MethodsX 7 (2020) 101150

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

Local climate zones classification method from Copernicus land monitoring service datasets: An ArcGIS-based toolbox

Ana Oliveira[a,*], António Lopes[b], Samuel Niza[a]

[a] IN+ Center for Innovation, Technology and Policy Research, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[b] Centre of Geographical Studies, Institute of Geography and Spatial Planning (IGOT), Universidade de Lisboa, Universidade de Lisboa, R. Branca Edmée Marques, 1600-276 Lisboa, Portugal

A B S T R A C T

Local Climate Zones (LCZ) have become a worldwide standard for identifying land cover classes, according to their climate-relevant morphological parameters. The LCZ's are mostly used to evaluate urban climate performance, particularly the relationship between the urban heat island effect (UHI) and the characteristics of the built-up environment. The World Urban Database and Access Portal Tools (WUDAPT) has provided a supervised LCZ classification method based only on moderate resolution free satellite imagery, mostly Landsat 7 or 8 (30 m pixel size, in the visible spectrum brands); however, its' results are less accurate for European cities. Conversely, alternative geographic information system (GIS)-based methods developed so far require information that is hardly available to all, such as building footprints or heights. Here, the ArcGIS based LCZ from Copernicus Toolbox (LCZC) provides an alternative classification method that uses only freely accessible information from the Copernicus Land Monitoring Service (CLMS), being possible to replicate it in 800 European urban locations. The method combines Urban Atlas (UA) and Corine Land Cover (CLC) with Tree Cover Density, Dominant Leaf Type and Grassland information, to produce a higher-resolution baseline shapefile that is classified according to each feature's dominant characteristics. The LCZC toolbox output is a LCZ raster map. It has been validated in five European cities: Athens, Barcelona, Lisbon, Marseille, and Naples.

- The LCZC toolbox provides an alternative LCZ GIS-based classification, based on freely accessible CLMS datasets.
- The use of CLMS shapefile higher-resolution inputs, particularly the UA and CLC datasets, ensures an output LCZ map that has greater detail and higher accuracy.
- The availability of CLMS information in 800 European urban areas guarantees that the method can be replicated in those locations.

DOI of original article: 10.1016/j.uclim.2020.100631

* Corresponding author.
E-mail addresses: anappmoliveira@tecnico.ulisboa.pt (A. Oliveira), antonio.lopes@campus.ul.pt (A. Lopes), samuel.niza@tecnico.ulisboa.pt (S. Niza).

https://doi.org/10.1016/j.mex.2020.101150

ARTICLE INFO
Method name: Local Climate Zones from Copernicus ArcGIS Toolbox (LCZC)
Keywords: Local climate zones, ArcGIS toolbox, Land cover classification, Copernicus, Urban atlas, Corine land cover, Urban climate
Article history: Received 9 May 2020; Accepted 13 November 2020; Available online 19 November 2020

Specifications table

Subject Area:	Environmental Science
More specific subject area:	Urban Climate
Method name:	Local Climate Zones from Copernicus ArcGIS Toolbox (LCZC)
Name and reference of original method:	The LCZC toolbox delivers classified maps of the Local Climate Zones scheme by Stewart et al. [1,2] and is an alternative open access method to the World Urban Database and Access Portal Tools (WUDAPT) [3]
Resource availability:	Requirements: ArcGIS Desktop version10.0 (or later), with Advanced License (http://desktop.arcgis.com/en/arcmap/10.3/get-started/system-requirements/arcgis-desktop-system-requirements.htm) The LCZC toolbox is provided with this article, as well as its' python scripts retrieved through the ArcGIS export functionality.

Method details

The Local Climate Zones (LCZ) from Copernicus Toolbox (LCZC) method entails a sequence of steps to reclassify several Copernicus Land Monitoring Service (CLMS) layers into a LCZ-based [1,2] classification, that can be used in urban climate-related studies, in 800 European urban regions. The method aims to provide an alternative solution to the satellite-based World Urban Database and Access Portal Tools (WUDAPT) supervised classification process [3], ensuring the greater accuracy and higher spatial resolution of Geographic Information Systems (GIS)-based methods [4–6] while preserving the ability to be freely reproducible. To process the LCZC tool, ArcGIS software with Advanced License is necessary. The list of inputs the LCZC requires is available in Table 1; all are mandatory, except Building Height (BH) which is only available from CLMS for capital cities. The toolbox is provided as supplementary material to this article in ArcGIS Toolbox format (*.tbx*) (Appendix 1), as well as the corresponding *python* scripts (Appendix 2), as exported through the Model Builder Arc-GIS functionalities.

