25 °C but extremes of almost 40 °C have been recorded in the past. The cities are subject to UHI effects, but these are exacerbated during extreme events.

III. Materials and Methods

A. Data Inputs

Landsat 8 imagery was downloaded from the US Geological Survey Earth Explorer site (http://earthexplorer.usgs.gov/) for both cities. For Kyiv, four scenes were used with the following dates (April 16, 2013; May 2, 2013; June 6, 2014; October 28, 2014) whereas for Lviv, five were used (May 24, 2014; June 9, 2014; March 8, 2015; March 24, 2015; April 9, 2015). These scenes had cloud cover of less than 4%. Although a fifth scene was downloaded for Kyiv, it resulted in linear artifacts in the LCZ map and was, therefore, omitted. Multiple scenes were downloaded because multitemporal information improves the LCZ classification as found by [12].

The algorithm to create the LCZ classification requires training data. These data should cover homogeneous areas that are as large as possible or at least the minimum size of an LCZ, i.e., around 1 km^2. Fig. 2 shows the training areas, whereas Table II contains details of these training areas, in particular the number of polygons digitized per LCZ and the area covered by the training areas in each city. In some cases, the number of polygons is small since the actual proportion of some LCZs in each city is small. A random stratified sample of 1125 pixels at the original resolution of 120 m was selected from the city of Kyiv. This was used for independent validation of the LCZ map of Kyiv.

To then undertake an independent comparison, two different datasets were used. The first is the GlobeLand30 land cover dataset at a resolution of 30 m that was recently developed by the National Geomatics Center of China [21] for 2010. This land cover dataset is freely available for downloading and contains nine classes including one for artificial surfaces. This latter class covers urban areas, roads, rural cottages, and mines. They used a supervised approach to first classify artificial surfaces followed by the application of a segmentation method. Artificial surfaces were then classified based on exceedance of a minimum threshold of 50% within the identified objects. Finally, manual verification was undertaken using high-resolution imagery from Google Earth. The user's accuracy was estimated at around 87% for this class, whereas the overall accuracy for all classes in this global product is around 80% [17].

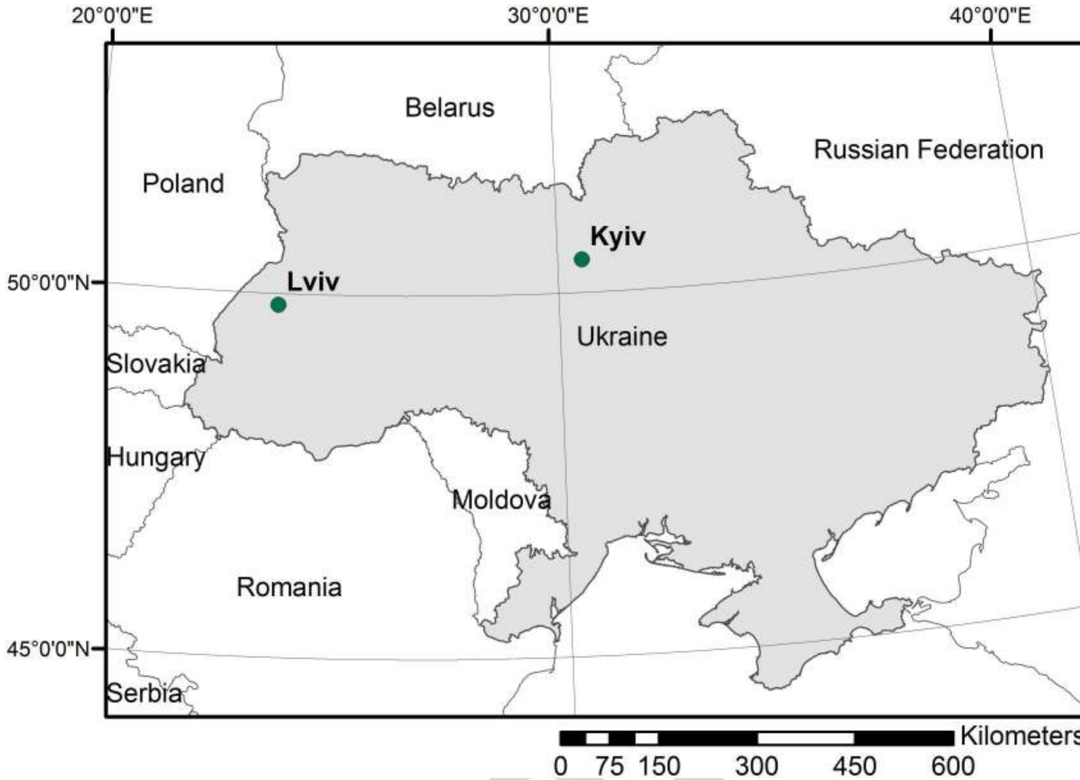

Fig. 1. Location of the cities of Kyiv and Lviv in Ukraine.

The second dataset for independent comparison is from OSM. OSM is a community-based mapping initiative in which volunteers map features such as buildings, roads, land use, and points of interest [22]. The data are openly available through an open database commons open data license and were downloaded from the GeoFabrik website in Germany (http://www.geofabrik.de). The features are organized as polygons, lines, and points. Only the point shapefile was used in this study in which points of interest of type cities, villages, and towns were extracted. These point locations are meant to correspond to the center of these features and will be used as an additional source of independent comparison with the LCZ classification of Kyiv and Lviv. Work is ongoing to investigate how OSM line and polygon features can be used in both LCZ classification and validation in the future.

B. *Methodology for LCZ Classification*

The workflow in [12] was employed to create the LCZ maps for Kyiv and Lviv. A modified version of this workflow is shown in Fig. 3. The Landsat 8 imagery was downloaded and the training areas were created for each city as outlined in Section III-A. The Landsat 8 imagery was then classified using a random forest classifier. Instead of using the SAGA software [23] from [12], the workflow was processed using R. Each Landsat 8 scene contains 11 bands, 8 of which are multispectral (at a resolution of 30 m), 1 is panchromatic (at 15 m resolution), and 2 are thermal (acquired at 100 m resolution, but delivered resampled to 30 m). Despite possible redundancy, all bands were used in the classification since random forest is relatively insensitive toward the number of features. All bands from the five scenes were resampled using the area mean to a common resolution of 120 m, which is within the range of 100–150 m recommended by [12]. Therefore, 48 inputs were provided to the random forest classifier for Kyiv (to include all four scenes) and 60 inputs were used in total for Lviv. Experimentation with the number of trees in the random forest classifier revealed a flattening out of the out of bag error curve at 128 trees (see Table III) so this was used as the final configuration to create the LCZ classifications of the two cities. Each tree in the classifier is constructed using a sample in which around one third of the observations are left out. Once all trees are constructed, the resulting class for a given set of inputs is based on majority voting. The out of bag error is the prediction error based on the trees that did not use a specific sample for training.

The LCZ map was then examined using Google Earth to look for any poorly classified areas. Based on this qualitative inspection, additional training areas were added and the classification was rerun. Using the advice provided in [11], the minimum number of training areas per class suggested was 4–5 (where it was possible to identify this number). Thus, areas with larger number of training areas (Table II) reflect attempts to improve the classification and represent additional training areas. This step is repeated as many times as necessary.

