Urban Climate 27 (2019) 64–89

Contents lists available at ScienceDirect

Urban Climate

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

Comparison between local climate zones maps derived from administrative datasets and satellite observations

Julia Hidalgo[a,*], Guillaume Dumas[a], Valéry Masson[b], Gwendall Petit[c], Benjamin Bechtel[d], Erwan Bocher[c], Michael Foley[e], Robert Schoetter[b], Gerald Mills[e]

[a] LISST UMR 5193, CNRS/UT2J, Toulouse, France
[b] CNRM UMR 3589, Météo-France/CNRS, Toulouse, France
[c] Lab-STICC UMR 6285, CNRS, Vannes, France
[d] Center for Earth System Research and Sustainability, Universität Hamburg, Germany
[e] School of Geography, UCD, Dublin, Ireland

ARTICLE INFO

Keywords:
Local climate zones
MApUCE
WUDAPT
Urban databases

ABSTRACT

This paper proposes a method for generating maps of Local Climate Zones (LCZs) within a GIS using administrative and 2.5D building databases. The LCZs are computed from morphological indicators and building typology, on vector reference spatial units that correspond to urban islets, i.e. blocks of buildings surrounded by nearby roads. The main originality is that, while mean building height criteria correspond exactly to the LCZ classification, a k-means statistical method is used to determine, for each city, the limits between compact and open (and sparsely built) LCZs for high-, mid- and low-rise LCZs, respectively. For example, in SO12 LCZ look-up tables and the WUDAPT-L0 mapping approach, "compact" LCZs correspond to a building density of over 40%. The resulting groups for Nantes, Toulouse and Paris for mid-height, treated with the proposed statistical method, are 36%, 37% and 33.8% respectively. The LCZ maps for these three cities are compared to the WUDAPT LCZ maps, the latter being obtained from satellite imagery at a resolution of 100 m. MApUCE LCZ maps show more spatial details, due to their finer resolution, and more variety in urban LCZs within each conurbation. This is very important for modeling micro-climatic effects on town peripheries.

1. Introduction

The presence of a city disrupts local and regional weather conditions, creating a distinctive urban climate at a hierarchy of scales by altering the surface-air exchanges of heat, moisture, mass and momentum (Oke, 1988). This effect is present in nearly every single meteorological variable and its magnitude depends on aspects of urban form and function, that is, the physical structure and pattern of occupation of the city. The most widely studied urban climate effect is the urban heat island (UHI), which describes the differences in surface, sub-surface and air temperatures in cities when compared to the surrounding 'natural' environment (Oke, 1982). Other impacts include changes to flow dynamics as the overlying air adjusts to a complex, rough and heterogeneous underlying surface which influences the dispersion of air pollutants and heat emitted near the ground (Hidalgo et al., 2009). Over the last few decades, urban climate science has made significant progress in linking the properties of the urban surface cover, including its extreme spatial heterogeneity, to changes in the overlying atmosphere (Oke et al., 2017). Significant gaps in our understanding of processes remain,

* Corresponding author.
E-mail address: julia.hidalgo@univ-tlse2.fr (J. Hidalgo).

https://doi.org/10.1016/j.uclim.2018.10.004
Received 10 November 2017; Received in revised form 2 August 2018; Accepted 23 October 2018
2212-0955/

but it is generally acknowledged that the outstanding issue for urban climate science is the need to transfer knowledge into urban decision-making (Ng, 2015).

One obstacle that inhibits this transfer is the absence of coherent and consistent urban databases suited to urban climate studies. These databases would contain information on both urban form (land-cover, materials and building dimensions) and function (occupation patterns) that could be used to support observational and modeling projects and inform climate mitigation and/or adaptation policies. This data gap is especially acute in the rapidly growing cities in poorer parts of the world (where the need is greatest) but even in wealthy cities, the data infrastructure that could support knowledge transfer is often incomplete, inconsistent, spatially coarse, imprecise and frequently unavailable. Ideally, the available data would be climatically meaningful and acquired using consistent methods to allow cross-city comparisons. Examples of recent efforts along these lines are the Global Human Settlement Layer at 100 m of resolution and the second generation Ecoclimap database (https://opensource.umrcnrm.fr/projects/ecoclimapsg/wiki).

One approach to getting suitable data for environmental and meteorological/climatic studies is to employ the Local Climate Zone (LCZ) typology to break the urban (and natural) landscape up into 'neighbourhoods' (ideally $\geq 1\,km^2$) that are relatively homogeneous in their make-up. The LCZ scheme, which has 10 urban classes and 7 natural ones, was designed to standardize the description of observation sites used in urban heat island studies (Stewart and Oke, 2012, hereinafter referenced as SO12), but it has two significant attributes: it is a 'universal' classification with limited cultural bias, which means that it can be applied to cities worldwide; and each LCZ type is linked to a series of variables that relate to the urban climate effect generally. These variables include the sky view factor, the aspect ratio, the mean building or tree height, the terrain roughness class, the building surface fraction, the pervious and impervious surface fractions, the thermal admittance, the albedo and the anthropogenic heat flux. The associated values are often referred to as urban canopy parameters (UCPs) as they describe the attributes of the urban surface in a manner that can be integrated into climate models. Mapping cities into LCZ types is an increasingly common practice in urban climate research, but there has been no cross-comparison of the methods in use, which include satellite images, aerial photographs, administrative databases and fieldwork.

One category of approaches can be described as 'bottom-up' as they are based on information acquired for individual cities. One approach is to sample the urban landscape using fieldwork; this requires considerable expertise as shown by Houet and Pigeon (2011) and Leconte et al. (2015) for Toulouse and Nancy, France respectively. More commonly, administrative or topographic data (on building footprints, heights, green spaces, etc.) and Geographic Information Systems (GIS) software are used. Raster-based methods superimpose a standard grid over the urban landscape and acquire information on selected variables (e.g. building height, sky view factor) at the scale of the individual cell. Each variable is stored in a layer and the gridded layers are combined using rules to generate LCZ types. This method has been used to generate LCZ maps for Hong-Kong (Zheng et al., 2017), for Nagpur in India (Kotharkar and Bagade, 2017), for three medium-sized Central European cities, Brno, Hradec Králové, and Olomouc in the Czech Republic (Geletic and Lehnert, 2016), and for Bilbao in Spain (Acero, 2012). Vector-based methods capture the boundary of an LCZ neighbourhood and represent a more precise delineation of contiguous neighbourhood types as individual objects; Unger et al. (2014) and Perera et al. (2012) have used this method for Szeged and Colombo, respectively.

In all of these studies, decision-making algorithms are employed that are based on typical UCP values proposed by SO12 (referred to henceforth as SO12). For a number of reasons, implementing a workflow to generate LCZ maps from urban data is not straightforward. First, the wide ranges of UCP values associated with each LCZ type overlap with those in other types. Second, the variable needed for the LCZ types may not be in the urban dataset and other exogenous indicators must be used as substitutes; for example, Geletic and Lehnert (2016) used indicators on the *Number of buildings per hectare*, the *Number of areas of continuous surface of crown cover* or the *Number of continuous fragments of all vegetation per ha*. Two of the required variables (albedo and anthropogenic heat flux) are especially difficult to obtain across the urbanized landscape and consequently are rarely included in the classification workflow. Third, the sequence of steps taken to classify a neighbourhood into an LCZ type is not fixed; some begin by classifying vegetated areas (Kotharkar and Bagade, 2017) whereas others start with the built categories (i.e. LCZ types 1 to 10). The advantage of these techniques is the opportunity to make best use of available data and local expertise to create LCZ maps with known accuracy and precision. The disadvantage is the variable nature of the underlying data, which makes cross-city comparison difficult and creating an international database a formidable challenge.

Another category of LCZ approaches are 'top-down' approaches, as they are based on satellite instruments that acquire information using passive or active sensors. Passive sensors respond to natural radiation that is emitted and reflected; active sensors record radiation emitted by the sensor that is reflected back from the target surface (e.g. synthetic aperture radar). The former can provide details on the Earth's surface when not obscured and are distinguished by their spectral and spatial resolution. Perhaps the best-known project is Landsat, which was launched in 1972. In its current mode (Landsat 8), it supports two instruments, the operational land imager and thermal infrared sensor, which collect information over 11 bands at spatial resolutions ranging from 15 to 100 m.

The World Urban Database and Access Portal Tools (WUDAPT) project utilizes these freely available data, which represent the lowest level (L0) of data in its information hierarchy, to map cities into LCZ types (Ching et al., 2017).

The protocol for generating LCZ maps employs supervised classification to automatically classify Landsat scenes using training data created by an urban expert (Bechtel and Daneke, 2012). Various tools within WUDAPT permit these data to be downloaded in a format suited to model applications; for example, Brousse et al. (2016) employed L0 data for Madrid to run the Weather Research Forecasting (WRF) model. The WUDAPT process is efficient, so cities can be mapped quickly and updated with little effort. Given the global cover provided by Landsat, this process provides the potential for creating a worldwide urban database, but the derived value is predicated on the quality of the training areas.

Fig. 1. spatial scales considered in the MapUCE project. Building (a), block (b) and islet, also called Reference Spatial Unit (RSU, c).

In WUDAPT, higher levels (Level 1, Level 2) provide more precise information on shapes (building height, street length, etc.) and functions (office, industrial, residential, etc.) for the whole city in Level 1 and with intra-city spatial variability in Level 2; but these levels are not directly achievable using the previously mentioned methods, and in this case several bottom-up methods are used.

This paper contributes to the ongoing research in this area by:

1. Proposing an original method for generating LCZ maps within a GIS framework using French administrative and topographic data acquired as part of the MApUCE project (section 2). The method takes regional variations into account and defines the UCPs associated with LCZ types more precisely. Here we present the LCZ output for three case study cities: Toulouse, Paris and Nantes (section 3).
2. Performing a geographical analysis and quantitative comparison of MapUCE maps with WUDAPT-L0 maps for these three cities (section 4 and section 5).

2. The MApUCE urban database

The *Applied Modeling and Planning Laws: Climate and Energy* (MApUCE) project is designed to facilitate the integration of climate information into relevant French legal documents and urban policies. This information includes quantitative data on the climate and energy consumption at the building scale. Currently these data are not available or, if they are, are scarce; for example most official weather stations are located outside cities and fine-scale energy consumption data are often considered economically sensitive information. As a result, there is little climate information for urban places and energy data are only available over long periods and large territories.

MApUCE has two objectives: to build a database and to integrate climate knowledge into policy. The database consists of indicators that are pertinent for urban climate and energy consumption studies. These indicators include information on land use, urban morphology, building type, and user behavior (Tornay et al., 2017; Schoetter et al., 2017). To overcome the obstacles to knowledge integration, legal and planning documents have been analyzed by experts to identify areas where regulations or incentives could be included that would promote climate-based actions. A few 'best cases' have been studied to understand the conditions that are needed for policy innovations and to encourage interactions between actors. Finally, based on urban planning agency requirements, the Urban Climate Maps methodology (see Ng and Ren, 2015) is being adapted to the French urban and legal context.

The MApUCE project is interdisciplinary in its structure and draws on partners in the following fields: law, urban climate, building energetics, architecture, sociology, geography and meteorology, as well as the National Federation of Urban Planning Agencies (FNAU).

2.1. The morphological database description

The MApUCE morphological database (Bocher et al., 2018) contains a set of indicators computed at three urban scales, corresponding to building, block and islet (Fig. 1) based on input dataset provided by French Institutes (Table 1).

MApUCE uses vector-based geographical data on buildings, cadastral parcels and roads so that the spatial resolution is close to the scale of urban features. The database is built in three steps consisting in:pre-processing, to reduce inconsistencies and to perform islet extraction through a Voronoi tessellation algorithm. Islets, called herein RSUs, are delimited by the centre line of road surfaces.

1. The computation of 64 morphological indicators organized into 5 categories: count, area, shape, distance and others (Table 2).
2. The attribution of an urban typology class to each building. A set of *Building typologies* has been used as described in Tornay et al. (2017) to synthesize the main metrics and to illustrate the urban fabric organization. At the RSU scale, the percentage of each type

Table 1
Datasets used to compute the morphological indicators in the MApUCE morphological database.

Dataset	Description
IGN-BDTopo	Topographic data, in vector format, provided by the French National Geographical Institute (IGN) (http://professionnels.ign.fr/bdtopo), which contains information on individual buildings.
Parcellaire IGN	Cadastral parcels, in vector format, provided by IGN (http://professionnels.ign.fr/bdparcellaire)
Gridded population	The population data is obtained from the French National Institute for Statistics and Economic Studies (INSEE) (https://www.insee.fr/en/accueil).

Table 2
Examples of morphological indicators for the different groups.

Group	Indicators	Examples
Count	Number of features	Number of buildings and blocks by RSU
Area	Areas or ratio of areas	Building floor; exposed building facades, road fraction, vegetative fraction, etc. by RSU
Shape	Volume, compactness, form factor, concavity, fractal dimension, contiguity	Building volume, block compactness and mean building height by RSU
Distance	Distance between one feature and *n* features	Distance between buildings, between buildings and roads for each RSU
Others	Passive volume, Main direction, Holes area	Building or blocks main direction; Block courtyard area

Table 3
Building typology used in MApUCE as described in Tornay et al., 2017. Identifiers were not translated from French to English to facilitate the identification of typologies in the MApUCE database.