Due to its' spatial resolution, a combination of the Urban Atlas (UA) and Corine Land Cover (CLC) shapefile datasets was chosen to establish the LCZ baseline vector layer for the procedure, and its' shapefile format was preserved throughout the process. Additional layers related to vegetation were also combined in the model to better distinguish non-urban classes – Tree Cover Density (TCD), Dominant Leaf Type (DLT) and Grassland (GRA). These were all converted from raster to vector shapefiles, classified according to equivalent LCZ based classes, and resulting maps were merged into the LCZ baseline layer, to quantify the dominant class in each polygon. Most built-up LCZ classes (LCZ's 1–10) were reclassified directly from the UA classes, by comparing both classifications' specifications in terms of built-up density, imperviousness degree (IMD) and typical land use/cover. As UA methodology does not allow to distinguish LCZ's 8 and 10, thus additional land use information from OpenStreetMap (OSM) was used to solve that specific gap. On the other hand, non-built-up land cover types (LCZ's A-G) were reclassified according to conditional sentences that filter through the CLC and the vegetation-based classes. The reclassification process was subject to several iterations between algorithm testing and correction, and it was found that combining CLC classes with the High-Resolution Layers related to Tree Cover (i.e. TCD, DLT and GRA) contributed to greater accuracy. Figure 1 summarizes the GIS-based workflow and Figure 2 illustrates the principal correspondence flows between UA and LCZ classes, from the example of Lisbon, where the flow thickness represent the proportion of each class. The method was implemented in ArcGIS software, using the Model Builder functionalities to produce the LCZC custom toolbox.