An additional experiment was undertaken in which the minimum, mean, and maximum value of the resampled 120 m bands were provided to the classifier, increasing the number of inputs (or features) from 48 to 144 for Kyiv and from 60 to 180 for Lviv. The idea was to determine whether providing additional information about the spectral variation to the classifier, which

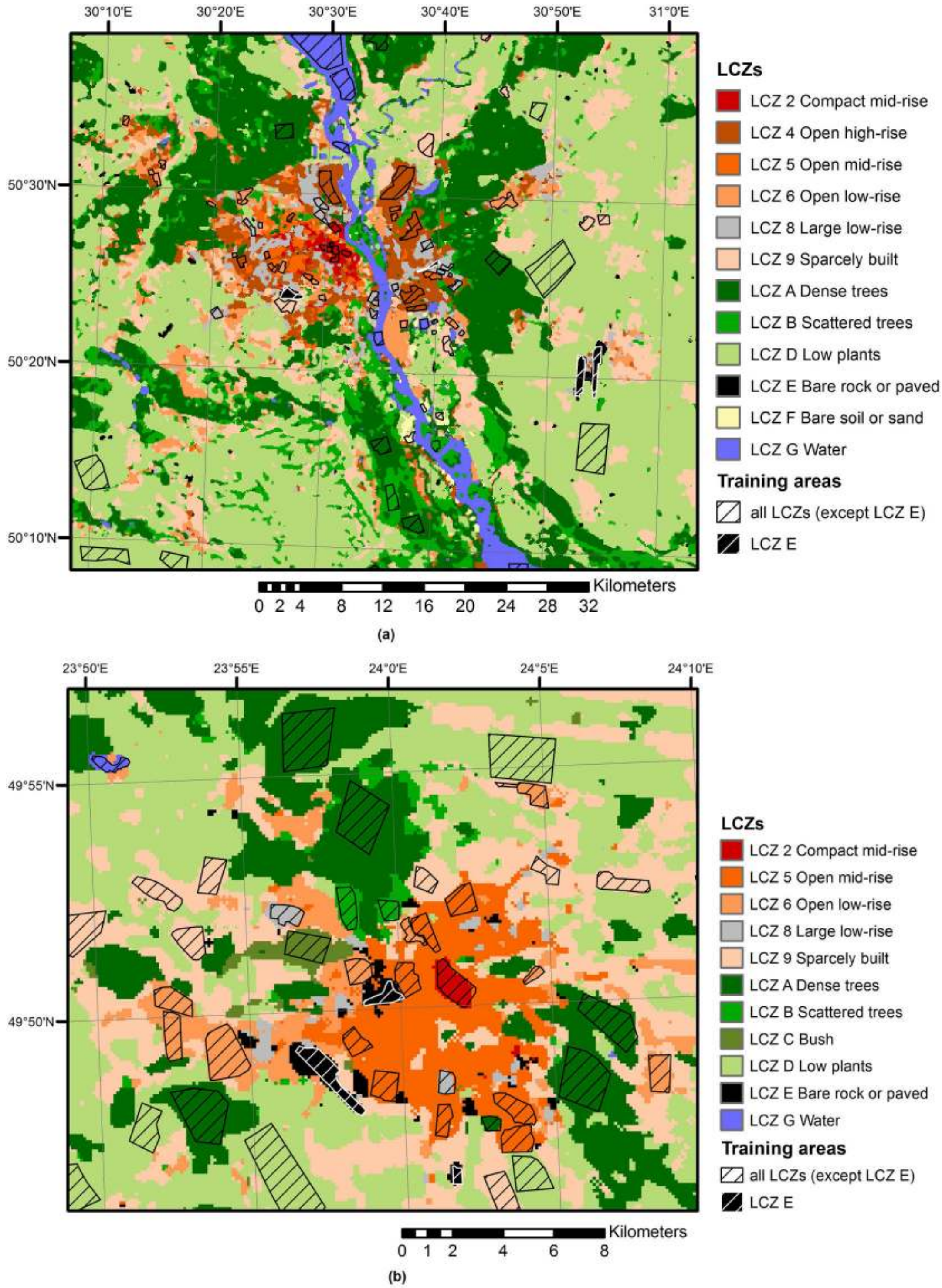

Fig. 2. Training areas in (a) Kyiv and (b) Lviv plotted on the LCZ maps.

would otherwise be lost in the resampling, might help to better discriminate between different LCZs.

Two new steps were then added to the workflow of [12]. The first was to undertake an independent validation using a random stratified sample (Fig. 3 item 1) as described in the section on data inputs. Such an approach has not yet been tried for validation of LCZ maps.

A postclassification filter of a two pixel window was then applied to the image to create more homogeneous LCZs. This is because LCZs are meant to be areas of around 1 km^2 since they must be large enough to have an effect on the local climate.

The second additional step to the workflow (Fig. 3 item 2) was to compare the map with other sources of independent data to determine the agreement. The GlobeLand30 land cover

TABLE II
TRAINING AREAS FOR KYIV AND LVIV

LCZ	Number of training areas		Area (km^2)	
	Kyiv	Lviv	Kyiv	Lviv
LCZ 2 Compact mid-rise	3	1	2.98	1.57
LCZ 4 Open high-rise	6	N/A	17.73	N/A
LCZ 5 Open mid-rise	7	7	3.70	7.65
LCZ 6 Open low-rise	12	8	9.04	8.50
LCZ 8 Large low-rise	16	2	9.66	1.43
LCZ 9 Sparsely built	7	9	10.50	10.50
LCZ A Dense trees	7	6	17.02	17.37
LCZ B Scattered trees	2	2	3.40	1.53
LCZ C Bush, scrub	N/A	1	N/A	1.56
LCZ D Low plants	6	5	53.16	15.11
LCZ E Bare rock or paved	6	3	8.41	2.46
LCZ F Bare soil or sand	4	N/A	1.43	N/A
LCZ G Water	6	1	31.05	0.53

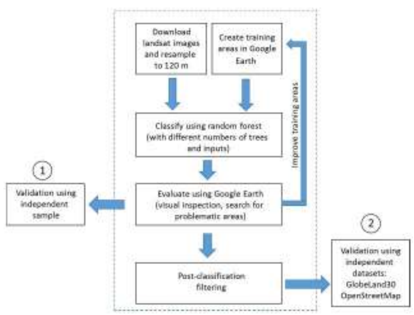

Fig. 3. LCZ workflow. The dotted lines contain the steps as outlined in [11], whereas the validation steps labeled 1 and 2 have been added here.

TABLE III
OUT OF BAG ERROR FOR DIFFERING NUMBERS OF TREES IN THE RANDOM FOREST CLASSIFICATION

Number of trees	Kyiv	Lviv
4	0.085	0.135
8	0.076	0.123
16	0.058	0.088
32	0.049	0.084
64	0.044	0.076
128	0.038	0.070
256	0.037	0.071

product and points of interest from OSM were overlaid onto the LCZ maps and a comparison was made, both visually and via confusion matrices to determine correspondence.

One of the proposed strengths of the LCZ classification is that it is a standardized approach so that it can theoretically be transferred from city to city. As outlined in Section II, Kyiv and Lviv differ in urban form so the LCZ classification can be used to examine these differences objectively. Therefore, official administrative boundaries for each city were applied to the LCZ maps to compare them in terms of what types of LCZs characterize each city and their relative sizes.

IV. RESULTS

A. *LCZ Classifications of Kyiv and Lviv*

Fig. 4 shows the LCZ map of Kyiv, which contains 12 out of the 17 LCZ types. The only compact LCZ is 2 (compact mid-rise) as there are no examples of LCZs 1 and 3 in Kyiv. LCZ 7 (lightweight low-rise) and LCZ 10 (heavy industry) are also not found in Kyiv. The Dnipro River clearly cuts the city in half with most of the urban types concentrated in a core around the river. The business district can clearly be seen on the western bank of the river. The whole western part of the city looks very heterogeneous, without a clear sense of structure. This is also seen very clearly when the city is viewed using Google Earth imagery, which is not shown here due to the size of the area. However, both cities can be viewed via the WUDAPT website (http://www.wudapt.org), which includes Google Earth imagery. This heterogeneity contrasts very sharply with much more organized cities such as those in North America and other parts of Europe. On the eastern side and to the north of the city is LCZ 4 (Open high-rise), which is characterized by large areas of newer residential buildings (i.e., post-1965 and also some post-1987). This part of the city looks more organized and may reflect more recent planning compared to the much older historical center. Areas of light industry are scattered throughout the city (LCZ 8—large low-rise). Around Kyiv is a considerable amount of greenspace (LCZs A, B, and D) with sparsely built settlements (LCZ 9) appearing as small clusters as one moves away from the center of the city. This leap frog development reflects urban satellite developments for a commuting population.

The LCZ classification of Lviv is given in Fig. 5. Like Kyiv, it has the same urban LCZ types although LCZ 4 (open high-rise) is absent. However, apart from a small central patch of LCZ 2 (compact mid-rise), the majority of the center is a large homogenous area of LCZ 5 (open low-rise). Examining photographs from Google street view reveals building architecture that is similar to the older residential part of the city of Vienna, reflecting the Austro-Hungarian history. The city's urban structure is more organized, which is in sharp contrast to the much more heterogeneous mix of LCZs seen in Kyiv. Areas of sparsely built settlements (LCZ 9) are also much larger and closer to the city center.