Description	Identifier and Legend
Extended low-rise Industrial, commercial or agricultural buildings. They are characterised by their simple morphology and their large Footprint	Ba
High-rise building A building of more than 12 storeys such as an apartment tower or office tower	Bgh
Continuous row of mid-rise Perimeter islet development: Connected buildings with street front elevations. This typology is often present in historic centres, urban fabric of the industrial	Icif
Discontinuous row of mid-rise A building complex in the centre of the urban islet	Icio
Detached mid-rise One or more buildings built in the centre of the islet	Id
Informal building Ephemeral constructions, non-traced on registers (caravans, temporary prefabricated buildings, etc.)	Local
Continuous row of low-rise Typical intermediary housing, terraced houses with patios, constructions typical of historic centres	Pcif
Discontinuous row of low-rise Street aligning terraced houses with gardens at the back	Pcio
Detached low-rise One or two-storey houses of at least four façades often located in the centre of plot of land	Pd
Semi-detached low-rise Town houses, terraced houses or houses detached on one side, with façades aligning the street	Psc
Other	na

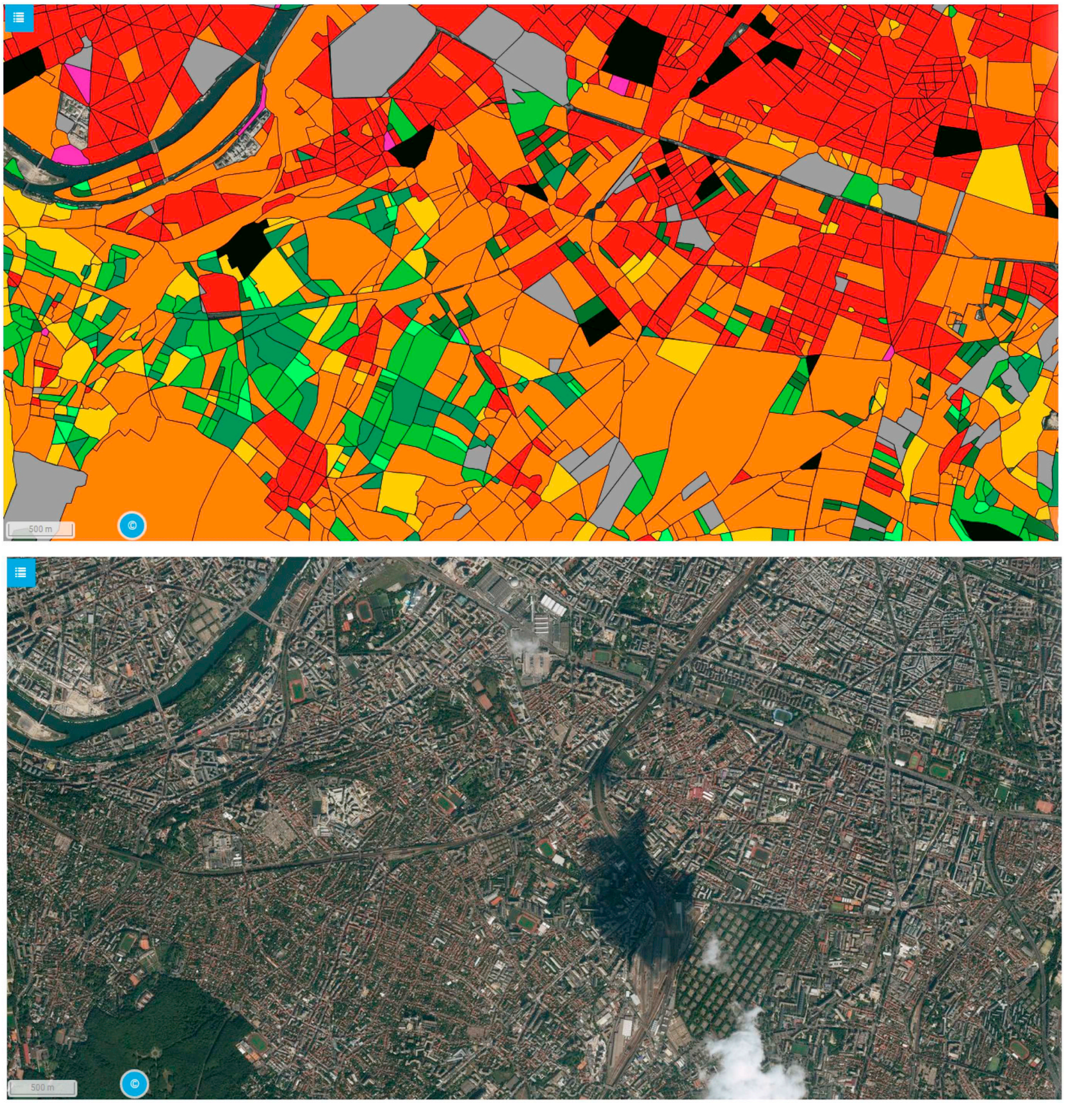


Fig. 2. Dominant building type and aerial map (ESRI World Imagery) at the RSU scale for a district of Paris. Code of colours corresponds to Table 3 (Other type is represented with a transparent colour).

is calculated and the primary and secondary types of building are identified. Fig. 2 shows an example of the dominant building typology for a district in Paris.

2.2. Case studies: urban units of Toulouse, Paris and Nantes

In MApUCE, the spatial extent of a city is defined by the notion of urban unit. In France, an urban unit (*Fr: "Unité Urbaine"*) is a statistical area defined by the National Institute of Statistics and Economics Studies, for the measurement of contiguously built-up areas. It is defined as 'a municipality or a group of municipalities which includes a continuously built up zone (where constructions are not > 200 m apart) and at least 2000 inhabitants'.[1] This definition is in accordance with United Nations recommendations for the measurement of contiguous built-up areas and other comparable units used in the United States, the 'Urbanized Area' and the 'urban area' definition shared by the United Kingdom and Canada.

In this study three urban units were chosen as case studies: Paris, Toulouse and Nantes. Paris is the capital and most populous city of France; in the 2010 census it had over 10 million inhabitants across an area of 1749 km^2.

[1] https://www.insee.fr/en/metadonnees/definition/c1501.

Table 4
Characteristics of the MApUCE database for Toulouse, Paris and Nantes.

	Urban unit		
	Toulouse	Paris	Nantes
Number of communes	100	334	38
Number of RSUs	15,611	61,029	15,724
Number of RSUs with buildings	12,366	47,357	10,626
Number of buildings	218,240	1,607,631	262,643
Number of blocks	206,280	100907	169,281
Dominant building type RSU			
Ba	1355 (10.9%)	3378 (7.1%)	1039 (9.8%)
Bgh	20 (0.2%)	321 (0.7%)	15 (0.1%)
Icif	606 (4.9%)	7151 (15.1%)	391 (3.7%)
Icio	1411 (11.4%)	8178 (17.3%)	846 (8.0%)
Id	1574 (12.7)	3273 (6.9%)	433 (4.1%)
Local	48 (0.4%)	624 (1.3%)	191 (1.8%)
Pcif	66 (0.5%)	1061 (2.2%)	487 (4.6%)
Pcio	471 (3.8%)	3556 (7.5%)	1252 (11.8%)
Pd	6291 (50.9%)	16,961 (35.8%)	4657 (43.8)
Psc	524 (4.3%)	2854 (6.1%)	1315 (12.3%)

Toulouse is the fourth-largest city in France with a population of over 800,000, occupying 1175 km^2; its large surface area, comparable to that of Paris but with one tenth the number of inhabitants, makes Toulouse one of the least dense cities in France. Nantes is the sixth-largest city in France with almost 600,000 inhabitants spread over 547 km^2.

Table 4 presents the characteristics of the MApUCE database for the three cities. The number of RSUs and buildings in Paris is much larger than for the other cities; for this reason the results in section 5 are only shown and discussed for the city centre and Greater Paris. Toulouse has the smallest number in terms of number of buildings, but a far higher percentage categorised as detached low-rise (pd), confirming its low density. High-rise buildings (bgh) are rare in France; continuous and discontinuous rows of mid-rise buildings (icif and icio) are the most prevalent building types in city centres. Extended low-rise buildings (ba) account for around 10% and informal urbanization is almost non-existent (~1%) compared to other parts of the world.

3. Deriving local climate zones using administrative datasets

To classify each RSU into one of the 17 LCZ types using the MApUCE database, indicators at the RSU and building scales are used to calculate four types of land cover: built-up areas, water bodies, vegetated areas and impervious surfaces. The process is complicated by the nature of RSUs, which vary in size (Fig. 2) and encompass one or more types of land-cover; to cope with the latter a semi-automatic method based on cluster classification has been developed.

When applying a classification method, three key aspects will determine the results: the limits and specificities of the dataset (section 3.1), the hierarchy of the data treatment (section 3.2) and the definition of cut thresholds for each parameter (section 3.3). Each of these is discussed in turn.

3.1. MApUCE database specificities relevant for the LCZ classification

The various morphological indicators included in the MApUCE database exhibit different levels of precision and quality that will affect the LCZ methodology. These specificities include the following:

- The position of water bodies and buildings is for the most part well represented in the IGN-BD TOPO. For this reason, the indicators *Building number*, *Building density*, *Building height*, and *Water fraction* are mostly reliable and will therefore be given priority in the LCZ classification.
- The representation of impervious surfaces in the IGN-BD TOPO is not comprehensive. Main roads are included, but pavements and car parks are missing. The impervious surfaces are therefore underestimated in the MApUCE database.
- The morphological indicators are calculated at the scale of the RSUs, which are roughly delimited by roads. As a result, some building types, such as the "Informal Building" class, are unlikely to become the dominant building type at RSU scale.
- The MApUCE database lacks information on buildings with very large energy throughput like power plants, steel or aluminium factories, garbage incineration plants, and so on. For this reason, we cannot identify heavy industry (LCZ 10) using the MApUCE database only.
- The urban vegetation was retrieved via SPOT 6–7 satellite images provided by EQUIPEX-GEOSUD (http://ids.equipex-geosud.fr/) at 1.5m × 1.5 m resolution between 2015 and 2017. Using the methodology and software developed by Crombette et al. (2014), the fusion of panchromatic and multi-spectral band images, it provided high vegetation (trees and bushes), low vegetation (grass) and no vegetation classes. No vegetation can correspond to bare ground or impervious surfaces. The available vegetation data thus

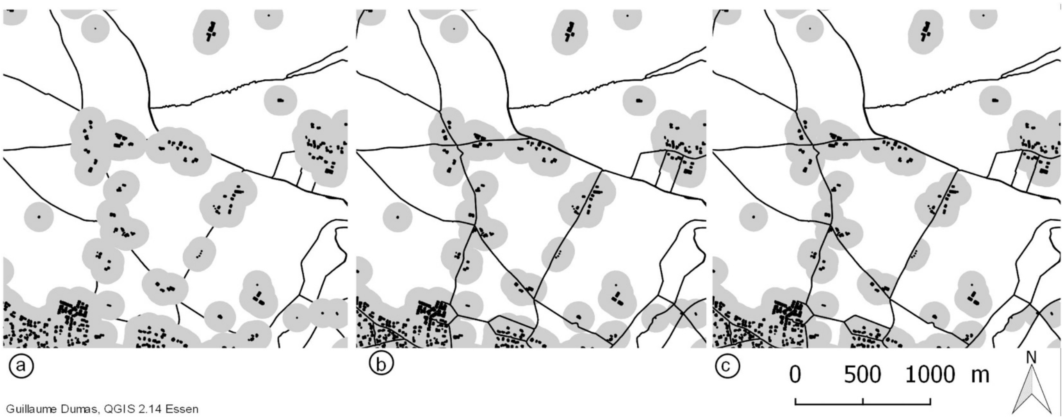


Fig. 3. Rules applied to outlined built areas inside the RSU. Black lines correspond to the RSU limits, grey stain corresponds to buffer surface and black stain to building footprint.

allows users to distinguish three LCZ types: dense trees (LCZ A), scattered trees (LCZ B) and low plants (LCZ D), but not bush/scrub (LCZ C) and bare ground (LCZ F). Moreover, the satellite images available are not the same for each city and this can have a strong impact on the vegetation representation. Images for Toulouse were taken in June, for Nantes in September and for Paris in May (West side) and August (East side).

The specificities of the MApUCE database will be taken into account in the following classification steps.

3.2. Hierarchy of the data treatment

The hierarchy proposed for the data treatment is as follows: first the RSUs with buildings are classified, followed by those that are mainly composed of water, roads and vegetation respectively.

The built-up RSUs are prioritised because of the reliability of these data and the need for an accurate discrimination among the urban LCZ types (LCZ 1 to 10). Water cover (LCZ G) is easily obtainable as it is well defined in the MApUCE database. Impervious areas (LCZ E) are also treated before vegetation areas due to the high number of parking spaces, roads or cemeteries with trees. Vegetation areas are treated at the end, once all other covers are considered well represented. Non-classified RSUs do not clearly fit into any category, so during data processing they are called ‘residue’ and are assigned to LCZ D, Low Plants.

3.3. Data pre-treatment

Small and completely built-up RSUs are mostly situated in city centres; by comparison large RSUs are mostly situated in the outskirts and are often composed of a mix of vegetated and built-up areas. In order to avoid over-classification of these large RSUs into vegetated LCZs, data pre-processing is applied to all RSUs with buildings to bring out the artificial surface.

The objective is to identify the built areas. For this purpose, a buffer of 100 m around the building perimeter is applied to the building scale database. The choice of this radius is arbitrary but justified by the micro-climatic effect of single buildings (Schmid et al., 1991) when including the urban environment around them (car parks, roads and public places). These new polygons are intersected with the RSU limits in order to define urbanized areas inside the RSU (Fig. 3.1). This makes it possible to evaluate the compactness of the built area within an RSU and better determine the LCZ density class while nevertheless retaining the MApUCE minimal surface unit, the RSU.

