City-descriptive input data for urban climate models: Model requirements, data sources and challenges

Valéry Masson, Wieke Heldens, Erwan Bocher, Marion Bonhomme, Bénédicte Bucher, Cornelia Burmeister, Cécile de Munck, Thomas Esch, Julia Hidalgo, Farah Kanani-Sühring, et al.

Version of Record: https://www.sciencedirect.com/science/article/pii/S2212095518303663
Manuscript_51c05c2eb7e7e06af450c262e226e8f0

City-descriptive input data for urban climate models: Model requirements, data sources and challenges

Valéry Masson[1], Wieke Heldens[2], Erwan Bocher[3], Marion Bonhomme[4], Bénédicte Bucher[5], Cornelia Burmeister[6], Cécile de Munck[1], Thomas Esch[2], Julia Hidalgo[7], Farah Kanani-Sühring[8], Yu-Ting Kwok[9], Aude Lemonsu[1], Jean-Pierre Lévy[10], Björn Maronga[8], Dirk Pavlik[6], Gwendall Petit[3], Linda See[11], Robert Schoetter[1], Nathalie Tornay[4], Athanasios Votsis[12], Julian Zeidler[2]

[1] National Center for Meteorological Research, Météo-France and CNRS, Toulouse, France
[2] DLR, German Remote Sensing Data Center Land Surface, Oberpfaffenhofen, Germany
[3] CNRS, Lab-STICC, Vannes, France
[4] Research Laboratory in Architecture, Toulouse, France
[5] Université Paris Est, LaSTIG, IGN, Saint-Mandé, France
[6] GEO-NET Umweltconsulting GmbH, Dresden, Germany
[7] National Center of Scientific Research (CNRS), LISST laboratory, Toulouse, France
[8] Leibniz University Hannover, Hannover, Germany
[9] Chinese University of Hong-Kong, Hong-Kong, China
[10] LATTS, Paris, France
[11] International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
[12] Finnish Meteorological Institute

Table of Contents

Abstract

Cities are particularly vulnerable to meteorological hazards because of the concentration of population, goods, capital stock and infrastructure. Urban climate services require multi-disciplinary and multi-sectorial approaches and new approaches in urban climate modelling. This paper classifies the required urban input data for both mesoscale state-of-the-art Urban Canopy Models (UCMs) and microscale Obstacle Resolving Models (ORM) into five categories and reviews the ways in which they can be obtained. The first two categories are (1) land cover, and (2) building morphology. These govern the main interactions between the city and the urban climate and the Urban Heat Island. Interdependence between morphological parameters and UCM geometric hypotheses are discussed. Building height, plan and wall area densities are recommended as the main input variables for UCMs, whereas ORMs require 3D building data. Recently, three other categories of urban data became relevant for finer urban studies and adaptation to climate change: (3) building design and architecture, (4) building use, anthropogenic heat and socio-economic data, and (5) urban vegetation data. Several methods for acquiring spatial information are reviewed, including remote sensing, geographic information system (GIS) processing from administrative cadasters, expert knowledge and crowdsourcing. Data availability, data harmonization, costs/efficiency trade-offs and future challenges are then discussed.

1) Introduction

1.1 Brief overview of urban atmospheric modelling

Cities are particularly vulnerable to meteorological hazards because of the concentration of population, goods, capital stock and infrastructure. In addition, they create their own urban micro-climate because they have a strong influence on the local meteorology (Oke et al., 2017). For example, heat waves, enhanced locally by the so-called 'Urban Heat Island' (UHI), can increase mortality rates in cities (Fouillet et al., 2006; Tan et al., 2010). Large cities can also interact with local

mesoscale flows, e.g., through intensification of the sea-breeze front due to thermal effects (Freitas et al., 2007), or, in contrast, they can slow down breezes due to enhanced friction (von Glasow et al., 2013). Urban climate also impacts the energy use of domestic heating and air conditioning (Ohashi et al., 2007; Santamouris et al., 2015; Kohler et al., 2016), which, in turn, enhances the UHI due to the associated heat release (de Munck et al., 2013; Wang et al., 2018). Thunderstorm activity can either be concentrated in different parts of the city or be enhanced above or downstream of the city (Shepherd, 2005). In such a complex multi-dimensional and multi-objective decision environment, pertinent and clear, decision-relevant information is indispensable for stakeholders, e.g., for fine-scale weather forecasting in cities during sporting events (Joe et al., 2017) or for urban planning (Hidalgo et al., 2018).