Table IV provides a comparison of the size of the LCZs in Kyiv and Lviv after official city boundaries were used to clip the LCZ maps. In absolute terms Kyiv is clearly much bigger, but when compared relatively, Lviv has more than 60% of urban LCZs compared to Kyiv, which has just under 40%. While Kyiv has almost 15% of its area covered by LCZ 4 (open high-rise), which is absent in Lviv, LCZ 5 (open mid-rise) is much more prevalent in Lviv than Kyiv. Lviv has a higher amount of LCZ 9 (sparsely built), which may reflect agricultural areas surrounding the city, whereas there are considerably more forested areas around the city of Kyiv. Water is also higher in Kyiv, reflecting the river that runs through the city.

The confusion matrix for the training data for Kyiv is shown in Table V, where the out of bag error was 3.82%. Table VI shows the results when using the additional inputs from the minimum and maximum values of the bands in addition to the mean. The out of bag error improves marginally to 3.5%. The overall accuracy is 96%, increasing slightly with the additional inputs to 97%. The natural classes are all captured extremely well with good results for the urban classes. However, there is some confusion between the compact and open urban classes. When considering all the inputs (Table VI),

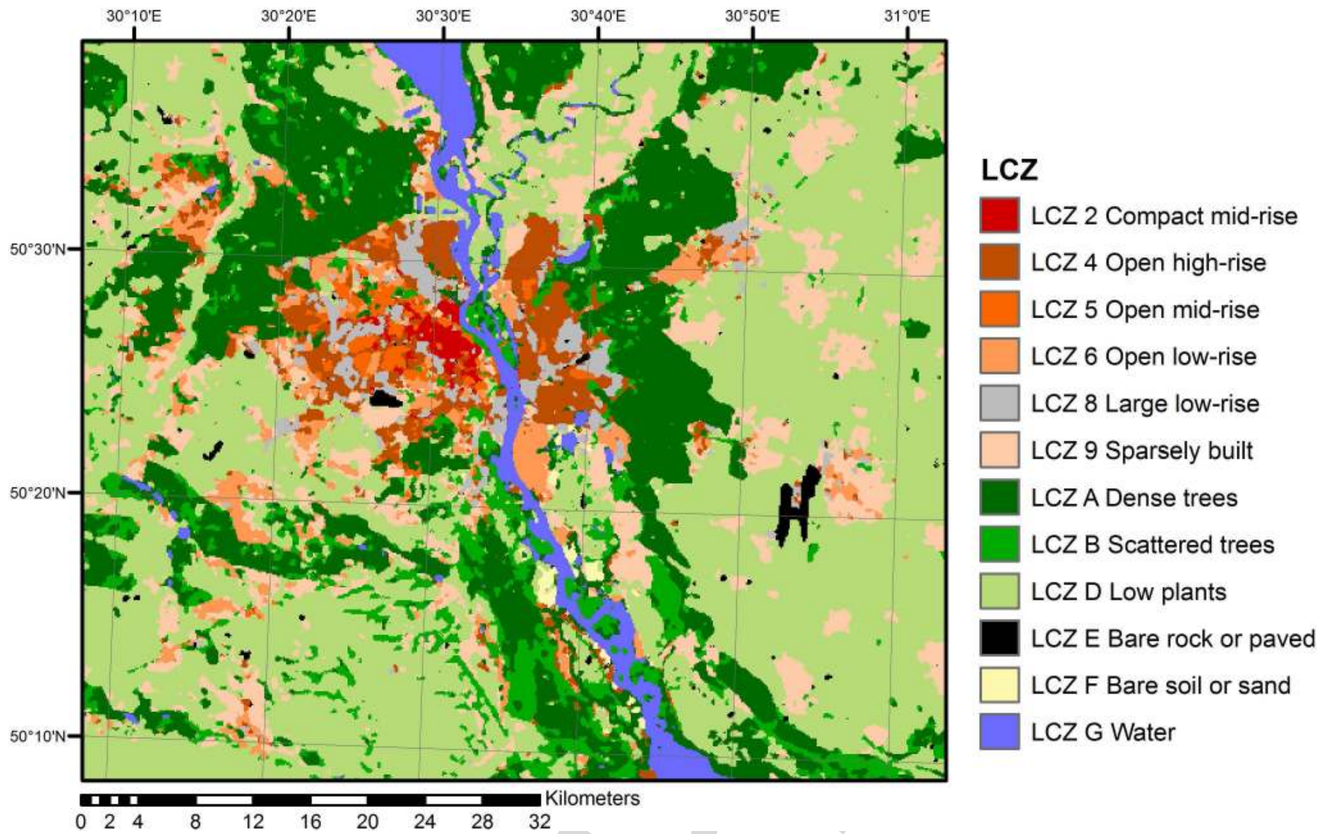

Fig. 4. LCZ map of Kyiv.

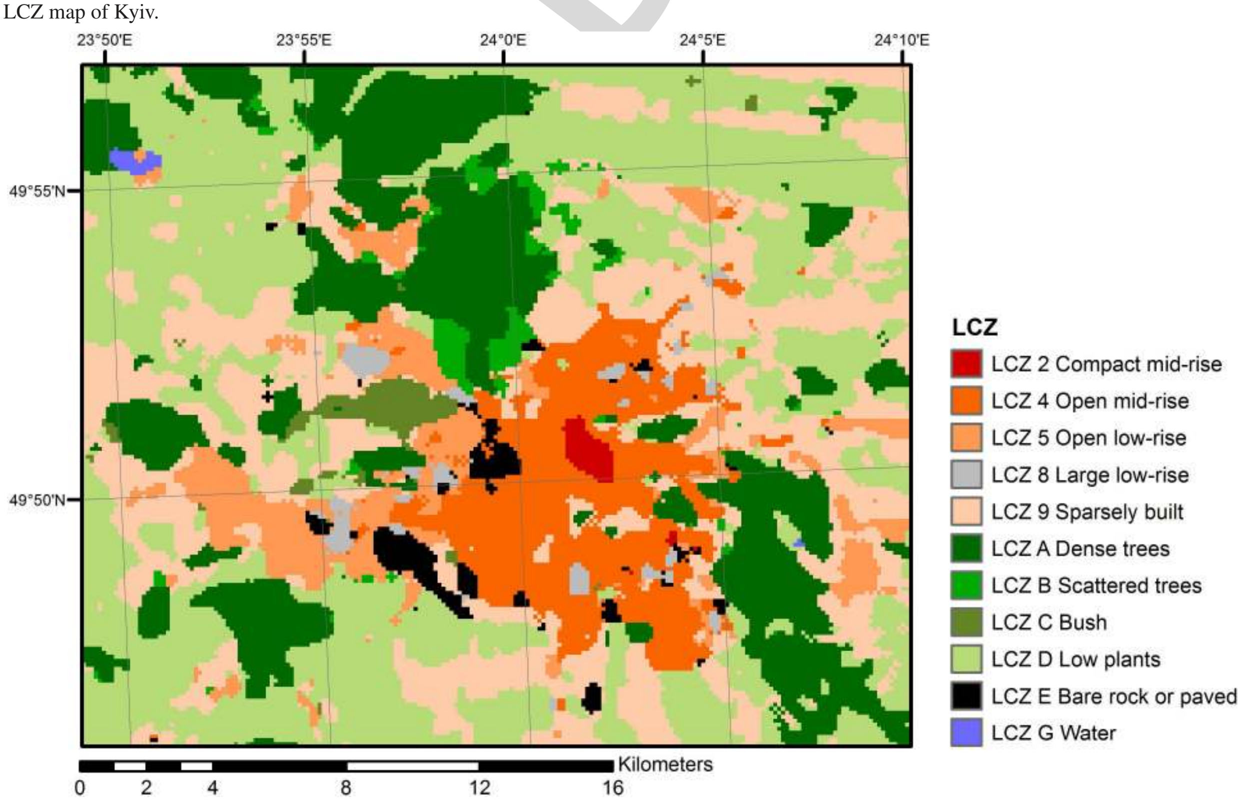

Fig. 5. LCZ map of Lviv.