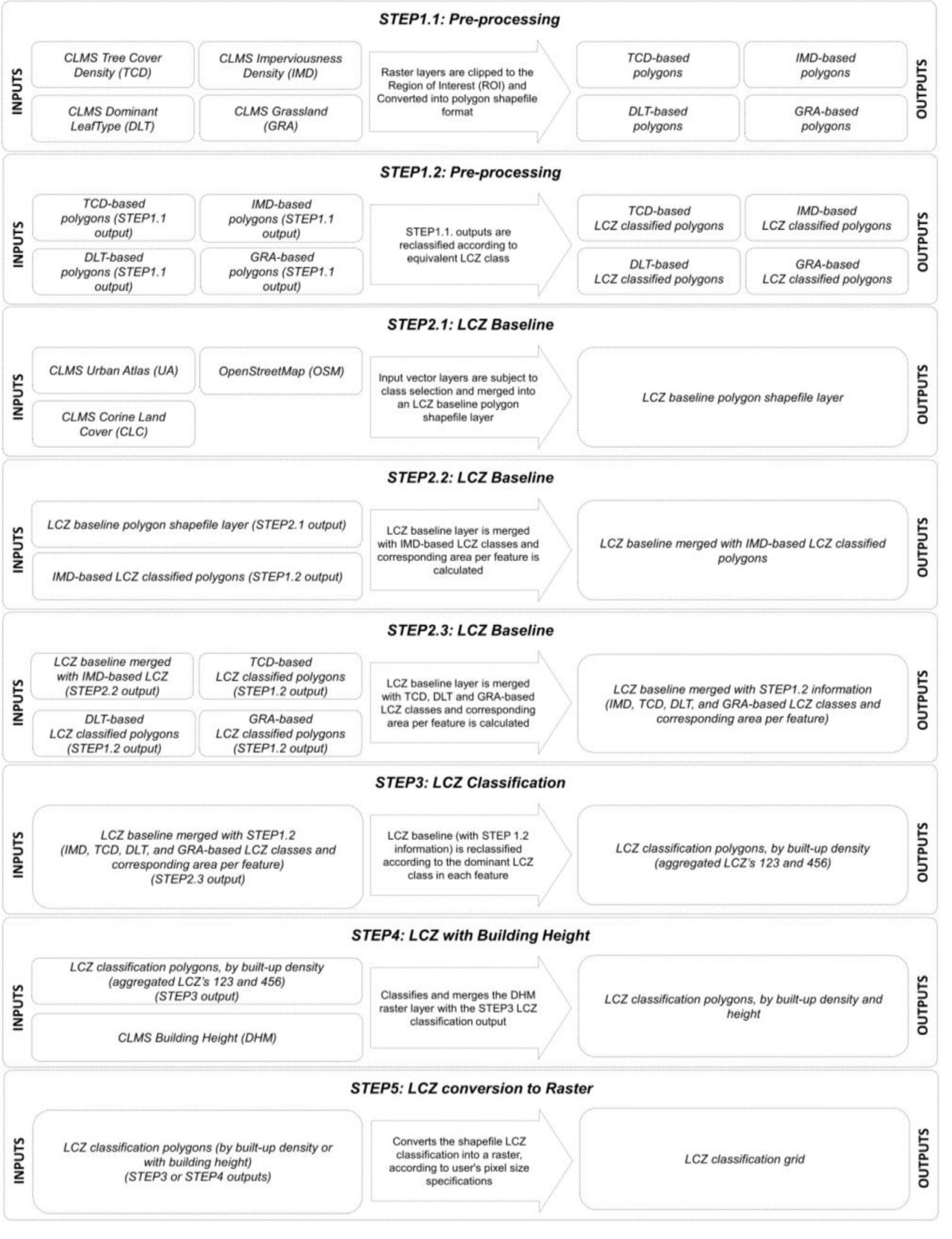

Fig. 1. LCZ GIS-based classification workflow from Copernicus Land Monitoring Service (CLMS) datasets. Detailed diagram based on Oliveira et al. [5].

Table 1
List of land use and land cover datasets used in this study.

Dataset	Version	Release date	Spatial resolution	Original format	Reference	Overall accuracy (%)
Urban Atlas (UA)	2012	2016–08–04	10 m	Vector Shapefile	[7]	88.03- 94.19 [3]
UA Building Height (BH)[1]	2012	2018–04–19	10m	Raster GeoTiff	[8]	n.a.
Corine Land Cover (CLC)	2012	2016–09–19	100 m	Vector Shapefile	[9]	83.6 - 89.7 [4]
Imperviousness Density (IMD)	2015	2018–03–22	20 m	Raster GeoTiff	[10]	89.63 - 98.37 [3]
Tree Cover Density (TCD)	2015	2018–03–22	20 m	Raster GeoTiff	[11]	85.03 - 92.27 [3]
Dominant Leaf Type (DLT)	2015	2018–04–13	20 m	Raster GeoTiff	[12]	85.30 - 95.63 [3]
Grassland (GRA)	2015	2018–09–09	20 m	Raster GeoTiff	[13]	85.50 - 92.67 [3]
OpenStreetMap (OSM)	n/a	n/a	n/a	Vector Shapefile	[14]	n.a.

[1] optional dataset[2]shapefile extracted from OSM, including land-use field column.
[3] as per [15].
[4] as per [16].

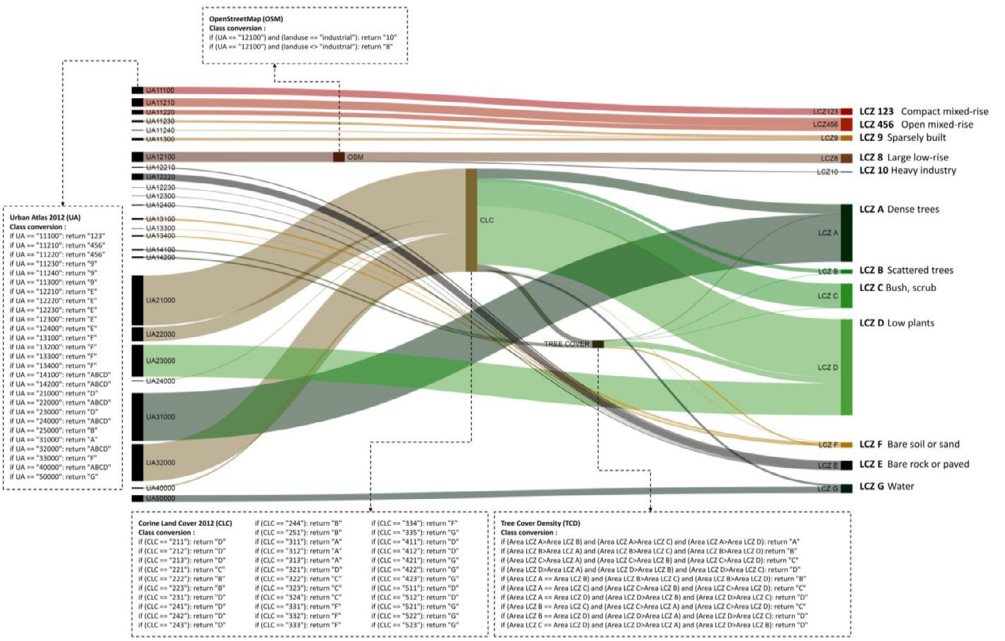

Fig. 2. Sankey Diagram of the LCZ reclassification algorithm – an example of Lisbon's LCZ_v1 map. The flows represent the correspondence between UA (left codes) and LCZ (right codes) classes, and their thickness depicts the proportion of the surface area. The text in boxes corresponds to the LCZ reclassification syntax. Detailed diagram based on Oliveira et al. [5].