Some rules were applied: within the RSU, each buffer zone that doesn't touch another one from the same RSU becomes a unique built polygon, Fig. 3(a). Each polygon receives a unique ID, and the *Built density*, the *Number of buildings*, the *Mean building height*, the *mean Minimum distance*, the *median Minimum distance* and the sum of each *Building typology* area are recalculated for each polygon. Any polygon that after being intersected by the RSU shape (black line in Fig. 3(b)) has no building on it, is integrated into the vegetated shape (visible in the right part of Fig. 3(c)).

3.4. Characterization of built areas (LCZ 1 to 10)

To characterize these built polygons and attribute an LCZ type, eight parameters are used (Fig. 4).

Even if some links can be directly made between the 11 *Building typologies* defined in MApUCE and the ten “urbanized” LCZs (Tornay et al., 2017) – for example the “Informal building” corresponds to LCZ 7– this link is not unequivocal for the other LCZ types. Besides, the M*ajority Building typology* cannot be the most representative within the RSU when thinking in terms of building morphology. To avoid this problem, *Building typologies* within an RSU are grouped together to obtain four classes based on their

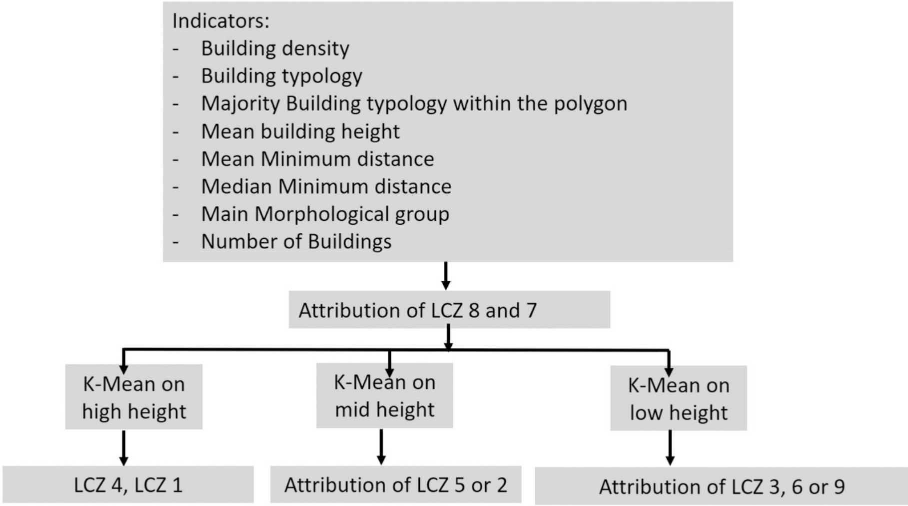


Fig. 4. Work-flow in the LCZ attribution process.

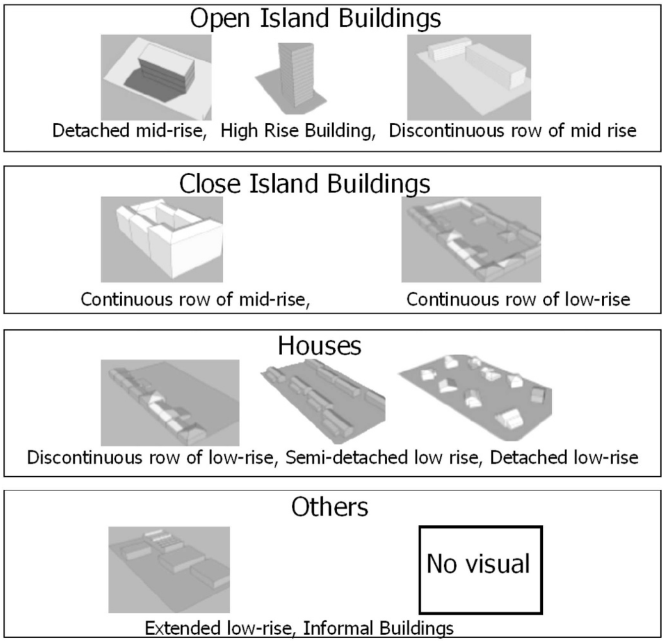


Fig. 5. Morphological groups used to verify if the Majority Building typology corresponds to the majority morphology. These groups are also used to identify LCZ 8 and LCZ7.

morphology (Fig. 5). The morphological groups are: open island buildings (*Building typologies*: Icio, Id and Bgh), closed island buildings (Icif and Pcif), houses (Psc, Pd and Pcio) and extended low-rise buildings. The surface fraction of these morphological groups is compared to the *Majority Building typology* indicator and a reclassification is applied if its value is higher. After this step, if "Extended low-rise" or "Informal building" is the main type of building, then LCZ 8 or LCZ 7 respectively are attributed.

Once these two LCZs are allowed (Fig. 4), the building typology is not used anymore and remaining urban polygons are separated using the *Mean building height* into the three height groups defined by SO12 (Low-rise: 0–10 m, Mid-rise: 10–25 m and High-rise: > 25 m). When exploring the relation between the LCZ classes and the *Building density* for the three case studies (Toulouse, Nantes and Paris), a high number of RSUs lie outside of the thresholds fixed by the SO12 *LCZ type table look-up* (Fig. 6). These thresholds are sometimes lower and often too high. For example, this is the case for those RSUs where the Majority Building typology is "Extended low-rise building", that often have values of *Building density* under 30%. This is not surprising, as the generic ranges

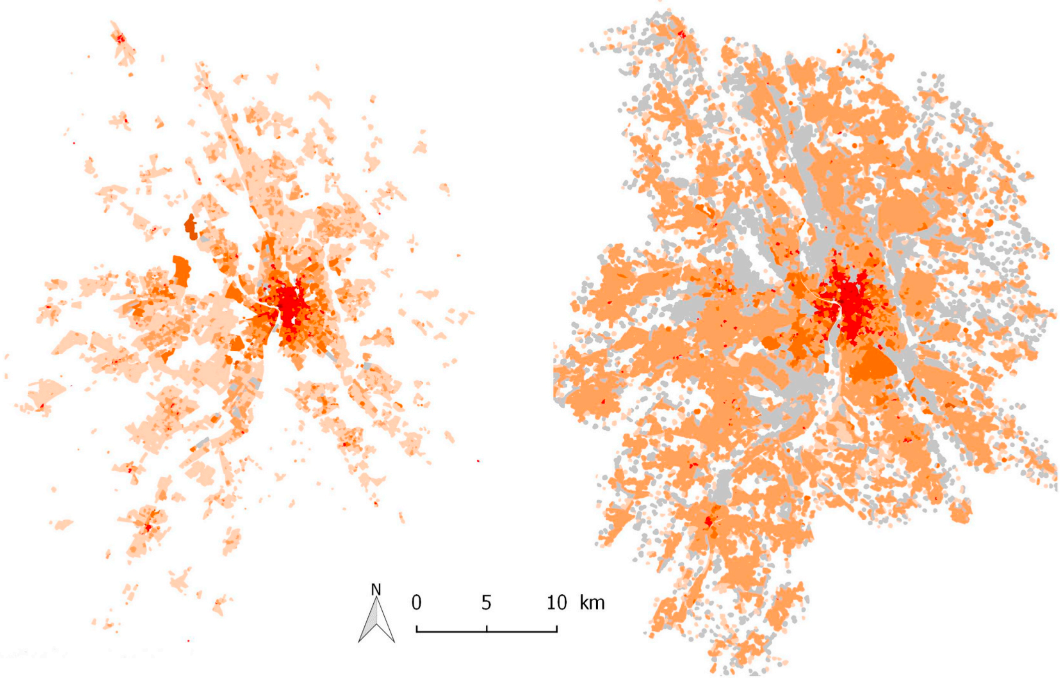


Fig. 6. Built areas (LCZ1 to LCZ10) classification for Toulouse using: left, thresholds fixed by the Stewart and Oke's LCZ type table look-up; right, thresholds obtained through the proposed supervised statistical method proposed in section 3.5.

provided by SO12 are based on expert knowledge and existing inventories, and they are not meant to capture the heterogeneity in the global real world.

3.5. A semi-automatic statistical method to obtain characteristic LCZ thresholds for each city

The issue now is to find pertinent indicator thresholds to dispatch the urban polygons in three, two and two groups for the Low-, Mid- and High-rise classes defined by SO12 but taking into account the building spatial configuration (*Building density*, *Distance between buildings*) characteristic of French urban morphologies.

For that purpose a semiautomatic statistical method was developed based on cluster analysis. In this study the statistical k-means method is used. This clustering method is convenient here because the final number of clusters is already fixed by the definition of LCZ types (three groups for low-rise LCZ types and two groups for mid-rise and high-rise types). In the k-means method a distance measurement determines how the similarity of two elements is calculated when forming the clusters. Here the Euclidean distance is used.

To define the cluster structure the following indicators at the RSU scale are used: the *Building density,* the *Mean* and the *Median of the Minimum Distance between buildings*. To avoid differences in scale among the variables, the mean and standard deviations are used to normalize their values. The indicators characterizing the distance between buildings are needed to differentiate between LCZ 6 and LCZ 9 because two areas with the same building density can present very different spatial arrangements of buildings (Fig. 7).

Exploratory tests showed that, when applying the k-means method, the classification is better if we impose more than just three classes to be re-grouped afterwards. For this study 10 groups were applied for the high-rise buildings, 10 groups for the mid-rise buildings and 20 groups for the low- rise buildings. A supervised group in final 3, 2 and 2 groups for, respectively, low-, mid- and high-rise classes proved a significant improvement and a finer classification, in particular for low-rise LCZ 3, 6 and 9 types.

The k-means method combined with a post supervised reclassification results in LCZ classes with thresholds that are characteristic for each site, thereby ensuring a better classification than if a single (and arbitrary) fixed threshold were applied. For example, in WUDAPT-L0, “compact” areas correspond to a *Building density* over 40%. The resulting mid-height groups for Nantes, Toulouse and Paris, treated with the same statistical method, are 36%, 37% and 33.8% respectively. Fig. 6 presents the classification for Toulouse when using SO12 LCZ type table look-up versus the result using the proposed supervised statistical method.

Fig. 7. Example of built areas (in blue and green) with similar building density (3%) but very different spatial configuration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.6. Thresholds for water, pavement and vegetation areas (LCZ A to G)

LCZ G is directly related to the *Water fraction indicator* based on the IGN BD TOPO as described in section 2. The RSU is classified here as LCZ G if the water fraction is higher than 50% of the RSU. Work is more arduous for paved surfaces, LCZ E, the road fraction is not reliable enough due to the linear definition of this indicator. It doesn't take into account the pavements or some car parks. Here the threshold value is fixed at 50% for the *Road fraction indicator,* considering that in reality it corresponds to over 60 to 70% of impervious surfaces in the RSU. Concerning the vegetated LCZ, with regard to the percentage of each vegetation class ("trees and bushes" and "grass"), if trees and bushes account for > 60%, the Dense trees LCZ A class is attributed. If both vegetated classes are over 40%, Scattered trees LCZ B class is attributed. Polygons where no buildings stand and the amount of high and low vegetation isn't enough to be classified are considered residual and are attributed to the LCZ D Low Plants class.

3.7. Final LCZ MApUCE maps

Fig. 8 shows the LCZ classification for Paris, Toulouse and Nantes based on administrative database from the MApUCE project. These maps reveal a typical European morphological structure. A denser city centre (mostly LCZ 2) is surrounded by large urban suburbs of residential housing (LCZ 6). The commercial and industrial developments (LCZ 8 and 10) are mostly concentrated along certain axes, which for the most part represent major traffic routes. However, there are also distinct differences. For instance, the dense urban core of Paris is large compared to the other two cities, while for Toulouse a large extensive development can be seen, but relevant patches of compact low rise (LCZ 3) and open mid rise (LCZ 5) can also be found adjacent to the town centre.