LCZs 4, 5, and 6 decrease in accuracy slightly but there is less confusion between LCZs 4 and 8. There are other small tradeoffs that can be observed when comparing Tables V and VI. However, there appears to be very small differences between the results with and without the additional inputs. The results for Lviv are similar to Kyiv. The out of bag error is slightly larger at 7% but the confusion matrix shows similar patterns.

TABLE IV
AREAS OF LCZS FOR KYIV AND LVIV CONTAINED WITHIN THE OFFICIAL CITY BOUNDARIES

LCZ	Area (km^2)		Area (%)	
	Kyiv	Lviv	Kyiv	Lviv
LCZ 2 Compact mid-rise	18.32	2.03	2.22	1.34
LCZ 4 Open high-rise	117.66	N/A	14.28	N/A
LCZ 5 Open mid-rise	35.64	48.20	4.32	31.82
LCZ 6 Open low-rise	48.56	13.55	5.89	8.95
LCZ 8 Large low-rise	66.61	5.43	8.08	3.58
LCZ 9 Sparsely built	34.52	25.40	4.19	16.77
LCZ A Dense trees	332.81	32.47	40.38	21.44
LCZ B Scattered trees	71.74	4.18	8.70	2.76
LCZ C Bush, scrub	N/A	5.00	N/A	3.30
LCZ D Low plants	52.26	8.15	6.34	5.38
LCZ E Bare rock or paved	3.33	7.06	0.40	4.66
LCZ F Bare soil or sand	3.01	N/A	0.37	N/A
LCZ G Water	39.73	N/A	4.82	N/A

B. Validation With Sample Data

The sample validation dataset described in Section III-A was used to assess the accuracy of the LCZ maps. Tables VII and VIII provide confusion matrices for Kyiv for the two different input datasets. Table VII contains results for the random forest classified with only the resampled mean of the bands as inputs while Table VIII shows the results when the minimum, mean, and maximum are included. The overall accuracy using the mean as inputs is 64%, where the poorest class is LCZ 4 (open high-rise). There is some confusion between LCZ 4 and other urban classes and LCZ E (bare rock or paved), and there are issues with LCZ 5 (open mid-rise), which is also mistaken for other classes. The overall accuracy improves slightly to 66% when including more inputs, where the user accuracy of some urban classes improves but the tradeoff is a slight decrease in the producer's accuracy. Although the effects of adding additional inputs is more pronounced on the independent validation dataset compared to the training data, it appears that there is little to be gained from adding these extra inputs to the classifier. Kyiv is very heterogeneous, particularly in the western part of the city, which may partly explain these accuracy figures. Further training data may be needed to improve the classification.

C. Comparison With GlobeLand30

Figs. 6 and 7 show the GlobeLand30 land cover map superimposed on the LCZ maps of Kyiv and Lviv, respectively. For Kyiv, the artificial surfaces appear to match the urban types extremely well from a visual point of view, including LCZ 9 (sparsely built) that covers scattered settlements around Kyiv. Large, homogeneous patches of forest cover and water are also captured well as are grassland and cultivated areas (corresponding to LCZ D low plants). However, there are some exceptions, e.g., Fig. 8(a) shows an area on high-resolution imagery from Google Earth where GlobeLand30 classifies the area as Forest and the area is LCZ D (low plants). The image contains a flood plain, which becomes inundated during flooding and is, therefore, left in a natural state. Thus, the LCZ map better captures this area than the GlobeLand30 product.

This overall correspondence is confirmed in Table IX, which contains a confusion matrix comparing the LCZ classification with the GlobeLand30 land cover product. The LCZs were first mapped onto the GlobeLand30 classes as follows.

1) Urban LCZs and LCZ E (since this latter one is an OR class of bare rock or paved) map onto artificial surfaces.
2) LCZs A and B map onto the Forest class.
3) LCZ C maps onto shrubland.
4) LCZ D maps onto cultivated land and grassland which were collapsed into a single class in the confusion matrix.
5) LCZ F maps onto bare soil or sand.
6) LCZ G maps onto water bodies.

There is no wetland class in the LCZ classification, and classes that are related to the tundra and snow were omitted. LCZ9 (sparsely built) could be either artificial surfaces, grassland or cultivated land. For the purpose of calculating correspondence between the two datasets, LCZ9 is mapped onto the GlobeLand30 class artificial surfaces.

Table IX shows that the overall correspondence between the two datasets was 83% for Kyiv. The user's and producer's accuracies were generally high except for classes that were simply not present (e.g., shrubland) or where there is no corresponding class (e.g., wetland).

For Lviv (Fig. 7), the visual comparison shows similar correspondence between the artificial surfaces class of GlobeLand30 and the urban types, with the exception of LCZ 9 (sparsely built), which often corresponds to the cultivated land class of GlobeLand30. This is not surprising as this class contains less than 20% artificial surfaces but still is considered an urban type in the LCZ classification. Correspondence with forests is also reasonably good although there are exceptions. For example, Fig. 8(b) shows an area on high-resolution imagery from Google Earth where GlobeLand30 classifies the area as artificial surfaces while the LCZ classification indicates LCZ B (scattered trees). The image clearly shows scattered houses but not an artificial surface fraction of greater than 50%. Although there are scattered trees, this could also be an example of LCZ 9 (sparsely built), in which case both maps would be wrong. Moreover, one large area of LCZ C (bush, scrub) has been classified as cultivated land in the GlobeLand30 product. However, it was difficult to tell from Google, Earth which one is actually correct. Thus, Google street view photographs were examined in this area and they revealed the presence of shrubs.

Table X contains the correspondence between the two products, which shows the overall agreement is at 75% and thus somewhat lower than for Kyiv. Table X shows that there is some confusion between water bodies, forest, and cultivated areas/grassland, whereas the highest agreement is for the urban LCZs.

D. Comparison With OSM

Figs. 9 and 10 show the city, towns, and villages from OSM overlaid on top of the LCZ maps of Kyiv and Lviv, respectively. A visual inspection shows that the OSM feature called city (which is single point of interest) falls in LCZ 2, which is to be expected as this is the business center of each city. The towns and villages also generally fall in urban classes as expected. Table XI summarizes the correspondence between the LCZs

T5:1 TABLE V
T5:2 CONFUSION MATRIX FOR KYIV USING THE MEAN AS INPUTS

LCZ	2	4	5	6	8	9	A	B	D	E	F	G	Sum	UA
2	231	40	10	4	21	0	0	0	0	0	0	0	306	0.75
4	11	1842	16	10	82	4	0	0	0	2	0	0	1967	0.94
5	13	106	218	25	4	6	0	1	0	0	0	0	373	0.58
6	3	24	3	894	3	23	0	0	1	0	0	0	951	0.94
8	2	142	1	6	799	0	0	1	1	8	1	0	961	0.83
9	1	4	0	31	0	1072	0	0	30	0	0	0	1138	0.94
A	0	0	0	0	0	1	1885	2	2	0	0	0	1890	1.00
B	0	0	0	0	0	0	6	371	5	0	0	0	382	0.97
D	0	1	0	0	0	13	0	1	5804	1	0	0	5820	1.00
E	0	4	0	0	9	4	0	0	14	908	0	0	939	0.97
F	0	0	0	0	2	0	0	0	1	0	143	0	146	0.98
G	0	0	0	0	0	0	0	0	0	0	0	3590	3590	1.00
Sum	261	2163	248	970	920	1123	1891	376	5858	919	144	3590		
PA	0.89	0.85	0.88	0.92	0.87	0.95	1.00	0.99	0.99	0.99	0.99	1.00	OA	0.96

Columns contain the training data while rows contain the results from the LCZ map.