The method involves five steps to complete the LCZ classification, from pre-processing of the datasets to the final LCZ maps. If several CMLS layer tiles are needed to contain a study region, they need to be merged into a new mosaic raster previously. The OSM information also must be retrieved previously in shapefile format, with the field column *land-use*.

Besides using ArcGIS-based tools for data conversion and merging, the LCZC workflow involves several *python*-based reclassification steps. Due to its' extension, these steps are provided as supplementary material Appendix 2, in *python*-format, for easier access to the algorithm conditional sentences details.

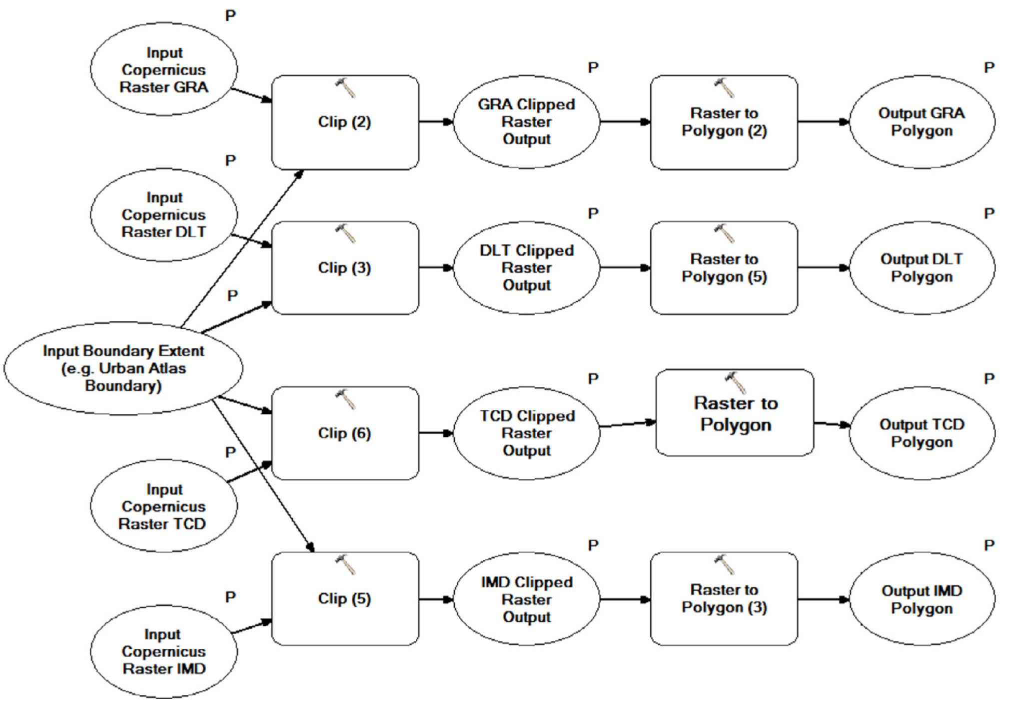

Fig. 3. ArcGIS toolbox model, STEP1.1. Pre-Processing: Copernicus High-resolution layers GRA, DLT, TCD, IMD are clipped and converted to shapefile.

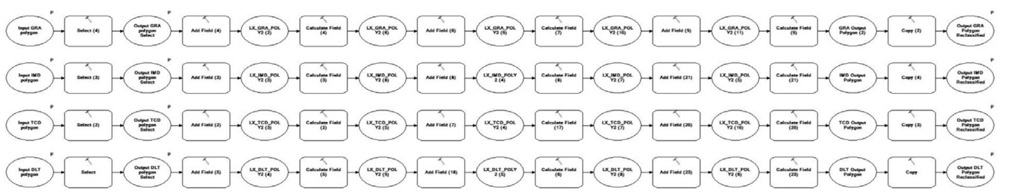

Fig. 4. ArcGIS toolbox model, STEP1.2. Pre-Processing: Copernicus High-resolution layers GRA, DLT, TCD and IMD are classified into LCZ-based classes.