4. WUDAPT LCZ maps

The WUDAPT project gathers and organises urban data by level of detail. The lowest level of data (Level 0 or L0) breaks down urban regions into LCZ types using a supervised classification method applied to LANDSAT multi-spectral sensing data; the result of the process is a raster-based database in which each cell (100 m) has an LCZ value. Higher levels of detail in WUDAPT will provide

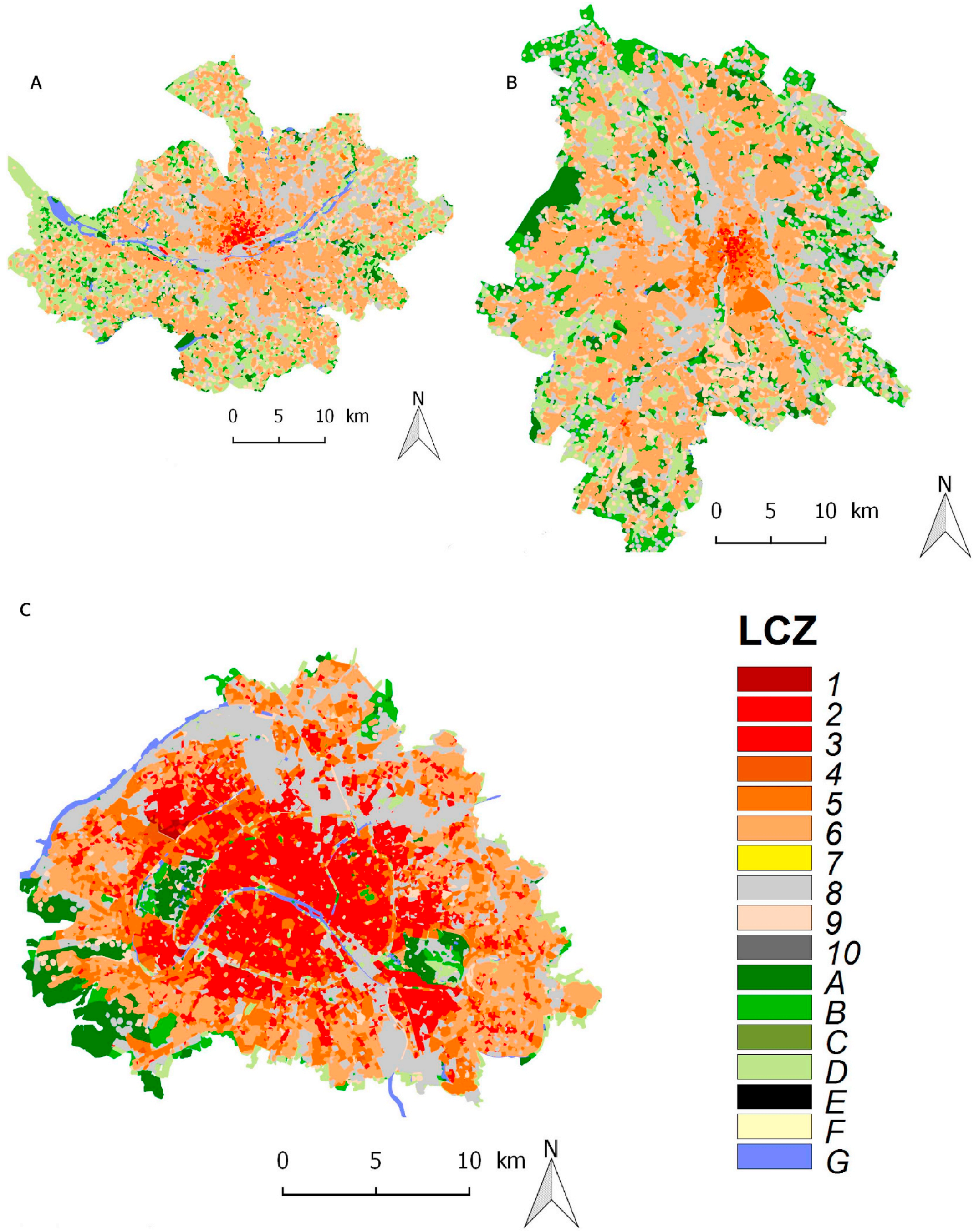


Fig. 8. LCZ classification for Paris, Toulouse and Nantes respectively. Due to the larger spatial extent of the Paris Urban Unit, only the city centre and the Greater Paris area are displayed.

more data on the features (e.g. buildings and trees) that make up urban areas and are needed by sophisticated climate models. As the WUDAPT-LCZ data are fundamentally based on the opinions of experts, the product is tested for reliability using a bootstrapping method (Kaloustian and Bechtel, 2016) that generated a set of accuracy measures (Bechtel et al., 2017). While this approach can detect inconsistencies in the classification, it does not represent an independent assessment of the WUDAPT-LCZ product, which would require alternative sources of data (Bechtel et al., 2019), such as that available within MApUCE.

The structure and status of the WUDAPT project are described elsewhere (Ching et al., 2017); here we focus on the WUDAPT-LCZ data for the three case-study cities. Fig. 9 shows the results for the three case studies, Toulouse, Nantes and Paris; the class frequency

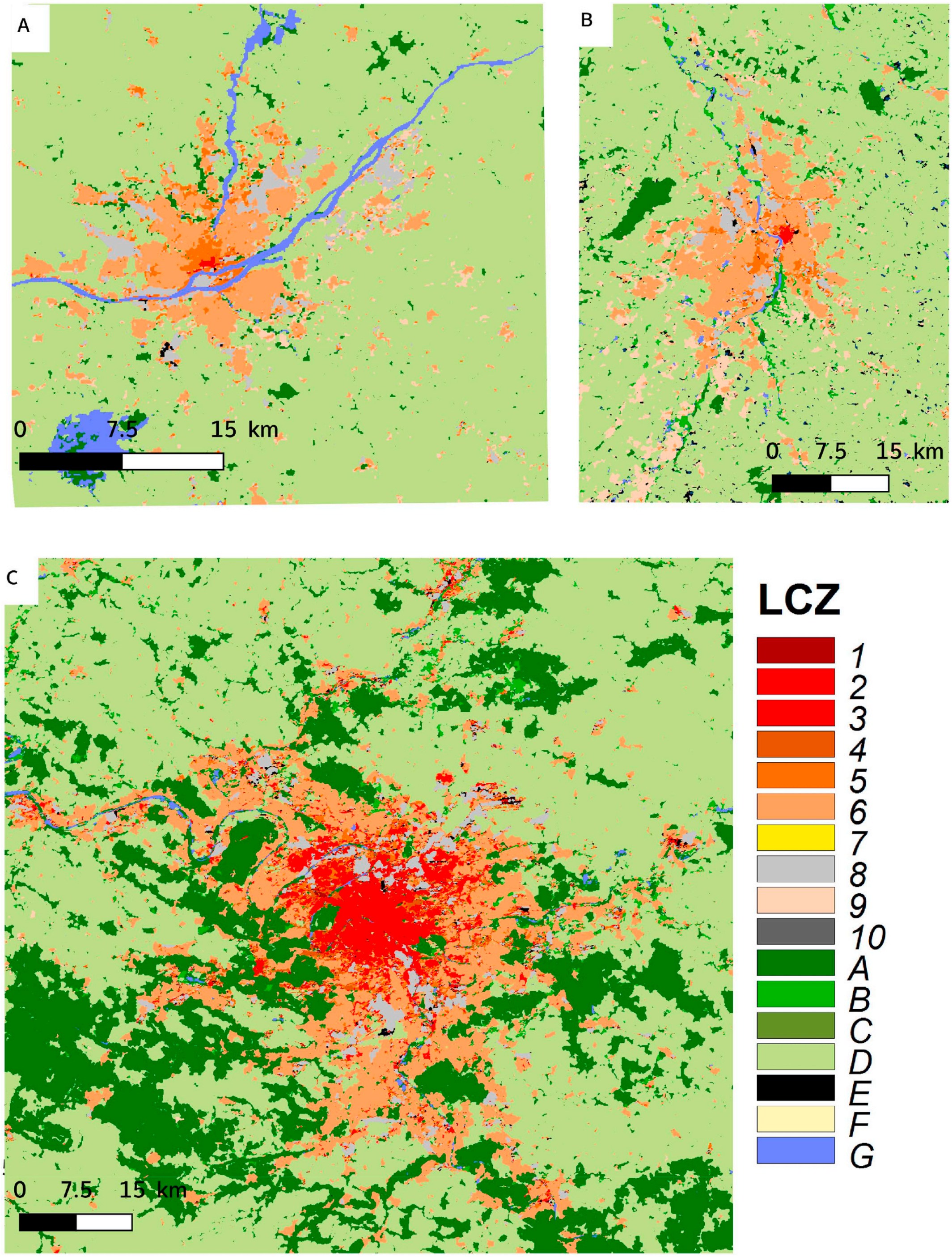


Fig. 9. WUDAPT-L0 Local Climate Zones maps at 100 × 100 m resolution for a) Nantes; b) Toulouse and c) Paris.

distributions (not shown) are dominated by the natural classes (in particular LCZ D, low plants) but it is important to recognise that this is largely a product of the size of the area under study. However, differences between cities are evident in the urban LCZ classes as well, Table 5: 45.7% of urban classes (LCZ1 to 10) in Paris are classified as compact (LCZ 1–3), compared to 0.37% in Nantes and 0.73% in Toulouse. Nantes, due to its old industrial port close to the city centre, has a large percentage (15.9%) of warehouse and industrial areas (LCZ8 and LCZ 10), compared to 10.62% in Paris and 9.89% in Toulouse. However, these findings are somewhat distorted by the large share of sparsely built areas (LCZ 9) in the Toulouse map (27.9%), which mostly consist of natural cover and are difficult to classify (Bechtel et al., 2016). Detailed cover fractions can be found in Table 5 in Appendix A.

Table 5
Percentage of total and individual LCZ surface for MApUCE (M) and WUDAPT-L0 (W) maps.

Tolouse					
LCZ	M (km^2)	M (%)	W (km^2)	W (%)	M-W (%)
1	0.77	0.12	0.00	0.00	100.00
2	6.51	1.03	2.72	0.58	58.22
3	1.52	0.24	0.72	0.15	52.63
4	0.44	0.07	0.54	0.12	−22.73
5	30.46	4.81	35.69	7.65	−17.17
6	379.64	59.99	250.57	53.72	34.00
8	169.26	26.74	46.11	9.89	72.76
9	44.29	7.00	130.07	27.89	−193.68
Total	632.89	100.00	466.42	100.00	26.30
Nantes					
LCZ	M (km^2)	M (%)	W (km^2)	W (%)	M-W (%)
1	0.00	0.00	0.00	0.00	0.00
2	3.74	0.84	1.00	0.36	73.26
3	2.46	0.55	0.02	0.01	99.19
4	0.18	0.04	0.82	0.29	−355.56
5	11.42	2.56	36.00	12.92	−215.24
6	279.12	62.65	111.03	39.85	60.22
8	103.89	23.32	44.35	15.92	57.31
9	44.73	10.04	85.37	30.64	−90.86
Total	445.54	100.00	278.59	100.00	37.47
Paris					
LCZ	M (km^2)	M (%)	W (km^2)	W (%)	M-W (%)
1	1.76	0.43	4.25	1.16	−141.48
2	98.59	24.26	114.15	31.28	−15.78
3	15.56	3.83	49.11	13.46	−215.62
4	6.90	1.70	35.53	9.74	−414.93
5	80.02	19.69	51.06	13.99	36.19
6	95.10	23.40	71.15	19.50	25.18
8	96.80	23.82	38.77	10.62	59.95
9	11.68	2.87	0.92	0.25	92.12
Total	406.41	100.00	364.94	100.00	10.20

4.1. Comparing MApUCE-LCZ with WUDAPT-LCZ

This section compares results produced by both approaches and highlights the potential and specificities of each. To that end, WUDAPT maps are vectorized in order to obtain polygons comparable to those in MApUCE maps. As stated in section 3, boundaries of vegetated areas for the MApUCE version are distorted by the RSU cutting definition. This comparison focuses only on built areas. Two types of analysis are presented, based on the total and individual LCZ surface and on a geographical analysis based on the *Building Density* and *Building Height* class.

A. Toulouse study case

Fig. 10a and Table 5 shows that total built areas are underestimated by 26.3% in the WUDAPT approach. This significant difference should be put in perspective, as the main differences appear in small suburban areas where there are few buildings and the choice of a 100-m buffer could contribute to overestimating built areas in these zones. Compact mid-rise class, LCZ 2, is underestimated in WUDAPT map and the city centre of Toulouse is the area the most impacted by this under-representation. In WUDAPT approach LCZ 6 is also under-represented comparing to MapUCE even if it has tendency to classify the large low-rise areas (LCZ 8) into LCZ 6. This is not the case with the MApUCE method. Nevertheless, those isolated big buildings that corresponds to agricultural activities and are then surrounded by big vegetated areas are classified by MApUCE method in LCZ 8 while it would be more suitable to be classified in LCZ9. For Toulouse, the WUDAPT approach catch those subtleties, but that strongly depends on the chosen training areas.

Sparsely built, LCZ 9, is less prevalent in MApUCE representation. The use of the mean and median minimum distances between buildings and the islet referential in the classification implies an under-representation of this typology compared with the WUDAPT

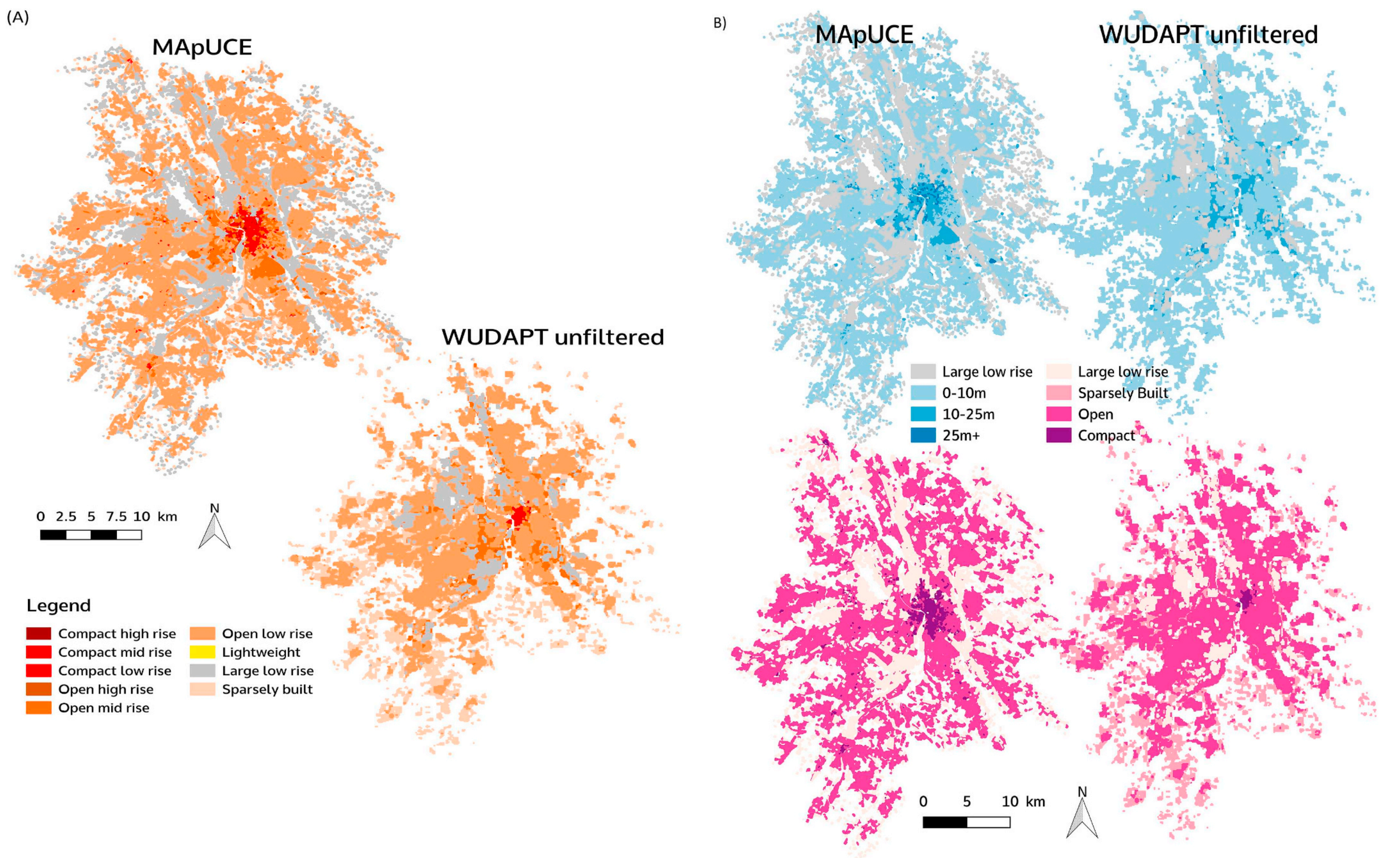


Fig. 10. Geographical comparison for Toulouse based on: (A) the built areas classification obtained through MapUCE and WUDAPT-L0 methods and (B) the Building Density and Building Height classes.

standards. French urban models for residential neighbourhoods are based on building distances of 5 to 15 m. It is of course possible to find buildings that are > 20 m apart, but it will be sporadic. That's why the representation of LCZ 6 and 8 is higher in the MApUCE method, and the LCZ 9 class is less prevalent than in the WUDAPT method.