T6:1 TABLE VI
T6:2 CONFUSION MATRIX FOR KYIV USING THE MINIMUM, MEAN, AND MAXIMUM AS INPUTS

LCZ	2	4	5	6	8	9	A	B	D	E	F	G	Sum	UA
2	239	36	7	4	20	0	0	0	0	0	0	0	306	0.78
4	11	1833	18	10	88	5	0	0	0	2	0	0	1967	0.93
5	13	114	210	22	7	7	0	0	0	0	0	0	373	0.56
6	4	17	11	890	5	24	0	0	0	0	0	0	951	0.94
8	8	86	4	6	850	1	0	0	0	4	2	0	961	0.88
9	0	6	2	21	0	1080	0	0	29	0	0	0	1138	0.95
A	0	0	0	0	0	0	1886	1	3	0	0	0	1890	1.00
B	0	0	0	0	0	0	6	372	4	0	0	0	382	0.97
D	0	0	0	0	0	4	0	0	5816	0	0	0	5820	1.00
E	0	1	0	0	11	3	0	0	16	908	0	0	939	0.97
F	0	1	0	0	1	0	0	0	0	1	143	0	146	0.98
G	0	0	0	0	0	0	0	0	0	0	0	3590	3590	1.00
Sum	275	2094	252	953	982	1124	1892	373	5868	915	145	3590		
PA	0.87	0.88	0.83	0.93	0.87	0.96	1.00	1.00	0.99	0.99	0.99	1.00	OA	0.97

Columns contain the training data while rows contain the results from the LCZ map.

T7:1 TABLE VII
T7:2 CONFUSION MATRIX FOR KYIV USING THE SAMPLE VALIDATION DATASET AND THE MEAN AS INPUTS

LCZ	2	4	5	6	8	9	A	B	D	E	F	G	Sum	UA
2	**34**	15	1	3	3	0	0	0	0	0	0	1	57	0.60
4	7	**37**	6	0	6	1	0	0	0	0	0	0	57	0.65
5	5	29	**20**	13	3	2	2	0	0	0	0	0	74	0.27
6	0	9	1	**44**	1	14	0	1	0	0	0	0	70	0.63
8	6	39	2	5	**70**	6	1	0	1	5	2	0	137	0.51
9	0	4	0	13	1	**56**	0	1	15	0	0	0	90	0.62
A	0	2	1	3	0	0	**111**	8	3	0	0	1	129	0.86
B	0	3	4	3	1	7	35	**31**	11	0	0	0	95	0.33
D	0	0	0	0	2	17	5	3	**150**	4	0	0	181	0.83
E	1	10	1	0	7	2	0	1	4	**11**	0	0	37	0.30
F	0	2	0	0	2	2	1	0	24	3	**58**	0	92	0.63
G	1	1	0	1	0	0	6	0	2	0	0	**95**	106	0.90
Sum	54	151	36	85	96	107	161	45	210	23	60	97	1125	
PA	0.63	0.25	0.56	0.52	0.73	0.52	0.69	0.69	0.71	0.48	0.97	0.98	OA	0.64

Columns contain the validation data while rows contain the results from the LCZ map.

TABLE VIII
CONFUSION MATRIX FOR KYIV USING THE SAMPLE VALIDATION DATASET USING THE MINIMUM, MEAN, AND MAXIMUM AS INPUTS

LCZ	2	4	5	6	8	9	A	B	D	E	F	G	Sum	UA
2	**40**	14	0	0	2	0	0	0	0	0	0	1	57	0.70
4	4	**43**	4	0	5	1	0	0	0	0	0	0	57	0.75
5	4	31	**23**	7	4	3	2	0	0	0	0	0	74	0.31
6	0	8	2	**46**	1	11	0	1	0	1	0	0	70	0.66
8	9	30	1	5	**78**	5	0	0	3	4	2	0	137	0.57
9	0	4	1	9	2	**56**	0	1	17	0	0	0	90	0.62
A	0	4	1	0	0	0	**112**	7	4	0	0	1	129	0.87
B	0	3	5	2	1	7	33	**33**	11	0	0	0	95	0.35
D	0	0	0	1	1	19	4	2	**146**	7	0	1	181	0.81
E	1	8	1	0	8	2	0	1	4	**12**	0	0	37	0.32
F	0	0	0	0	3	5	1	0	22	0	**61**	0	92	0.66
G	1	0	1	1	0	0	2	1	1	0	1	**98**	106	0.92
Sum	59	145	39	71	105	109	154	46	208	24	64	101	**1125**	
PA	0.68	0.30	0.59	0.65	0.74	0.51	0.73	0.72	0.70	0.50	0.95	0.97	OA	0.66

Columns contain the validation data while rows contain the results from the LCZ map.

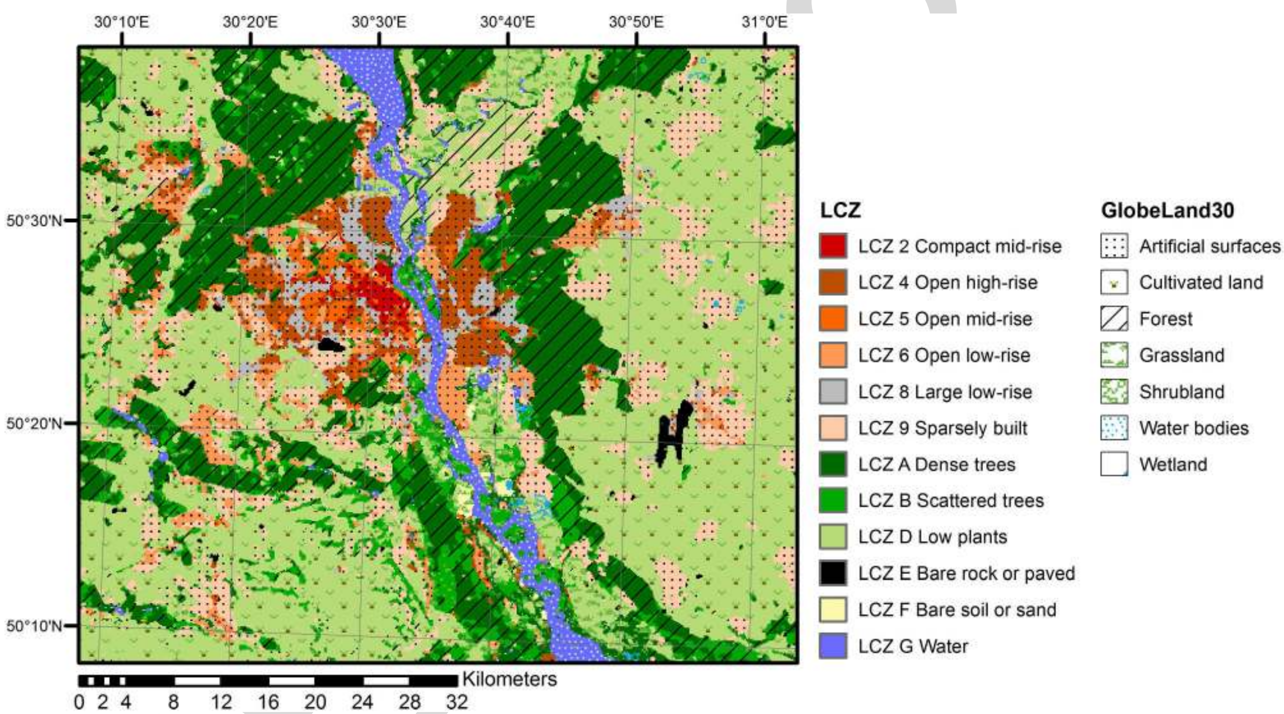

Fig. 6. LCZ map of Kyiv compared with the GlobeLand30 land cover product.

and the city, towns, and villages. In the case of Kyiv, all towns fall in urban classes or LCZ E (bare rock or paved), whereas one town in Lviv falls in LCZ A (dense trees), indicating a misclassification. For villages in Kyiv, 6 out of 136 locations fall in nonurban classes (roughly 4%) while all villages in Lviv fall in urban classes or LCZ E (bare rock or paved). Thus, the results show a good correspondence between the points of interest for the city, towns, and villages and the LCZ classification.

V. DISCUSSION

The LCZ methodology is simple to implement using freely available satellite imagery and software, as per the original goal of WUDAPT [12]. The results also illustrate that the LCZ classification provides a standardized way of mapping and comparing cities. Although Kyiv and Lviv have similarities due to their geographical proximity, they are also quite different cities in terms of size, topography, and urban form. The LCZ classification provides a way of clearly visualizing and quantifying these differences in a standardized, transferable manner. However, there are challenges in working with small cities such as Lviv. For example, finding sufficient training areas of a large enough size was much more difficult for Lviv than Kyiv.