The LCZC tool includes 11 separate custom models, for more proficient use of computational capabilities. These models process the input data in 5 workflow steps, as follows:

STEP 1. Input raster layers are preprocessed, which includes:

STEP 1.1. Pre-Processing: Clips the Copernicus High-Resolution raster Layers TDC, DLT, GRA and IMD according to the UA Boundary, and converts the clipped result to polygon shapefile (see Fig. 3). The raster to polygon conversion is based on a regular squared polygon grid, in which the grid cells have the same size of the original raster pixels, ensuring that values are unchanged (i.e., each polygon feature corresponds to a pixel of equivalent size and value).

STEP 1.2. Pre-Processing: Polygons are reclassified into relevant LCZ-based classes, according to conditional algorithms applied to the raster's gridcode values (see Fig. 4).

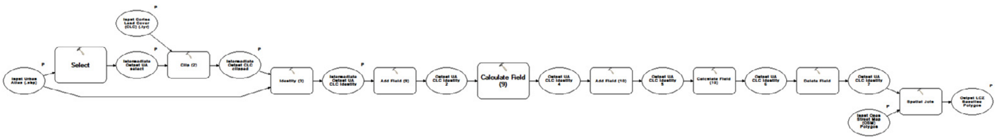

Fig. 5. ArcGIS toolbox model, STEP2.1. LCZ Baseline: UA and CLC layers are subject to class selection and merged into an LCZ baseline feature layer. OSM land-use field is added through the Spatial Join function.

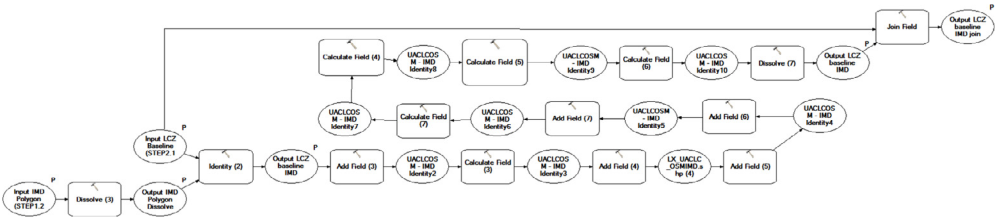

Fig. 6. ArcGIS toolbox model, STEP2.2. LCZ Baseline: LCZ baseline is subject to adding fields that account the area per IMD LCZ-based class.

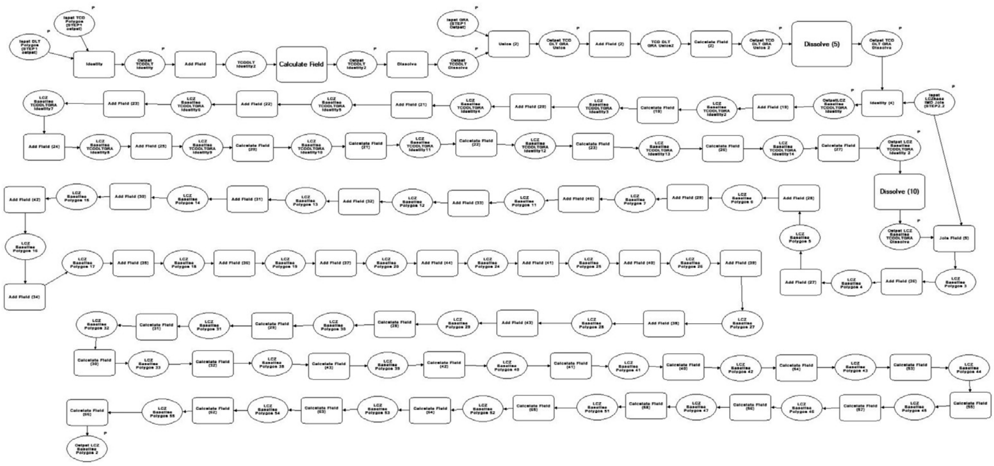

Fig. 7. ArcGIS toolbox model, STEP2.3. LCZ Baseline: the same procedure from STEP2.2 is applied to the TCD, DLT and GRA shapefiles, by adding fields to the LCZ baseline shapefile and calculating the area of each LCZ-based class.

STEP 2. LCZ baseline shapefile is assembled:

STEP 2.1. LCZ Baseline: Assembles the LCZ baseline polygon layer, based on conditional selections and merging functions, applied to the UA, CLC and OSM inputs (see Fig. 5). UA features are chosen to represent urban classes, while CLC features are preferred in non-urban land cover typologies. Selected features of both layers are merged through the Identity tool. The OSM land use information is added through the Spatial Join function.

STEP 2.2. LCZ Baseline: Calculates area per IMD class, adding the resulting field to the LCZ baseline polygon layer (obtained in the previous STEP 2.1) (see Fig. 6).