When comparing by *Building height* or *Building density* classes, some differences clearly appear (Fig. 10b). The compact zones (LCZ 1, LCZ2 and LCZ3) are more prevalent on the MApUCE map and allow for the capture of small town centres, which is very important for modeling micro-climatic effects on town peripheries. The road structure to the north, south-east and south-west, appears clearly when analysing both height class maps. In this sense the MApUCE map shows better resolution and can capture the internal large avenues lined with high buildings represented by the internal red arcs in the centre of Toulouse.

B. Paris and Nantes case studies

Due to the size of the Paris Urban Unit (see figures in Appendix B), only the city centre and Greater Paris are analyzed here. As dense built areas are captured well by the WUDAPT method, and this area is highly artificial, both methods capture comparable urban extension. In terms of LCZ typologies, Paris is almost exclusively composed of Compact Mid Rise buildings, LCZ 2, over-estimated of about 15.8% compared to MApUCE. Differences in LCZ 5 and LCZ 6 extension, are quite similar, under representation of WUDAPT with 36.2% and 25.2% respectively. The biggest differences being found for LCZ 3 and LCZ 4 with a high over-representation by the WUDAPT methodology, 215% and 414% respectively. In France, most LCZ 4 classes correspond to areas called in French “les grands ensembles” (“large units”), a particular type of 1960s urban planning situated near the roadside belt. Even if training areas were taken from that kind of territory, the WUDAPT method classified zones between dense and open as LCZ 4 instead of LCZ 5.

The MApUCE method shows a balanced profile with four main LCZs: Compact Mid Rise, Open Mid Rise, Large Low Rise, and Open Low rise; predominant LCZs for the WUDAPT maps are: Compact Mid Rise and Open Low Rise.

For Nantes, built areas are particularly under-represented in the WUDAPT method, with a surface difference of 37%. Nevertheless, the LCZ distribution identified by both methods is not so different. The most frequent LCZs are the same (LCZ6, LCZ8 and LCZ9) but their hierarchy slightly differs. First LCZ 6 Open Low Rise, is the most frequent in both approaches. Then, LCZ 8 Large Low Rise, and lastly, LCZ 9 Sparsely Built are the most frequent in MAPUCE while is the opposite in WUDAPT. This is explained by the fact that, for Nantes, the WUDAPT methodology identified as LCZ9 some areas classified as LCZ6 and LCZ8 by MApUCE method. Finally, density classes (Compact, Open, Sparse, Large Low Rise) or height classes (High-rise, Mid-rise, Low-rise, Large Low Rise) are correctly spatialized for this territory.

4.2. Detailing LCZ indicators for French cities from the MApUCE database

In this section the analysis is focused on the LCZ ranges of urban canopy model parameters fixed on the table lookup in SO12. The objective here is to detail the typical borders of the interval of LCZ indicators for French cities.

Out of the ten indicators proposed in the table lookup in SO12, the MApUCE morphological DB calculates seven:

- The Mean building height
- The Building surface fraction
- The Pervious surface fraction
- The Impervious surface fraction
- The Aspect ratio H/W
- The Roughness class
- The Sky view factor

However, in the morphological DB in MApUCE there are no quantitative elements with which to estimate *Albedo*, *Surface admittance* and *Anthropogenic heat flux*.

In order to characterize the medium-size cities of the conurbation as well, the first five indicators were computed for both the central commune and the immediate periphery of each case study. In this study, the *Sky view factor* was not explored.

A) Indicator calculation

Two LCZ indicators already feature on the MApUCE DB, the *Mean building height* and the *Building surface fraction (*called *Building density* in the MApUCE DB*)*. The *Building density* corresponds in this study to the built surface with respect to the surface of the buffer polygon. This can be an entire RSU, for completely urbanized ones close to the city center, or a fraction of it for those of the rural areas that were divided into several zones during the buffering process. The *Mean building height* and the *Building density* are then recalculated for each polygon.

Table 6
Detailed LCZ distribution based on the central commune and the immediate periphery of each case study from the MApUCE database.

LCZ	Toulouse		Paris		Nantes	
	Centre (%)	Periphery (%)	Centre (%)	Periphery (%)	Centre (%)	Periphery (%)
1	1.31	0.03	0.00	1.03	0.00	0.00
2	25.77	1.23	78.03	37.30	15.12	0.42
3	2.83	1.03	0.78	14.60	8.24	3.11
4	1.38	0.22	2.80	1.78	0.63	0.05
5	10.26	3.51	9.09	16.30	12.55	3.04
6	44.65	78.62	1.25	16.94	53.92	73.13
8	11.53	11.27	3.83	9.59	7.36	10.53
9	2.28	4.08	4.22	2.47	2.18	9.74

The *Aspect ratio* corresponds to the *Mean building height* of the polygon divided by the *Mean minimum Distance between buildings*. As in Stewart and Oke's study, values higher than three are truncated.

The *Pervious and impervious surface fraction* are calculated using the vegetation maps at 1.5 m resolution combined with the buffer zones and the building footprints. In this case the "no vegetation" category is used. On the central commune and the immediate periphery where there is no unattributed land (or very little), we assume that this category accurately represents the mix of buildings and paved surfaces. To calculate the *impervious surface fraction* indicator, the building footprints are extracted from the buffer zones and zonal statistics are applied to the intersection of this information and the "no vegetation" category on the vegetation map. The *Pervious surface fraction indicator* over the buffer zone is calculated as 1- *impervious surface fraction - the building footprint surface.*

B) Presentation of results

Three levels of analysis are presented and discussed, in terms of: overall LCZ spatialization (frequency count of LCZ's), the mean indicator magnitude for the central commune and the communes situated in the immediate periphery (also called first crown), and finally the indicator variability at the LCZ scale.

In terms of overall LCZ spatialization (Table 6), it is possible to observe morphological differentiation linked to the administrative division. For example, for Paris, if the town centre is composed of 78% of LCZ 2 Compact Mid Rise, the first crown comprises 37% and the enterprise zones range from 3.8% to 9.5%. LCZ 5 and 6 account for a mere 9.1% and 1.25% respectively in the town centre, but increase to 16.3% and 19.9% in the first crown.

This type of analysis can be repeated on Toulouse and Nantes. This makes it possible to observe the overwhelming presence of LCZ 6 on the peripheral territory with 78.6% of the LCZs coupled with LCZ 8 at 11.2%. For the centre of Toulouse, the LCZ 6 increases to 44.6% and the LCZ 2 appears at 25.7%. For Nantes, we find the same phenomenon with LCZ 6 and 2 for the town centre accounting for 53.9% and 15.1% and an LCZ 6 predominant with 73.1% on the periphery. It is interesting to note that the profiles of the town centre and the periphery of Toulouse and Nantes are very similar. This can be seen in the general profile presented in the previous section, but when the territory is broken up it is all the more striking.

When focusing on the mean indicator magnitude for the central commune and the immediate periphery, we may observe that (Table 7):

- For the three sites and all LCZs, the *Mean building height* values fall on the "standard" thresholds fixed by SO12 because in our classification method the heights are fixed according to these LCZ thresholds (0–10 m 10–25 m and > 25 m). It can also be observed that buildings are consistently higher for the city centre than for the periphery. The variability seems to be higher for Paris, in particular for LCZ 4 and 8, than for Toulouse or Nantes.
- In general terms for these French cities, values of the *Building surface fraction* seem to be close to the standard ones (the mean value falls on the range) for LCZ 1, 2 and 3. Values are slightly smaller (the mean value ± σ, falls on the range) for LCZ 4, 5 and 6. And values are far from the standard ones (the mean value ± σ, does not fall on the range) for LCZ 8 and 9. Evidently differences may be observed between the cities, for example, differences in values for the city centre of Toulouse and for the periphery of Nantes are higher.
- The partition of impervious and pervious surfaces in the standard look-up table seems to be quite well represented (all the values fall on or are close to the standard range) for Paris and Nantes, in particular for LCZ 2, 3, 5 and 6. Large low-rise (LCZ 8) seems to be consistently more vegetated, in particular for Toulouse.
- From the *Aspect ratio* values, French cities seem to be denser than the average, in particular for central towns where higher values of H/W are found. The H/W values fall on the standard threshold for just one LCZ – LCZ1, situated on the peripheral communes in

Table 7
Mean value of the indicators and its standard deviation for the central commune and the immediate periphery from the MApUCE database. Green colour indicates that mean value falls on the "Standard Threshold" (ST) proposed by SO12. Orange that the mean value ± σ, falls on the ST and red colour that the mean value ± σ, does not fall on the ST.