Since the random forest classifier provides an out of bag error, there is theoretically no need for an additional test dataset.

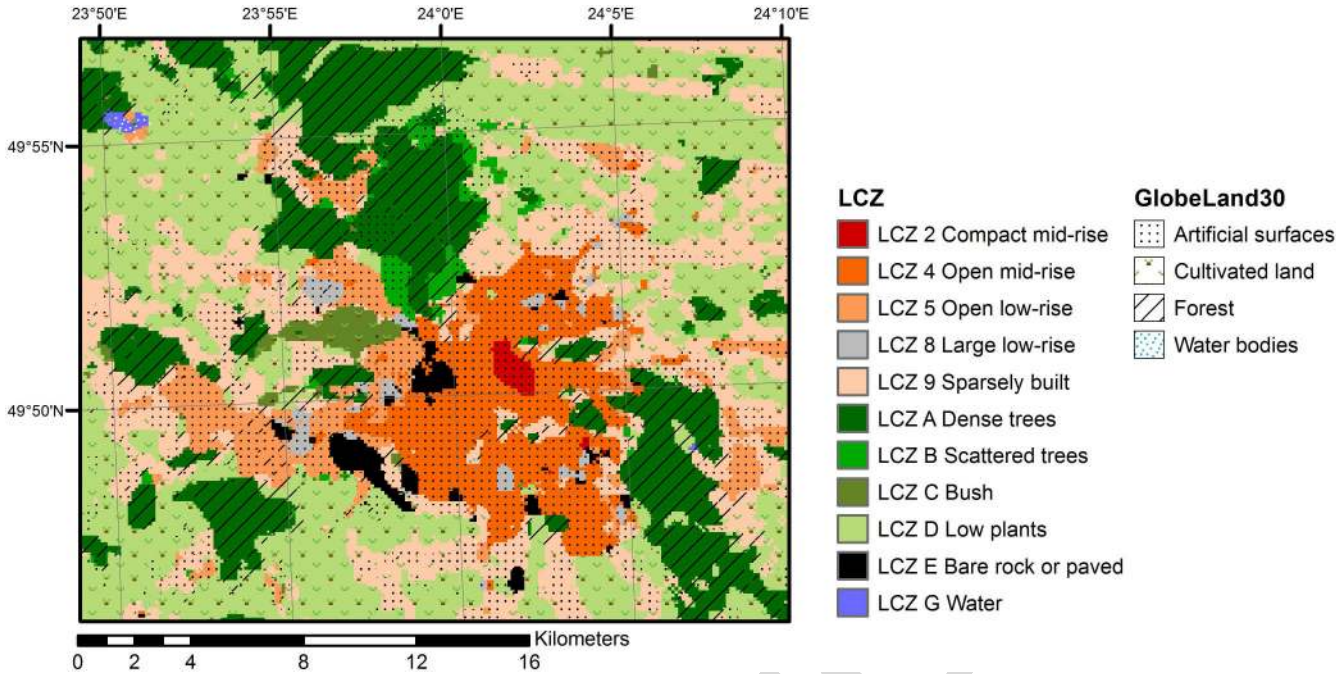

Fig. 7. LCZ map of Lviv compared with the GlobeLand30 land cover product.

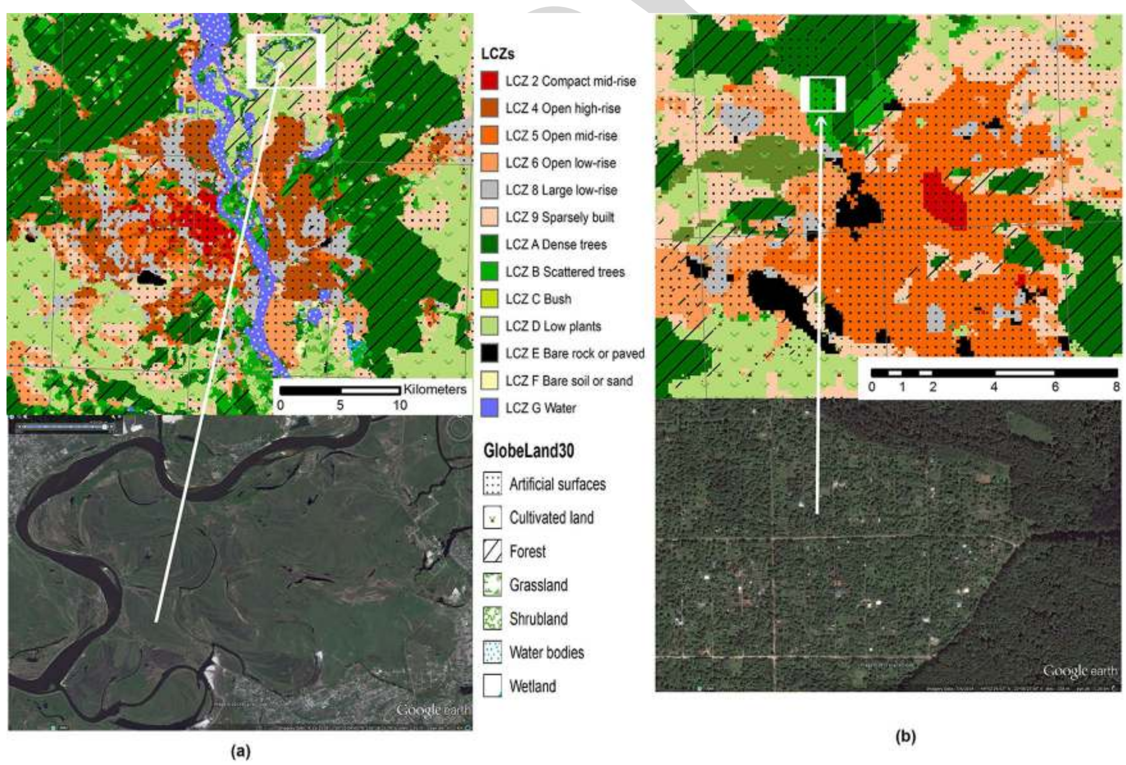

Fig. 8. Examples of disagreeing areas between the LCZ map and GlobeLand30 in (a) Kyiv and (b) Lviv with Google Earth imagery for comparison.

However, validation was undertaken in this study using an independent test dataset to provide additional confidence in the classification. The results, applied only to Kyiv, indicated that the classification accuracy is similar to other land cover products but that there is still room for improvement. However, independent validation using pixels of 120 m is clearly problematic since LCZs are meant to be homogenous areas of 1 km^2 or larger and a postclassification filter is applied to remove small occurrences of LCZ types that are not representative of the larger zone. Validation using larger pixels of at least 1 km^2 may improve the validity of this approach and will be investigated in the future.

Comparison with additional datasets did provide additional confidence in the LCZ classifications of both cities.

TABLE IX
CONFUSION MATRIX FOR KYIV COMPARING LCZS TO GLOBELAND30

	LCZs								
GlobeLand30	Cultivated land and Grassland	Forest	Shrubland	Wetland	Water bodies	Artificial surfaces	Bareland	Sum	UA
Cultivated land and Grassland	10 2637	9595	0	0	575	5282	608	118697	0.86
Forest	5479	54 125	0	0	230	2198	5	62 037	0.87
Shrubland	274	673	0	0	109	144	0	1200	0.00
Wetland	207	138	0	0	3	43	12	403	0.00
Water bodies	465	2176	0	0	8179	300	38	11 158	0.73
Artificial surfaces	11497	3936	0	0	296	50 005	50	65 784	0.76
Bareland	0	0	0	0	0	0	0	0	N/A
Sum	120 559	70 643	0	0	9392	57 972	713	259 279	0.86
PA	0.85	0.77	N/A	N/A	0.87	0.86	0.00	OA	0.83