STEP 2.3. LCZ Baseline: Calculates area per TCD, DLT and GRA classes, adding the resulting field to the LCZ baseline polygon layer, from previous 2.2 (see Fig. 7).

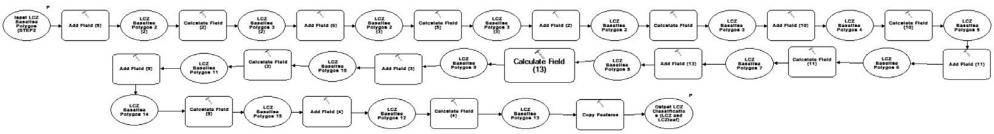

Fig. 8. ArcGIS toolbox model, STEP3. LCZ Classification: LCZ baseline is reclassified according to the dominant LCZ class, without building height information. Two classification fields are created in the attribute table: LCZ and LCZ_leaf (LCZ with added DLT classification).

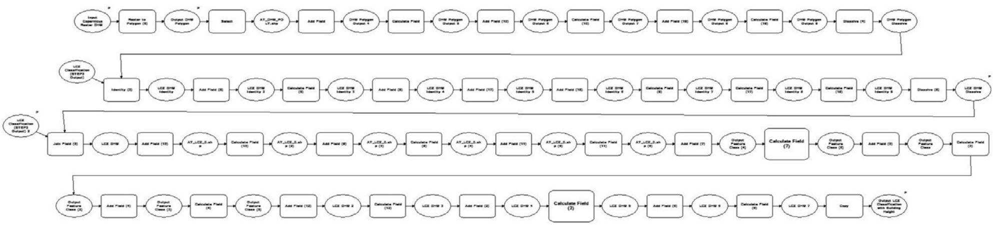

Fig. 9. ArcGIS toolbox model, STEP4. LCZ with Building Height: reclassifies the LCZ baseline LCZ and LCZ_leaf fields, by decoding LCZ's 123 and 456 polygon features according to dominant BH class.

STEP 3. LCZ classification: LCZ baseline dataset from previous STEP2.3 is reclassified as two LCZ classification fields (*LCVZv1* and *LCZv1_leaf*), both without building height information (i.e. Urban LCZ classes are merged by density, e.g. LCZ's 1, 2 and 3 are classified as LCZ '123′). The LCZ classification fields created are named *LCZ* (containing classes 123, 456, 8, 9, 10, A, B, C, D, E, F and G), and *LCZ_leaf* (containing the same classes but separating LCZ's A and B according to DLT, deciduous or coniferous) (see Figure 8).

STEP 4. LCZ with Building Height: converts, classifies and merges the BH raster layer with the STEP3 LCZ classification output; LCZ's 123 and 456 are reclassified according to dominant BH: LCZ 1 Compact high-rise LCZ 2 Compact midrise LCZ 3 Compact low-rise LCZ 4 Open high-rise LCZ 5 Open midrise LCZ 6 Open low-rise. Urban features without BH information remain as LCZ 123 and LCZ 456. The two fields created, *LCZ_BH* and *LCZ_leaf_BH*, correspond the reclassification of the two fields from the previous STEP3 (see Figure 9).

STEP 5. LCZ conversion to Raster: converts the shapefile LCZ classification into a raster, according to user's pixel size specifications. STEP5 has 4 alternative models, according to the LCZ desired content: (a) *LCZv1* - classification without DLT or BH; (b) *LCZv1_leaf* - classification with DLT but without BH; (c) *LCZv1_BH* classification with BH, but without DLT; and (d) *LCZv1_leaf_BH* classification with BH and DLT. The resulting raster uses a numerical codification for LCZ's classes, and the corresponding attribute table also contains 2 string fields with LCZ's names and description, and a numerical field with the corresponding area (m^2) (see Figure 10 and Table 2).

The presented method was tested in 5 Southern European cities [5]: Athens, Barcelona, Lisbon, Marseille and Naples [4]. Even though Stewart and Oke describe 300 m as a reasonable minimum radius for the LCZ classification [2,17], the resulting LCZ shapefile datasets were converted to a 50 m pixel raster format, where each pixel value depicts the LCZ class that has the greatest area. Each city's LCZ classification was subject to accuracy assessment analysis, by randomly selecting samples of pixels per class, stratified by surface area coverage. About 550 samples per each city were classified according to the dominant LCZ type (based on satellite true color imagery and 3D information from Google Earth). The sample's classification was compared with the LCZC toolbox output, and results re-arranged into a confusion matrix. Average overall accuracy (OA) was 81% and Kappa coefficient 0.79. Correctly classified pixels varied according to LCZ class, as built-up LCZ classes 1–10 revealed, on average, 90,0% agreement, but non-built-up LCZ classes A-G had fewer correctly classified sample