TOULOUSE Indicators/LCZ	LCZ1		LCZ2		LCZ3		LCZ4		LCZ5		LCZ6		LCZ8		LCZ9	
	C	P	C	P	C	P	C	P	C	P	C	P	C	P	C	P
Mean Building Height (m)			15.65	13.55	8.52	7.37	29.67	31.71	14.28	13.92	7.03	5.74	8.60	7.40	5.53	5.57
σ			3.79	3.06	1.05	1.85	4.27	6.46	3.64	3.80	1.51	1.34	3.13	2.96	1.92	1.81
Theoritical Thresholds	>25m		10 to 25m		3 to 10m		>25m		10 to 25m		3 to 10m		3 to 10		3 to 10	
Building surface fraction (%)			49%	36%	41%	45%	23%	21%	19%	18%	22%	18%	17%	15%	2%	3%
σ			14%	8%	5%	9%	9%	8%	6%	6%	8%	7%	10%	11%	2%	3%
Theoritical Thresholds	40 to 60		40 to 70		40 to 70		20 to 40		20 to 40		20 to 40		30 to 50		10 to 20	
Impervious surf. fraction (%)			26%	21%	23%	24%	35%	26%	25%	23%	18%	12%	33%	30%	21%	17%
σ			10%	14%	9%	15%	19%	17%	15%	15%	11%	11%	16%	19%	19%	23%
Theoritical Thresholds	40 to 60		30 to 50		20 to 50		30 to 40		30 to 50		20 to 50		40 to 50		0 to 20	
Pervious surf. fraction (%)			25%	43%	36%	31%	42%	53%	56%	59%	61%	70%	51%	55%	78%	80%
σ			17%	18%	11%	18%	20%	23%	15%	16%	14%	13%	21%	24%	20%	23%
Theoritical Thresholds	0 to 10		0 to 30		0 to 30		30 to 40		20 to 40		30 to 60		0 to 20		60 to 80	
Aspect ratio H/W (0 to 3)			2.91	2.52	2.89	1.62	2.78	1.97	1.91	1.58	1.66	0.97	1.00	0.76	0.48	0.31
σ			1.46	1.28	1.02	0.98	1.47	1.06	0.98	1.00	0.87	0.56	0.69	0.58	0.56	0.40
Theoriticals Thresholds	> 2		0.75 to 2		0.75 to 1.5		0.75 to 1.25		0.3 to 0.75		0.3 to 0.75		0.1 to 0.3		0.1 to 0.25	
PARIS Indicators/LCZ	1		2		3		4		5		6		8		9	
	C	P	C	P	C	P	C	P	C	P	C	P	C	P	C	P
Mean Building Height (m)		29.35	19.07	15.02	8.40	7.89	30.86	33.84	18.63	14.98	6.60	7.31	11.32	10.37	4.88	5.53
σ		5.93	3.94	3.47	1.54	1.51	11.82	8.92	3.99	3.69	2.12	1.49	8.05	4.38	2.04	1.84
Theoritical Thresholds	>25m		10 to 25m		3 to 10m		>25m		10 to 25m		3 to 10m		3 to 10		3 to 10	
Building surface fraction (%)		60%	49%	48%	45%	53%	23%	19%	19%	19%	14%	22%	20%	37%	1%	2%
σ		29%	10%	21%	8%	26%	8%	13%	7%	6%	11%	7%	19%	25%	2%	2%
Theoritical Thresholds	40 to 60		40 to 70		40 to 70		20 to 40		20 to 40		20 to 40		30 to 50		10 to 20	
Impervious surf. fraction (%)		34%	46%	41%	48%	40%	60%	56%	58%	50%	46%	52%	51%	53%	53%	57%
σ		26%	10%	20%	10%	24%	20%	24%	21%	25%	31%	23%	25%	24%	37%	36%
Theoritical Thresholds	40 to 60		30 to 50		20 to 50		30 to 40		30 to 50		20 to 50		40 to 50		0 to 20	
Pervious surf. fraction (%)		6%	5%	11%	7%	6%	18%	25%	23%	30%	40%	26%	29%	10%	46%	41%
σ		11%	9%	14%	9%	10%	19%	23%	23%	25%	37%	24%	31%	16%	37%	36%
Theoritical Thresholds	0 to 10		0 to 30		0 to 30		30 to 40		20 to 40		30 to 60		0 to 20		60 to 80	
Aspect ratio H/W (0 to 3)		2.94	2.99	2.98	3.00	2.98	2.80	2.70	2.49	2.71	1.26	2.81	1.46	2.51	0.42	0.97
σ		1.49	1.47	1.32	1.39	1.12	1.45	1.42	1.34	1.12	1.04	0.82	1.17	1.09	0.41	0.99
Theoriticals Thresholds	> 2		0.75 to 2		0.75 to 1.5		0.75 to 1.25		0.3 to 0.75		0.3 to 0.75		0.1 to 0.3		0.1 to 0.25	
NANTES Indicators/LCZ	1		2		3		4		5		6		8		9	
	C	P	C	P	C	P	C	P	C	P	C	P	C	P	C	P
Mean Building Height (m)			14.62	12.28	7.25	5.93	37.67	31.20	14.15	13.00	6.02	4.72	7.60	6.56	4.33	4.25
σ			3.25	2.81	1.63	1.93	22.02	8.77	3.16	2.33	1.67	1.26	2.65	2.20	1.58	1.34
Theoritical Thresholds	>25m		10 to 25m		3 to 10m		>25m		10 to 25m		3 to 10m		3 to 10		3 to 10	
Building surface fraction (%)			44%	36%	40%	44%	14%	5%	18%	16%	21%	14%	21%	14%	2%	1%
σ			11%	6%	7%	9%	6%	1%	6%	6%	7%	7%	14%	10%	2%	1%
Theoritical Thresholds	40 to 60		40 to 70		40 to 70		20 to 40		20 to 40		20 to 40		30 to 50		10 to 20	
Impervious surf. fraction (%)			48%	46%	48%	44%	68%	43%	57%	59%	49%	48%	59%	44%	48%	34%
σ			9%	13%	7%	15%	11%	0%	14%	17%	13%	21%	17%	24%	33%	26%
Theoritical Thresholds	40 to 60		30 to 50		20 to 50		30 to 40		30 to 50		20 to 50		40 to 50		0 to 20	
Pervious surf. fraction (%)			7%	17%	11%	13%	18%	52%	23%	24%	29%	38%	19%	42%	49%	65%
σ			6%	12%	6%	13%	13%	0%	15%	18%	15%	24%	19%	28%	34%	26%
Theoritical Thresholds	0 to 10		0 to 30		0 to 30		30 to 40		20 to 40		30 to 60		0 to 20		60 to 80	
Aspect ratio H/W (0 to 3)			2.95	2.33	2.97	2.90	2.14	0.42	2.36	1.61	2.55	1.55	1.51	0.94	0.57	0.38
σ			1.47	1.05	1.21	1.47	1.18	0.30	1.27	1.09	1.11	1.02	1.10	0.75	0.61	0.53
Theoriticals Thresholds	> 2		0.75 to 2		0.75 to 1.5		0.75 to 1.25		0.3 to 0.75		0.3 to 0.75		0.1 to 0.3		0.1 to 0.25	

Paris. While values are not so far from the standard ones in Toulouse and Nantes, values for LCZ 3, 4, 5, 6 and 8 are invariably higher for the periphery of Paris.

An example of sheets presenting ranges of urban canopy parameters according to the framework proposed by SO12 can be found in Appendix C.

Indicator variability at the LCZ scale. Fig. 11 clearly shows the advantage of shifting from WUDAPT level 0 to WUDAPT level 1. Effectively, in actual fact there is an intra-LCZ variability that should be described to properly evaluate the impact of the urban

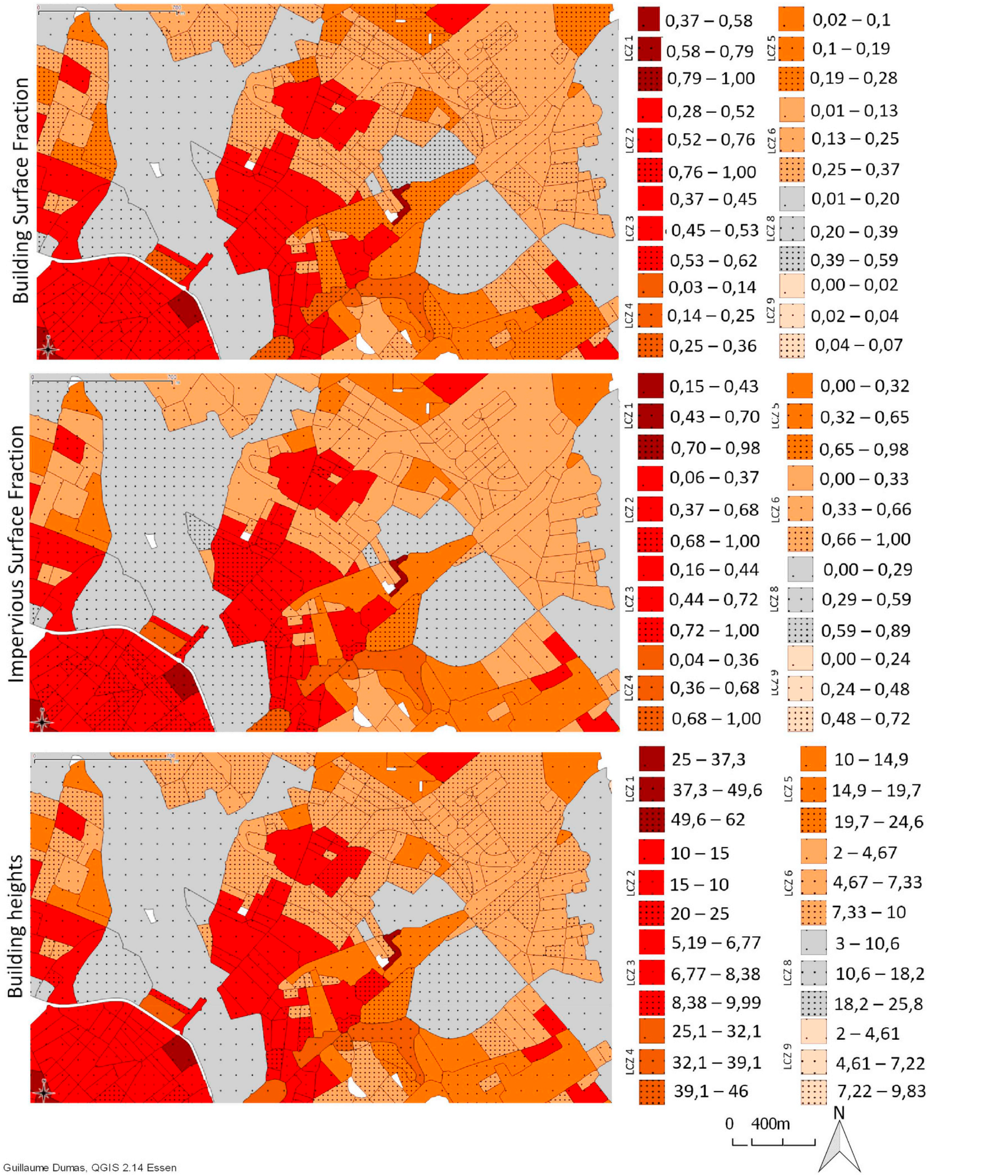


Fig. 11. Intra-LCZ variability in a neighbourhood of Toulouse.

structure on the atmospheric boundary layer. In this example, for a neighbourhood situated in the north-western part of Toulouse, it can be clearly seen that the *Mean building height*, the *Mean building fraction* and the *Impervious surface fraction* for a given LCZ varies without correlation between the parameters.

5. Conclusions

This paper aims to contribute to the scientific discussions of the pertinence and limits of the various ways ("bottom-up" vs "top-down") of obtaining Local Climate Zone maps.

From a methodological point of view the paper presents an original semi-automatic method based on cluster classification to obtain GIS-based Local Climate Zone maps. For this purpose, a rich administrative dataset developed during a national funded project, the MApUCE project, is used to obtain MApUCE-LCZ maps for three French case studies, Paris, Nantes and Toulouse. The choice was made to work with vector data that offer better comprehension of the boundaries of urban elements than rasterized data and make it easier to share results with practitioners. The work-flow in the classification is as follows: first areas with buildings are classified and then those composed mainly of water, roads and vegetation respectively. The paper also presents maps obtained through the WUDAPT level 0 methodology and comment the differences observed.

Geographical analysis based on the comparison between MApUCE and WUDAPT results for total and individual LCZ surface and *Building Density* and *Building Height* indicators shows that:

- Compared to the MApUCE approach, in general terms, WUDAPT L-0 under-represent urbanized areas.
- Main differences appear in the small suburban areas where there are few buildings. These differences are mainly concentrated on LCZ 6, 8 and 9.
- Even if the city cores are quite well represented in the WUDAPT approach in terms of urban extension, MApUCE has better resolution for these areas and the structures of town centres are consequently better captured.
- In the WUDAPT approach, for Toulouse key differences are concentrated on LCZ 6. For Paris LCZ 4 is highly over-represented; and for Nantes, even if LCZ distribution were comparable, built areas are particularly under-represented.

SO12 provided general LCZ ranges for ten urban canopy model parameters. The morphological database in MApUCE allows users to verify and detail these ranges for French cities and to analyse their intra-LCZ variability. Main findings are:

- Relatively significant morphological differentiation was observed between the central commune and the immediate periphery. The geographical domain used in numerical simulations of Urban Heat Island, and more generally urban micro-climate, used to be larger than the size of the city, in order to capture the local meteorological interactions between the city and the countryside. Therefore, it seems important to widen the urban structure analysis to peripheral cities when preparing the urban surface input data for climatic simulations.
- As already pointed out by other authors, to obtain satisfactory LCZ classification, it is important to take into account specific regional features concerning the borders of the intervals of the LCZ physical properties. When available, a high resolution administrative database is a solution; but if it is not available, other solutions exist based on manual sampling from field campaigns or crowd-sourcing and multi-source satellite images.
- Even if the LCZs are supposed to have a homogeneous climatic footprint on the near-surface atmosphere, their characteristic size is over 500x500m and, for modeling purposes, the internal variability of the building structure inside this area must be evaluated and represented in the urban surface input data. Indeed, while WUDAPT level 0 morphological parameters may be sufficient for Numerical Weather Prediction and Very-High resolution models, both typically using 1 to 4 km resolutions, hectometric-scale and Large eddy Simulation models are now currently used in the urban climate community for micro-climate studies, and a finer description of the city internal structure is desirable.

The advantage of the use of the Local Climate Zone classification is the use of the same referential on the urban structure comprehension that allows comparison of morphological portraits. On this subject, Toulouse and Nantes are two cities which possess many similar morphological characteristics. The difference between the two has more to do with materials and their characteristics: white stones and light soils in Nantes, red brick and dark pavement for Toulouse, for example. It is at this level of detail that the limits of the morphological database in MApUCE or WUDAPT-L0 maps can be seen, and it is for this reason that a complementary architectural database was also developed in MApUCE. This architectural database is based on building use, building construction date and geographical location (Tornay et al., 2017) and makes it possible to architecturally differentiate cities and areas within a city during modeling exercises. The need for increase in architectural description arises in the WUDAPT community, and two ways are currently used to fill this gap: architectural expertise (as in Tornay et al., 2017), that can be organized at global scale, and crowd-sourcing, in order to gather some information for many buildings in each LCZ for many cities.