TABLE X
CONFUSION MATRIX FOR LVIV COMPARING LCZS TO GLOBELAND30

	LCZs							
GlobeLand30	Cultivated land and Grassland	Forest	Shrubland	Wetland	Water bodies	Artificial surfaces	Sum	UA
Cultivated land and Grassland	11 171	605	375	0	27	5053	17 231	0.65
Forest	702	6142	36	0	12	967	7859	0.78
Shrubland	0	0	0	0	0	0	0	N/A
Wetland	1	0	0	0	0	10	11	0.00
Water bodies	7	16	0	0	21	7	51	0.41
Artificial surfaces	666	489	22	0	0	9236	10 413	0.89
Sum	12 547	7252	433	0	60	15 273	35 565	
PA	0.89	0.85	N/A	N/A	0.35	0.60	OA	0.75

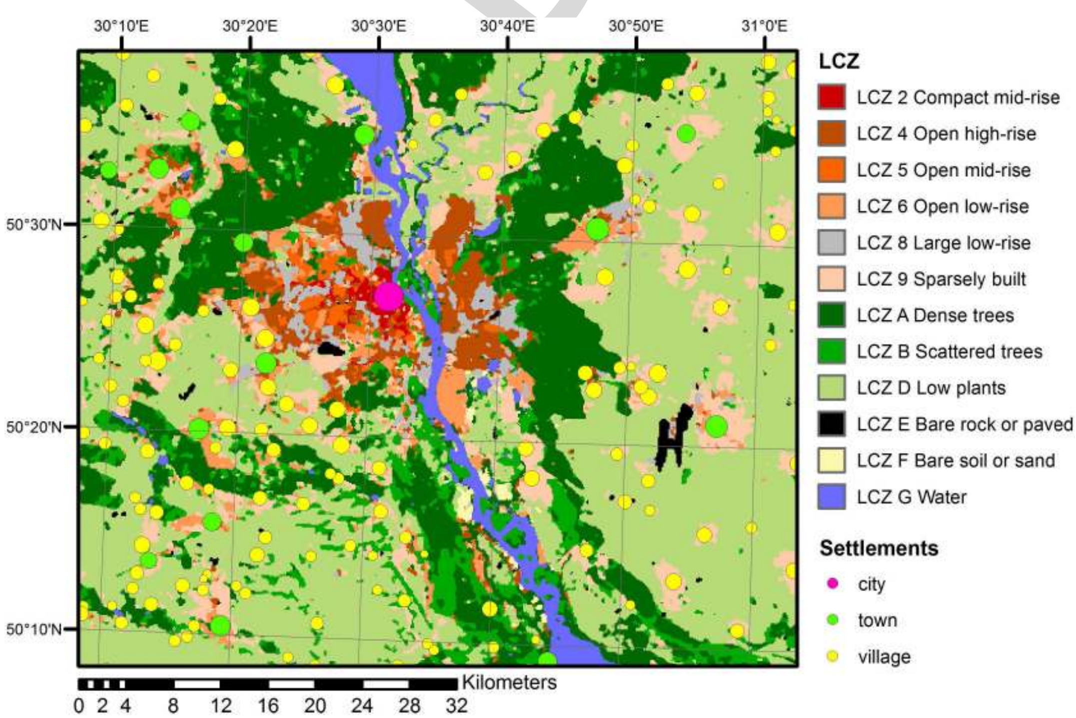

Fig. 9. LCZ map of Kyiv with locations of settlements according to OSM. © OSM contributors.

However, both external datasets have their own errors so agreement between them is subject to some uncertainty. The illustrative examples (Fig. 5) showed that a comparison with external datasets should be treated with appropriate caveats. Comparison with *in-situ* temperature measurements and thermal remote sensing may be other ways to help validate the classification. Validation is clearly an area that will require more attention in the future if LCZs are to be used with confidence in urban climate modeling or as inputs to other applications.

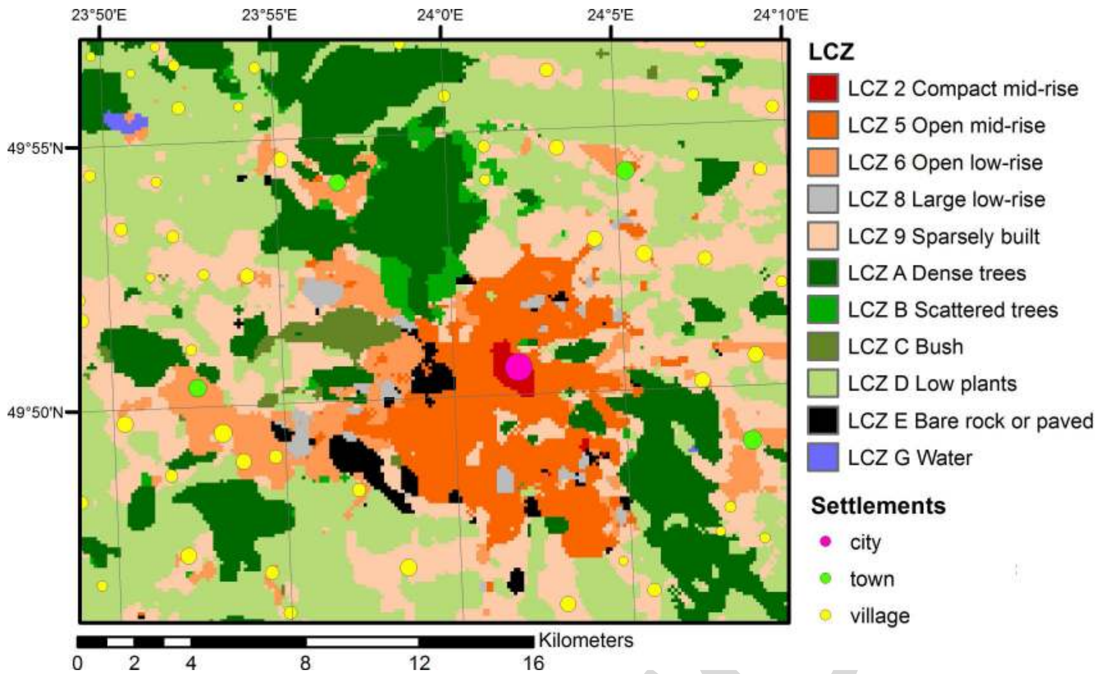

Fig. 10. LCZ map of Lviv with locations of settlements according to OSM. © OSM contributors.

TABLE XI
COMPARISON OF CITY, TOWN, AND VILLAGE LOCATIONS FROM OSM IN RELATION TO THE LCZS FOR KYIV AND LVIV

LCZ	City		Town		Village	
	Kyiv	Lviv	Kyiv	Lviv	Kyiv	Lviv
2	1	1	0	0	0	0
4	0	N/A	6	N/A	1	N/A
5	0	0	0	1	0	0
6	0	0	6	1	21	5
9	0	0	2	1	81	33
A	0	0	0	1	1	0
B	0	N/A	0	N/A	5	N/A
E	0	0	1	0	27	13
Total	1	1	15	4	136	51

VI. Conclusion

In this paper, we applied a methodology for LCZ classification as first outlined in [11] in order to assess the transferability of this concept to two cities in the same climatic zone but that are quite different in urban form and topography, i.e., Kyiv and Lviv. The results demonstrated that LCZs are a generically applicable, culturally neutral classification for urban areas that allowed these cities to be compared in a standardized way. To a certain degree, the heterogeneous versus more homogenous pattern of LCZs in Kyiv and Lviv, respectively, does tell us something about the way cities are organized and could form a framework for further explanation of the patterns of urban form. However, we recognize that these cities and others classified in [12] are in the Global North so we need to further test the classification in cities located in the Global South before we can adequately assess transferability. Some efforts have already been made in this direction with the classification of Khartoum [13].

The workflow in [11] was also extended to consider different methods of validation, in particular validation using an independent dataset and comparison with other sources of information, i.e., OSM and the GlobeLand30 land cover product. The maps will continue to be improved in those areas where confusion between LCZs persists and then contributed to the WUDAPT initiative, which has the overarching goal of creating LCZ classifications for all major cities globally. It will be possible to visualize and download the data for urban climate modeling purposes or for use in many other types of applications that require a detailed delineation of the urban landscape. LCZs will also form the basis of a sampling framework for collecting more detailed information about the urban form and function of cities in the future [4], [24]. More information can be found at: http://www.wudapt.org.

References

[1] I. D. Stewart and T. R. Oke, "Local climate zones for urban temperature studies," *Bull. Amer. Meteorl. Soc.*, vol. 93, no. 12, pp. 1879–1900, May 2012.