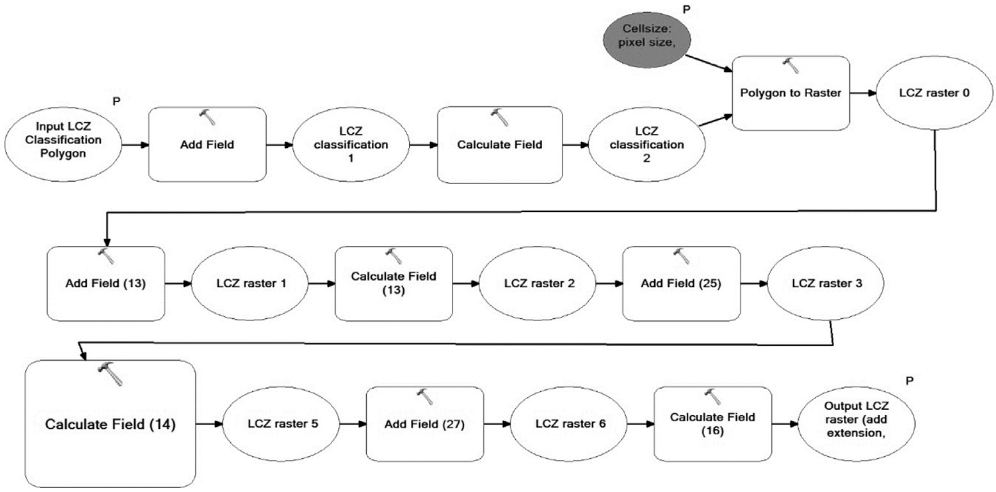

Fig. 10. ArcGIS toolbox model, STEP5a. LCZ conversion to Rater: the classified LCZ baseline shapefile is converted to Raster format, according to the user's chosen spatial resolution. There are 4 STEP5 models, one per each alternative LCZ field.

Table 2
Alternative LCZ classifications and the corresponding list of LCZ classes contained in the attribute tables[1].

Numerical code	LCZ class	Description	LCZv1.tif *Without BH Without DLT*	LCZv1_leaf.tif *With DLT Without BH*	LCZv1_BH *With BH Without DLT*	LCZv1_BH_leaf *With DLT Without BH*
100,100	1	Compact high-rise			X	X
100,200	2	Compact midrise			X	X
100,300	3	Compact low-rise			X	X
100,123	123	Compact mix-rise	X	X	X	X
100,400	4	Open high-rise			X	X
100,500	5	Open midrise			X	X
100,600	6	Open low-rise			X	X
100,456	456	Open mix-rise	X	X	X	X
100,800	8	Large low-rise	X	X	X	X
100,900	9	Sparsely built	X	X	X	X
101,000	10	Heavy industry	X	X	X	X
110,100	A	Dense trees	X	X	X	X
110,110	A coniferous	Dense trees coniferous		X		X
110,120	A deciduous	Dense trees deciduous		X		X
110,200	B	Scattered trees	X	X	X	X
110,210	B coniferous	Scattered trees coniferous		X		X
110,220	B deciduous	Scattered trees deciduous		X		X
110,300	C	Bush scrub	X	X	X	X
110,400	D	Low plants	X	X	X	X
110,500	E	Bare rock or paved	X	X	X	X
110,600	F	Bare soil or sand	X	X	X	X
110,700	G	Water	X	X	X	X

[1] X = Class contained in the dataset.

pixels. The lowest accuracies occur in low-density vegetation types since only the dense trees class (LCZ A) proved to have 80% correct results, on average. LCZ 123 (compact urban fabric), LCZ 8 and LCZ 10 revealed the most noteworthy agreement, being correct in approximately 95% samples. LCZ 456 (open urban fabric) and LCZ 9 (sparsely built) were found to be less accurate, even though above the 80% threshold. This agrees with the great diversity of suburban neighbourhood typologies, more difficult to group in one class (an issue also present in the UA dataset). It should be noted that, while the LCZC toolbox uses the high-resolution CLMS layers to improve the Urban Atlas and Corine Land Cover miss-classification rates, particularly in rural areas, it is still limited by its inputs overall accuracy. This agrees with the fact that the resulting LCZ's maps revealed an overall accuracy that is slightly lower than that of its input datasets. Nonetheless, the LCZC toolbox aims to provide a readily available tool for LCZ mapping in European cities, improving the accuracy reached through satellite-based alternatives, and to be useful on a metropolitan scale.