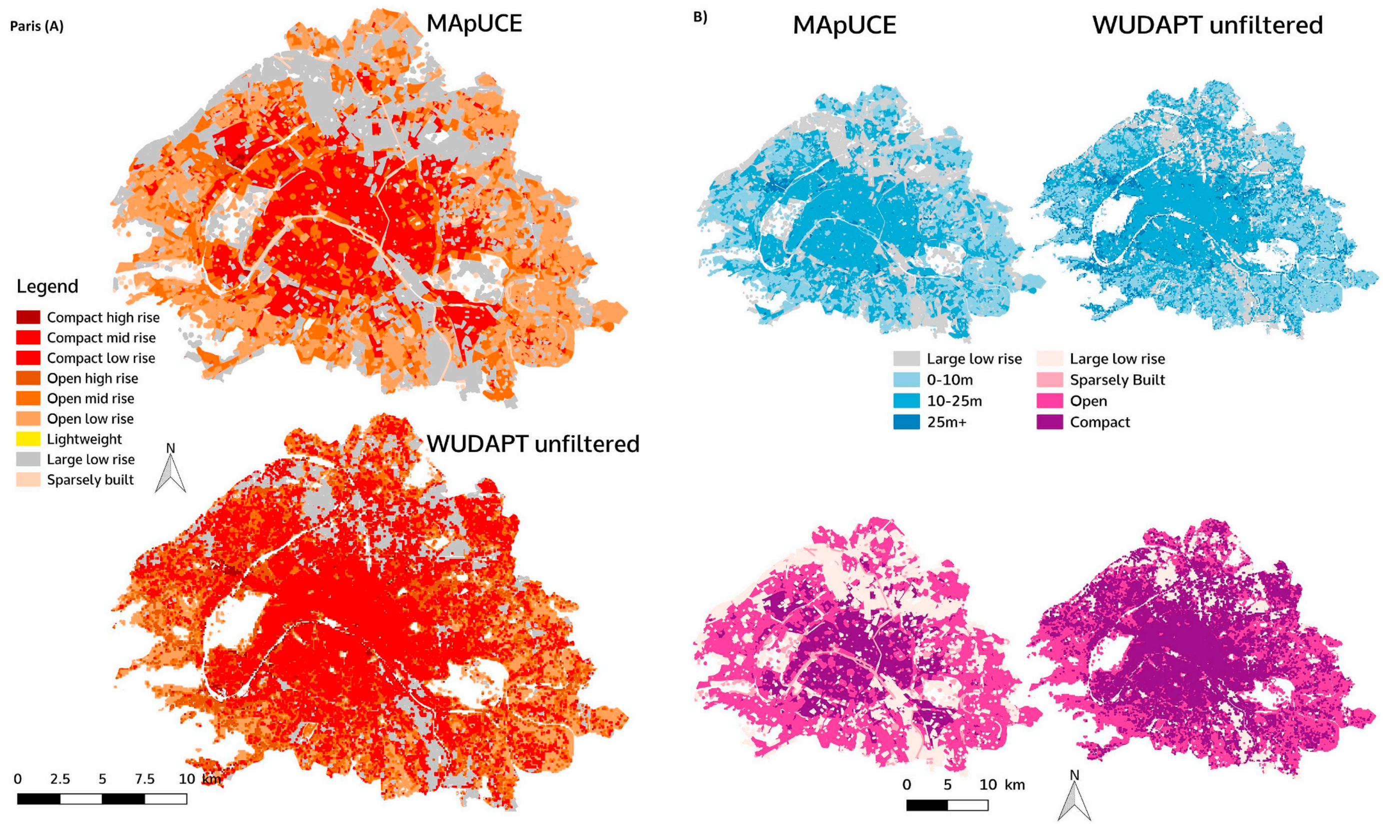
Acknowledgements

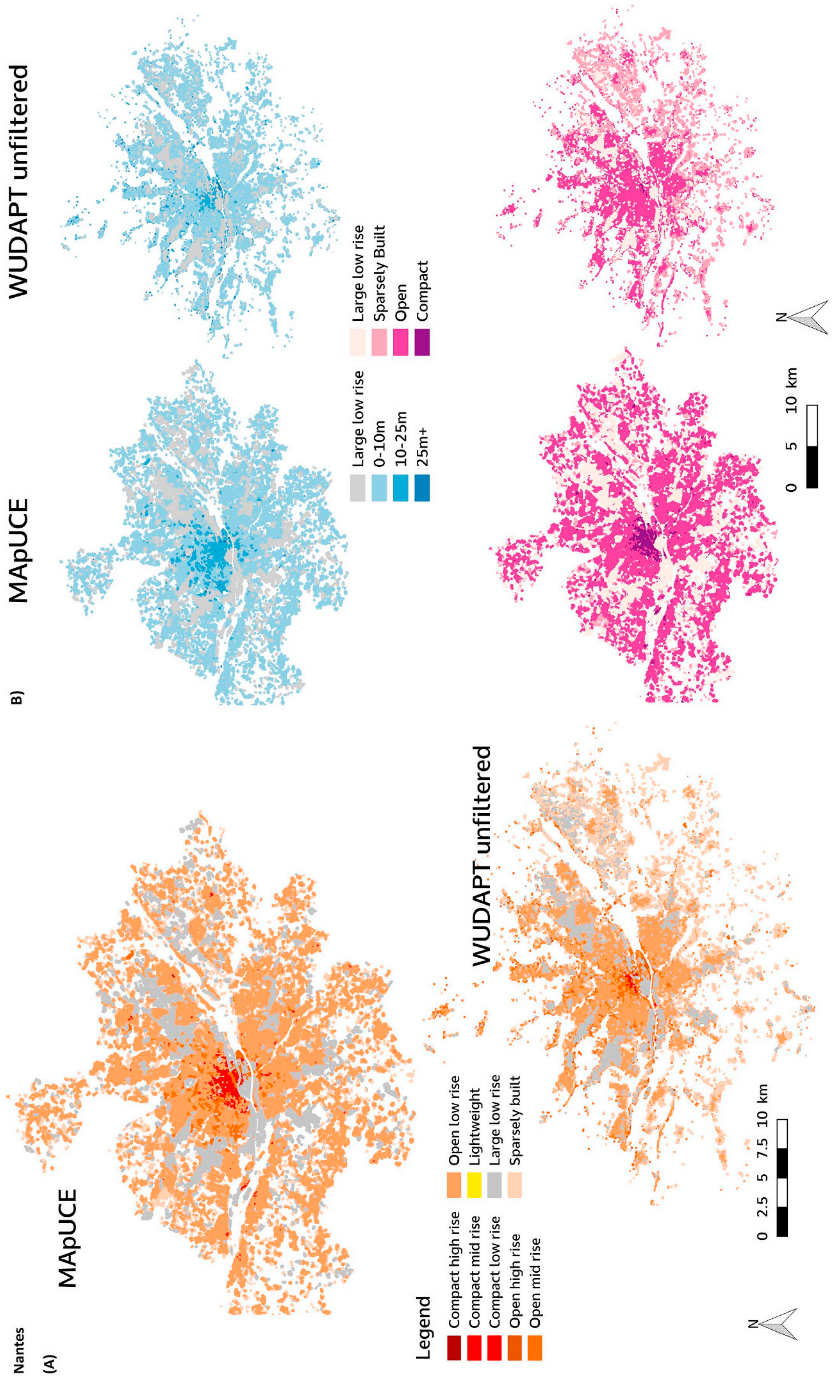
This work was funded by the French National Research Agency (Agence Nationale de la Recherche) under grant #ANR-13-VBDU-004. It was partly supported by the Cluster of Excellence 'CliSAP' (EXC177), University of Hamburg, funded through the German Science Foundation (DFG)

Appendix A. WUDAPT maps meta data and evaluation results

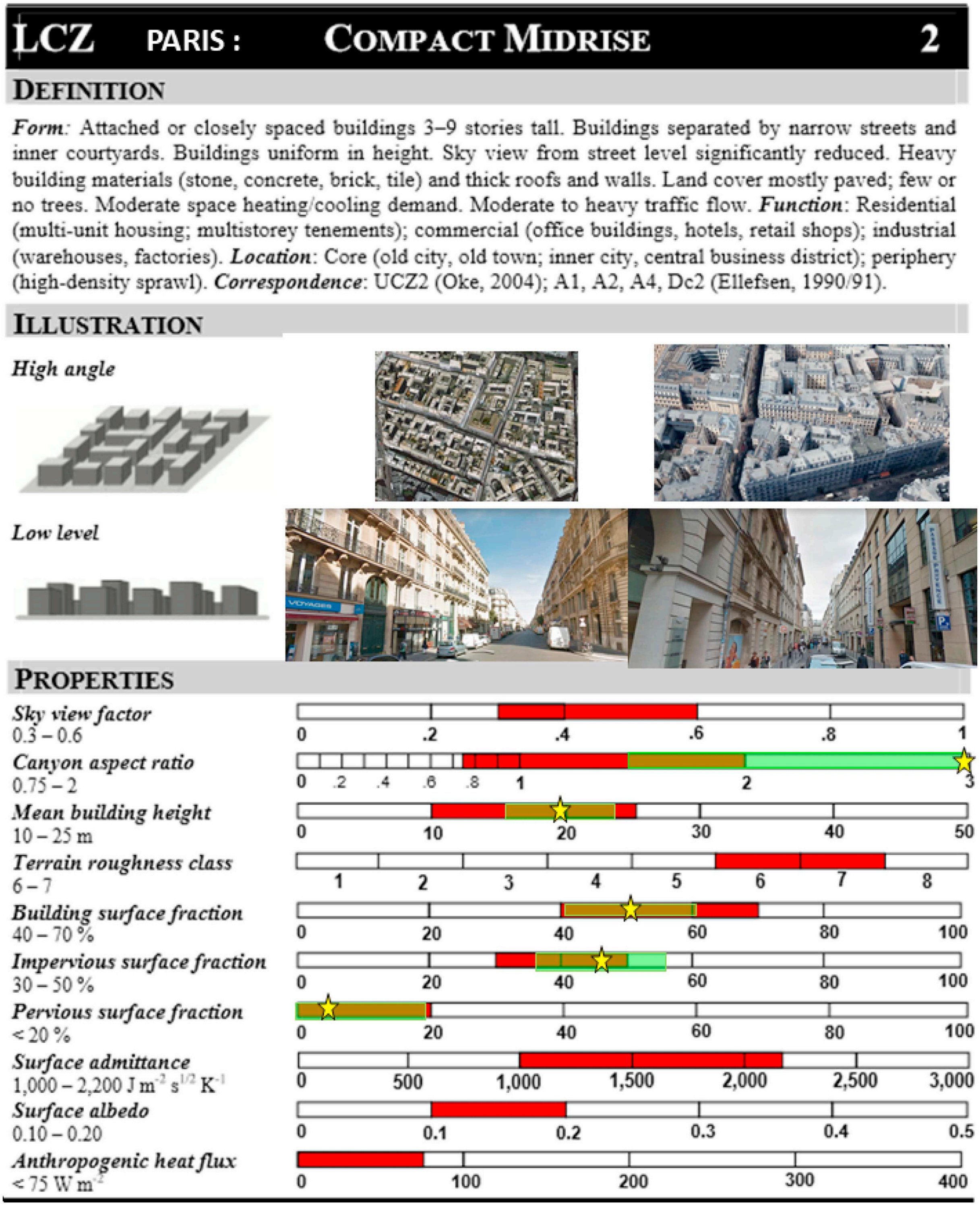
<table>
<tr><th></th><th>Nantes</th><th>Paris</th><th>Toulouse</th></tr>
<tr><td>Meta data</td><td></td><td></td><td></td></tr>
<tr><td>Docversion</td><td>0.241</td><td>0.241</td><td>0.241</td></tr>
<tr><td>Training data creator</td><td>Gwendall Petit</td><td>Guillaume Dumas, WUDAPT team</td><td>Julia Hidalgo</td></tr>
<tr><td>Training data ID</td><td>Nantes(2018-07-26)_{Gwendall_Petit,Benjamin_Bechtel}</td><td>Paris_GuillaumeDumas_MF_20161001</td><td>Toulouse_JuliaHidalgo_20160119</td></tr>
<tr><td>Satellite imagery</td><td>LANDSAT:
LC82010272013246LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC82010272013342LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC82010272014073LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC82010272014105LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC82010272014137LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 ></td><td>LANDSAT:
LC81990262014139LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC81990262015270LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 ></td><td>LANDSAT:
LC81980302014196LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC81980302014244LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC81980302015103LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC81980302015215LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 > ,
LC81980302015263LGN00 < B1|B10|B11|B2|B3|B4|B5|B6|B7 ></td></tr>
<tr><td>Acquisition times</td><td>years: 2013–2014 months: < Apr,Dec,Mar,May,Sep ></td><td>years: 2014–2015 months: < May,Sep ></td><td>years: 2014–2015 months: < Apr,Aug,Jul,Sep ></td></tr>
<tr><td>Training area sizes</td><td>N (size) of traing areas: LCZ2:12(0.43 km^2), LCZ3:1(0.021 km^2), LCZ4:3(0.1 km^2), LCZ5:21(1.3 km^2), LCZ6:31(3.8 km^2), LCZ8:25(4.1 km^2), LCZ9:7(0.55 km^2), LCZA:11(1.8 km^2), LCZD:25(19 km^2), LCZE:9(0.5 km^2), LCZG:22(7.1 km^2),</td><td>N (size) of traing areas: LCZ1:3(0.31 km^2), LCZ2:10(9.3 km^2), LCZ3:9(1.1 km^2), LCZ4:8(0.97 km^2), LCZ5:7(1.1 km^2), LCZ6:18(6 km^2), LCZ8:10(5.5 km^2), LCZ9:2(0.95 km^2), LCZA:11(31 km^2), LCZB:6(2.7 km^2), LCZD:9(71 km^2), LCZE:10(1.5 km^2), LCZG:6(1.5 km^2)</td><td>N (size) of traing areas: LCZ2:3(1.1 km^2), LCZ3:5(0.4 km^2), LCZ4:2(0.087 km^2), LCZ5:8(1.2 km^2), LCZ6:22(5.3 km^2), LCZ8:16(4 km^2), LCZ9:10(3.9 km^2), LCZ10:1(0.095 km^2), LCZA:12(17 km^2), LCZB:10(1.1 km^2), LCZD:24(15 km^2), LCZE:11(1.2 km^2), LCZF:10(3.5 km^2), LCZG:13(0.91 km^2)</td></tr>
<tr><td>LCZ fractions</td><td>URBAN: 12 NATURAL[E]: 88 LCZ: < 2:0.064, 4:0.0023, 5:1.1, 6:7.2, 8:2.5, 9:0.82, A:1.9, D:84, E:0.021, G:2.9 ></td><td>URBAN: 18 NATURAL[\E]: 82 LCZ: < 1:0.02, 2:1.4, 3:1, 4:0.26, 5:0.74, 6:12, 8:2.1, 9:0.2, A:22, B:1.2, D:59, E:0.15, G:0.29 ></td><td>URBAN: 19 NATURAL[\E]: 81 LCZ: < 2:0.095, 3:0.017, 4:0.0052, 5:1, 6:9.4, 8:1.5, 9:6.8, 10:0.00097, A:4.3, B:1.6, D:73, E:0.084, F:1.7, G:0.47 ></td></tr>
<tr><td>Extend of scenes</td><td>Lat: 47.22, Lon: −1.564, UTM zone 30 N X: 590328–634,228 Y: 5212460–5,252,060</td><td>Lat: 48.88, Lon: 2.186, UTM zone 31 N X: 390330–506,330 Y: 5361570–5,460,370</td><td>Lat: 43.6, Lon: 1.377, UTM zone 31 N X: 348152.9–395,652.9 Y: 4795126–4,859,926</td></tr>
<tr><td>Cellzise</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td>Bootstrapping results</td><td></td><td></td><td></td></tr>
<tr><td>Mean OA</td><td>0.87</td><td>0.87</td><td>0.71</td></tr>
<tr><td>Mean kappa</td><td>0.81</td><td>0.79</td><td>0.64</td></tr>
<tr><td>Mean OA_urb</td><td>0.74</td><td>0.76</td><td>0.63</td></tr>
<tr><td>Mean OA_built-up</td><td>0.96</td><td>0.97</td><td>0.90</td></tr>
<tr><td>Mean WA</td><td>0.96</td><td>0.96</td><td>0.92</td></tr>
<tr><td>% Mode class urban</td><td>0.70</td><td>0.75</td><td>0.70</td></tr>
<tr><td>% Mode class all</td><td>0.89</td><td>0.83</td><td>0.77</td></tr>
<tr><td>Status</td><td>Passed</td><td>Passed</td><td>Passed</td></tr>
<tr><td>Review</td><td></td><td>Accepted</td><td>Accepted</td></tr>
</table>