[2] T. R. Oke, "The energetic basis of the urban heat island," *Q. J. R. Meteorol. Soc.*, vol. 108, no. 455, pp. 1–24, Jan. 1982.

[3] European Environment Agency, *Urban Atlas*. Copenhagen, Denmark: Copenhagen, 2010.

[4] L. See *et al.*, "Developing a community-based worldwide urban morphology and materials database (WUDAPT) using remote sensing and crowdsourcing for improved urban climate modelling," presented at the Joint Urban Remote Sensing Event (JURSE), Lausanne, Switzerland, 2015.

[5] J. Ching *et al.*, "WUDAPT: Facilitating advanced urban canopy modeling for weather, climate and air quality applications," in *Proc. IEEE 94th Amer. Meteorol. Soc. Annu. Meeting*, Atlanta, GA, USA, 2014, pp. 1–7 [Online]. Available: https://ams.confex.com/ams/94Annual/webprogram/Paper236443.html.

[6] 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 Sens. Environ.*, vol. 154, pp. 192–201, Nov. 2014.

[7] M. Bochow, H. Taubenbock, K. Segl, and H. Kaufmann, "An automated and adaptable approach for characterizing and partitioning cities into urban structure types," in *Proc. 2010 IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Honolulu, HI, USA, Jul. 25–30, 2010, pp. 1796–1799, doi: 10.1109/IGARSS.2010.5652972.

[8] I. Walde, S. Hese, C. Berger, and C. Schmullius, "From land cover-graphs to urban structure types," *Int. J. Geogr. Inf. Sci.*, vol. 28, no. 3, pp. 584–609, Mar. 2014.

[9] S. E. Gill, J. F. Handley, A. R. Ennos, S. Pauleit, N. Theuray, and S. J. Lindley, "Characterising the urban environment of UK cities and towns: A template for landscape planning," *Landscape Urban Plann.*, vol. 87, no. 3, pp. 210–222, Sep. 2008.

[10] S. Pauleit and F. Duhme, "Assessing the environmental performance of land cover types for urban planning," *Landscape Urban Plann.*, vol. 52, no. 1, pp. 1–20, Nov. 2000.

[11] B. Bechtel and C. Daneke, "Classification of local climate zones based on multiple earth observation data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1191–1202, Aug. 2012.

[12] B. Bechtel *et al.*, "Mapping local climate zones for a worldwide database of the form and function of cities," *ISPRS Int. J. Geo-Inf.*, vol. 4, no. 1, pp. 199–219, Feb. 2015.

[13] B. Bechtel, L. See, G. Mills, and M. Foley, "Classification of local climate zones using SAR and multi-spectral data in an arid environment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2016, to be published.

[14] M. F. Hamm, *Kiev: A Portrait, 1800–1917*. 4: printing and 1: paperback printing, Princeton, NJ, USA: Princeton Univ. Press, 1993.

[15] Kiev Department of Statistics, "Population," 2014 [Online]. Available: http://kievobl.ukrstat.gov.ua/content/.

[16] Wikipedia, "List of European Cities by Population," 2015 [Online]. Available: https://en.wikipedia.org/wiki/.

[17] Wikipedia, "Kiev," 2015 [Online]. Available: https://en.wikipedia.org/wiki/Kiev.

[18] Wikipedia, "Lviv," 2015 [Online]. Available: https://en.wikipedia.org/wiki/Lviv.

[19] Lviv Ecotour, "Why Lviv has no street grid," *West Ukraine Genealogy, Tours and Transport*, 2016 [Online]. Available: http://lvivecotour.com/why-lviv-has-no-street-grid/.

[20] Lviv Department of Statistics, "Population," 2014 [Online]. Available: http://www.lv.ukrstat.gov.ua/eng/main_eng.php.

[21] J. Chen *et al.*, "Global land cover mapping at 30 m resolution: A POK-based operational approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 103, pp. 7–27, May 2015.

[22] F. Ramm, J. Topf, and S. Chilton, *OpenStreetMap: Using and Enhancing the Free Map of the World*. Cambridge, England: UIT Cambridge, 2011.

[23] O. Conrad *et al.*, "System for automated geoscientific analyses (SAGA) v. 2.1.4," *Geosci. Model Dev.*, vol. 8, no. 7, pp. 1991–2007, Jul. 2015.

[24] L. See *et al.*, "Generating WUDAPT's specific scale-dependent urban modeling and activity parameters: Collection of level 1 and level 2 data," presented at the 9th International Conference of Urban Climate (ICUC9), Toulouse, France, 2015.

Olha Danylo received the M.Sc. degree in mathematical and computer modeling and the Ph.D. degree in mathematical simulation from Lviv Polytechnic National University, in 2012 and 2015, respectively.

She is a Research Scholar in the Ecosystems Services and Management (ESM) Program with the International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. Her research interests include the use of GIS and machine-learning algorithms to analyze spatial and remote-sensing data, which is aimed at improving land cover mapping and contributes to the science of urban climatology.

Linda See received the Ph.D. degree in spatial applications of fuzzy logic from the University of Leeds, Leeds, U.K., in 1999.

She is a Research Scholar in the ESM Program with the International Institute for Applied Systems Analysis, Laxenburg, Austria. She was a Senior Lecturer in Computational Geography and GIS for more than a decade with the University of Leeds. She spent one year at the Max Planck Institute for Atmospheric Sciences in Germany followed by four years with the Food and Agriculture Organization, Rome, Italy, working in agrometeorology and early warning for food security.

Dr. See is an active member of the Geo-Wiki Team and has worked on a number of crowdsourcing aspects including quality issues, community building, and a branch of Geo-Wiki concerned with validation of urban land cover. She is currently a Chair of the Working Group 1 of the EU COST network: TD1202 Mapping and the Citizen Sensor.

Benjamin Bechtel received the Ph.D. degree in geosciences from the University of Hamburg, Hamburg, Germany, in 2012.

He is a Research Associate with the Cluster of Excellence CliSAP, University of Hamburg. His research interests include urban remote sensing, in particular the characterization of urban surfaces for applications in urban climatology.

Dr. Bechtel is a Reviewer for numerous international journals and Guest Editor for *Remote Sensing* in the special issue "The Application of Thermal Urban Remote Sensing to Understand and Monitor Urban Climates." He was the recipient of the Dissertation Award 2013 for physical geography of the Verband der Geographen an Deutschen Hochschulen (VGDH).

Dmitry Schepaschenko received the Engineer's degree in forestry from Moscow State Forest University, Mytischi, Russia, in 1988, the Ph.D. degree in soil science from Dokuchaev Soil Science Institute, Moscow, Russia, in 1993, and the Professor degree in ecology from Moscow State Forest University, in 2005.

He is a Research Scholar in the ESM Program with the International Institute for Applied Systems Analysis, Laxenburg, Austria. He also holds a part-time Professorial position at the Department of Pedology with Moscow State Forest University. His research interests include global and regional land cover, forest and biomass mapping, full carbon account of Northern Eurasia, remote sensing applications, modeling of structure, productivity and growth of forests, and adaptation and mitigation strategies under global change.

Steffen Fritz received undergraduate degree in physics and geography from the University of Tuebingen, Tuebingen, Germany, in 1995, the M.Sc. degree in GIS from the University of Durham, Durham, U.K., in 1996, and the Ph.D. degree in GIS and wildland research from the University of Leeds, Leeds, U.K., in 2001.

He is a Head of the Earth Observation Systems (EOS) Group in the ESM Program with the International Institute for Applied Systems Analysis, Laxenburg, Austria. He is a Senior Expert in GIS, remote sensing, data interoperability, land use and land cover as well as policy-related land-use modeling. He has authored in the field of earth observation, crowdsourcing, fuzzy logic, remoteness mapping, global and regional vegetation monitoring, crop yield and crop acreage estimation, and wild land research.

Dr Fritz is a Coordinator of a European Research Council (ERC) grant entitled "Harnessing the power of crowdsourcing to improve land cover and land-use Information." He is the driving force behind Geo-Wiki, a tool for visualization, crowdsourcing, and validation of land cover and is currently a Co-Chair of the EU COST network: TD1202 Mapping and the Citizen Sensor.