Acknowledgments

AO acknowledeges "Fundação para a Ciência e Tecnologia (FCT) – Portugal" for providing financial support to the research, through the individual Ph.D. grant number PD/BD/52304/2013.

The Copernicus Land Monitoring Service datasets used in this study were produced "with funding by the European Union".

Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.

We acknowledge the use of imagery provided by services from the Global Imagery Browse Services (GIBS), operated by the NASA/GSFC/Earth Science Data and Information System (ESDIS, https://earthdata.nasa.gov) with funding provided by NASA/HQ.

AO would like to thank all the Zephyrus research group for guidance, particularly Professors António Lopes and Ezequiel Correia (CEG, IGOT, Universidade de Lisboa) for providing helpful supervision throughout the development of the GIS-based tool.

AO gratefully acknowledges the editors and reviewers by their helpful suggestions to improve an earlier version of the manuscript.

Declaration of Competing Interest

The Authors confirm that there are no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mex.2020.101150.

References

[1] I.D. Stewart, T.R. Oke, Local climate zones for urban temperature studies, Bull. Am. Meteorol. Soc. (2012), doi:10.1175/BAMS-d-11-00019.1.
[2] I.D. Stewart, T.R. Oke, E.S. Krayenhoff, Evaluation of the "local climate zone" scheme using temperature observations and model simulations, Int. J. Climatol. (2014), doi:10.1002/joc.3746.
[3] B. Bechtel, P. Alexander, J. Böhner, J. Ching, O. Conrad, J. Feddema, G. Mills, L. See, I. Stewart, Mapping local climate zones for a worldwide database of the form and function of cities, ISPRS Int. J. Geo Inf. (2015), doi:10.3390/ijgi4010199.
[4] A. Oliveira, A. Lopes, S. Niza, Local climate zones datasets from five southern European cities: Copernicus based classification maps of Athens, Barcelona, Lisbon, Marseille and Naples, Data Br. 31 (2020), doi:10.1016/j.dib.2020.105802.
[5] A. Oliveira, A. Lopes, S. Niza, Local climate zones in five southern European cities: an improved GIS-based classification method based on Copernicus data, Urban Clim. 33 (2020) (in press), doi:10.1016/j.uclim.2020.100631.
[6] R. Wang, C. Ren, Y. Xu, K.K.L. Lau, Y. Shi, Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: a case study of Hong Kong, Urban Clim. (2018), doi:10.1016/j.uclim.2017.10.001.
[7] European Environment Agency (EEA), Urban atlas 2012, (2016). https://land.copernicus.eu/local/urban-atlas/urban-atlas-2012.
[8] European Environment Agency (EEA), Building height 2012, (2018). https://land.copernicus.eu/local/urban-atlas/building-height-2012.
[9] European Environment Agency (EEA), Corine land cover 2012, Https://Www.Eea.Europa.Eu/Data-and-Maps/Data/External/Corine-Land-Cover-2012. (2016).
[10] European Environment Agency (EEA), Imperviousness density, (2018). https://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness/status-maps/2015.
[11] European Environment Agency (EEA), Tree cover density, (2018). https://land.copernicus.eu/pan-european/high-resolution-layers/forests/tree-cover-density.
[12] European Environment Agency (EEA), Dominant leaf type, (2018). https://land.copernicus.eu/pan-european/high-resolution-layers/forests/dominant-leaf-type.

[13] European Environment Agency (EEA), Grassland 2015, (2018). https://land.copernicus.eu/pan-european/high-resolution-layers/grassland/status-maps/2015.
[14] OpenStreetMap contributors. (2015) Planet dump [Data file from $date of database dump$]. Retrieved from https://planet.openstreetmap.org, [Data file from June 2018], (2015).
[15] European Environment Agency (EEA), Validation sampling scheme over Europe for global, panEuropean and local validation purposes, (2019). https://land.copernicus.eu/user-corner/technical-library/validation-sampling-scheme-over-europe-for-global-pan-european-and-local-validation-purposes.
[16] European Environment Agency (EEA), Corine land cover 2012 final validation report, (2017). https://land.copernicus.eu/user-corner/technical-library/clc-2012-validation-report-1.
[17] I.D. Stewart, T.R. Oke, Local climate zones for urban temperature studies, Bull. Am. Meteorol. Soc. (2012), doi:10.1175/BAMS-d-11-00019.1.