Appendix B. Geographical comparison between MApUCE and WUDAPT classification for Paris and Nantes on: (A) the built area classification obtained through MApUCE and WUDAPT-L0 methods and (B) the Building Density and Building Height classes




Nantes
(A)
MApUCE
WUDAPT unfiltered
Legend
Compact high rise
Compact mid rise
Compact low rise
Open high rise
Open mid rise
Open low rise
Lightweight
Large low rise
Sparsely built
0 2.5 5 7.5 10 km
N
B)
MApUCE
WUDAPT unfiltered
Large low rise
0-10m
10-25m
25m+
Large low rise
Sparsely Built
Open
Compact
0 5 10 km
N

Appendix C. Example of sheets presenting ranges for LCZ 2 urban canopy model parameters according to the framework proposed by SO12 for Paris, Toulouse and Nantes



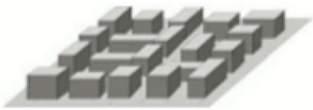
LCZ TOULOUSE : COMPACT MIDRISE (529 USR) 2

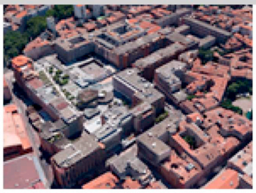
DEFINITION

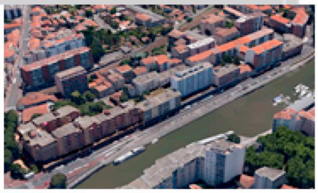
Form: Attached or closely spaced buildings 3–9 stories tall. Buildings separated by narrow streets and inner courtyards. Buildings uniform in height. Sky view from street level significantly reduced. Heavy building materials (stone, concrete, brick, tile) and thick roofs and walls. Land cover mostly paved; few or no trees. Moderate space heating/cooling demand. Moderate to heavy traffic flow. ***Function***: Residential (multi-unit housing; multistorey tenements); commercial (office buildings, hotels, retail shops); industrial (warehouses, factories). ***Location***: Core (old city, old town; inner city, central business district); periphery (high-density sprawl). ***Correspondence***: UCZ2 (Oke, 2004); A1, A2, A4, Dc2 (Ellefsen, 1990/91).

ILLUSTRATION

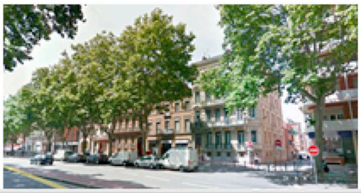
High angle

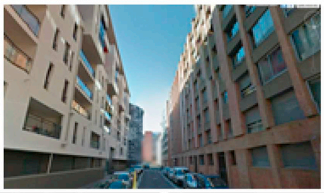




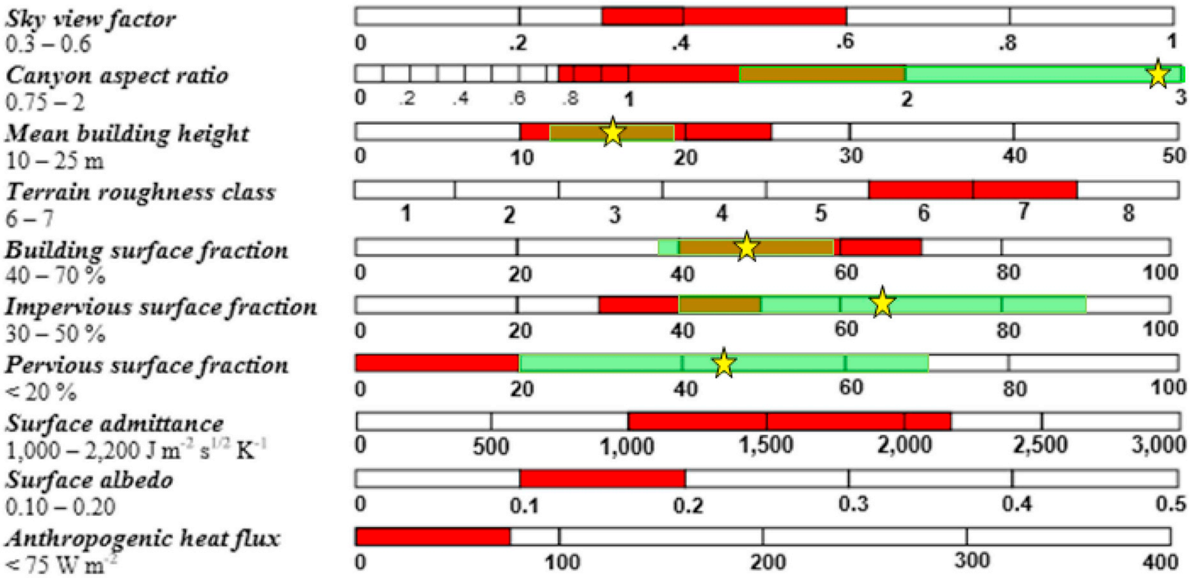


Low level





PROPERTIES



References

Acero, J.A., 2012. Urban Climate Modelling. Development of Climate Evaluation Methods for Urban Planning Purposes. Ph.D. dissertation thesis. University of Kassel. ISBN (978-84-88734-09-9).

Bechtel, B., Alexander, P.J., Beck, C., Böhner, J., Brousse, O., Ching, J., Demuzere, M., Fonte, C., Gál, T., Hidalgo, J., Hoffmann, P., Middel, A., Mills, G., Ren, C., See, L., Sismanidis, P., Verdonck, M.-L., Xu, G., Xu, Y., 2019. Generating WUDAPT Level 0 data – Current status of production and evaluation. Urban Clim. 27, 24–45.

https://doi.org/10.1016/j.uclim.2018.10.001.
Bechtel, B., Daneke, C., 2012. Classification of local climate zones based on multiple earth observation data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5 (4), 1191–1202. https://doi.org/10.1109/JSTARS.2012.2189873.
Bechtel, B., Demuzere, M., Sismanidis, P., Fenner, D., Brousse, O., Beck, C., Coillie, F., Conrad, O., Keramitsoglou, I., Middel, A., Mills, G., Niyogi, D., Otto, M., See, L., Verdonck, M.-L., 2017. Quality of Crowdsourced Data on Urban Morphology—The Human Influence Experiment (HUMINEX). Urban Sci. 1, 15. https://doi.org/10.3390/urbansci1020015.
Bechtel, B., Pesaresi, M., See, L., Mills, G., Ching, J., Alexander, P.J., Feddema, J.J., Florczyk, A.J., Stewart, I., 2016. Towards consistent mapping of urban structures – global human settlement layer and local climate zones. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI-B8 1371–1378. https://doi.org/10.5194/isprs-archives-XLI-B8-1371-2016.
Bocher, Erwan, Petit, Gwendall, Bernard, Jérémy, Palominos, Sylvain, 2018. A Geoprocessing Framework to Compute Urban Indicators: The MApUCE Tools Chain. Urban Climate. Vol. 24. Elsevier, pp. 153–174.
Brousse, O., Martilli, A., Foley, M., Mills, G., Bechtel, B., 2016. WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. Urban Climate 17, 116–134. https://doi.org/10.1016/j.uclim.2016.04.001.
Ching, J., Mills, G., Bechtel, B., See, L., Feddema, J., Wang, X., Ren, C., Brousse, O., Martilli, A., Neophytou, M., Mouzourides, P., Stewart, I., Hanna, A., Ng, E., Foley, M., Alexander, P., Aliaga, D., Niyogi, D., Shreevastava, A., Bhalachandran, P., Masson, V., Hidalgo, J., Fung, J., Andrade, M., Baklanov, A., Dai, W., Milcinski, G., Demuzere, M., Brunsell, N., Pesaresi, M., Miao, S., Mu, Q., Chen, F., Theeuwes, N., 2017. World Urban Database and Access Portal Tools (WUDAPT), an urban weather, climate and environmental modeling infrastructure for the anthropocene. Bull. Amer. Meteor. Soc. https://doi.org/10.1175/BAMS-D-16-0236.1. Accepted.
Crombette, P., Le Corre, S., Tinel, C., 2014. Traitement d'images satellitaires à Très Haute Résolution Spatiale et identification de zones à enjeux dans l'aménagement des Trames Vertes urbaines, Revue Française de Photogrammétrie et de Télédétection. http://www.sfpt.fr/rfpt/index.php/RFPT/article/view/130. https://hal-univ-tlse2.archives-ouvertes.fr/hal-01140009.
Geletic, J., Lehnert, M., 2016. GIS-based delineation of local climate zones: the case of medium-sized central European cities. Moravian Geographical Reports 24 (3), 2–12. https://doi.org/10.1515/mgr-2016-0012.
Hidalgo, J., Masson, V., Baklanov, A., Pigeon, G., Gimeno, L., 2009. Advances in urban climate modelling. ANYAS (Annals of the New York Academy of Sciences) issue "Trends and Directions in Climate Research" 1146, 354–374 (ISBN: 978-1-57331-732-0).
Houet, T., Pigeon, G., 2011. Mapping urban climate zones and quantifying climate behaviors–an application on Toulouse urban area (France). Environ. Pollut. 159 (8), 2180–2192.
Kaloustian, Noushig, Bechtel, Benjamin, 2016. Local climatic zoning and urban heat island in Beirut. Procedia Engineering 169, 216–223. https://doi.org/10.1016/j.proeng.2016.10.026.
Kotharkar, R., Bagade, A., 2017. Local climate zone classification for Indian cities: a case study of Nagpur. Urban Climate. https://doi.org/10.1016/j.uclim.2017.03.003.
Leconte, F., Bouyer, J., Claverie, R., Pétrissans, M., 2015. Using local climate zone scheme for UHI assessment: evaluation of the method using mobile measurements. Build. Environ. 83, 39–49.
Ng E, 2015. Adapting Asian Cities to climate & Urban Climatic changes. Open Plenary session, ICUC9, Toulouse.
Ng, E., Ren, C., 2015. The Urban Climatic Map: A Methodology for Sustainable Urban Planning. Routledge, Abingdon, Oxon; New York, NY.
Oke, T.R., 1982. The energetic basis of the urban heat island. Quart. J. Roy. Meteor. Soc. 108, 1–24.
Oke, T.R., 1988. The urban energy balance. Prog. Phys. Geogr. 12, 471–508.
Oke, T., Mills, G., Christen, A., Voogt, J., 2017. Urban climates. In: Urban Climates (p. I). Cambridge University Press, Cambridge.
Perera, N.G.R., Emmanuel, R., Mahanama, P.K.S., 2012. Mapping “Local Climate Zones” and Relative Warming Effects in Colombo, Sri.
Schmid, H.P., Cleugh, H.A., Grimmond, C.S.B., Oke, T.R., 1991. Spatial variability of energy fluxes in suburban terrain. Boundary Layer Meteorology 54 (3), 249–276.
Schoetter, R., Masson, V., Bourgeois, A., Pellegrino, M., Lévy, J.-P., 2017. Parametrisation of the variety of human behaviour related to building energy consumption in TEB (SURFEX v. 8.2). Geoscientific Model Development 10, 2801–2831.
Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93 (12), 1879–1900. https://doi.org/10.1175/bams-d-11-00019.1.
Tornay, N., Schoetter, R., Bonhomme, M., Faraut, S., Lemonsu, A., Masson, V., 2017. GENIUS: a methodology to define a detailed description of buildings for urban climate and building energy consumption simulations. Urban Climate 20, 75–93.
Unger, J., Lelovics, E., Gál, T., 2014. Local climate zone mapping using GIS methods in Szeged. Hung. Geogr. Bull. 63 (1), 29–41.
Zheng, Y., et al., 2017. GIS-based mapping of local climate Zone in the high-density city of Hong Kong. Urban Climate. https://doi.org/10.1016/j.uclim.2017.05.008.