Numerical modelling has become a useful tool for analyzing detailed urban meteorology and to design urban climate services, e.g., to evaluate risks or benefits of urban planning or adaptation strategies. To simulate urban climate, atmospheric models need to have an adequate representation of the influence of the city on the exchanges with the atmosphere above. This is done using Urban Canopy Models (UCMs). These UCMs (such as Masson, 2000; Kusaka et al., 2001; Martilli et al., 2002; Lee and Park, 2008; Lemonsu et al., 2012; Schubert et al., 2012; Wouters et al., 2016) are surface schemes that aim to represent the energy, water and momentum exchanges between the urban surface and the atmosphere. They are often based on a simplified description of the 3D shape of the city, e.g., for mesoscale modelling, the 'urban canyon' approximation is often used. Masson (2006) describes several types of UCM according to their degree of complexity and realism, and Grimmond et al. (2010, 2011) further classify each physical component of these schemes, underlining the need for data to describe the city. When going to the microscale, individual buildings are explicitly resolved by Obstacle Resolving Models (ORMs) (Krayenhoff et al., 2015; Resler et al., 2017; Salim et al., 2018). These ORMs create a high demand for detailed input data.

1.2 Scale issues: mesoscale and microscale

Mesoscale UCMs require suitable surface input data to provide reliable boundary conditions for atmospheric models. The relevant and necessary spatial scale of these input data, however, strongly depends on a) the atmospheric scale to be investigated, and b) the scale and level of detail of the UCM applied. For mesoscale applications, this ranges from estimations of the roughness length as a first order approximation of the friction due to urban canopies (Grimmond and Oke, 1999), to the more advanced street canyon approach (e.g. Masson, 2006; Schubert et al., 2012), which requires the determination of typical canyon parameters at the grid cell size of the atmospheric model. Note that canyon here encompasses both impervious (e.g., roads, sidewalks, car parks, etc.) and pervious (e.g., vegetated) areas. For mesoscale models, this might be at the kilometer down to the hectometer scale. However, for microscale models that operate at the meter scale, individual elements in the urban canopy layer such as buildings, streets and trees are resolved explicitly. Today, urban microscale models such as PALM/PALM-4U (Maronga et al., 2015, 2019a,b) are able to simulate city quarters at a grid spacing of 1 m and entire cities at 10 m. It is obvious that such fine grid spacing requires the input data to have a different level of detail. Moreover, many typical parameters that would normally be derived must be replaced by the actual conditions of individual urban structures, such as specific buildings or tree configurations in streets and parks. The building envelope is then no longer represented by a simple street canyon but by individual surface elements that - as in reality - have different insulation, window fraction, surface albedos, and which cast shadows on each other. The high relevance of this information has been highlighted by Resler et al. (2017) in a systematic sensitivity study of material parameters, where the sensitivity of the modelled surface temperature reached up to 5°C with a variation in the surface albedo by ± 0.2. This paper

illustrates that different data acquisition strategies and data sources are required to fulfill the individual requirements of numerical models.

1.3 Coverage issues: from city-scale to global modelling

Urban climate studies have been traditionally focused on the analysis of energy exchange and turbulent processes using field campaigns or they have investigated the spatial distribution of the temperature field from meteorological stations (Arnfield, 2003). Almost all of these studies have focused on a specific site or on a given city where the physical description of the city was done locally. This tendency has continued with the arrival of UCMs in the early 2000s, and weather forecasting models that cover countrywide areas with UCMs (Seity et al., 2011) are now operational. There is now more common use of atmospheric models for urban climate studies, and even recently, the development of regional and global climate models (Oleson et al., 2011) with UCMs. This underlines the necessity to have homogeneous descriptions and methods for producing urban parameters at a fine resolution - typically one or a few urban blocks. This should be done in comparable ways for a set of representative cities for scientific studies, and ideally worldwide.

The objective of this paper is to categorize the urban parameters needed for UCMs and then to discuss their availability and acquisition. We start with a high-level categorization of 5 major types as follows:

- Land use and land cover;
- Morphological parameters;
- Architectural parameters;
- Socio-economic parameters and building use; and
- Urban vegetation.

For each type, several possible ways to obtain spatially explicit parameters are presented and discussed. The final section considers the availability and access to the data, the quality of such data sets and the associated trade-offs between them.

1.4 Fit for purpose

The requirements for these urban parameters will depend on the purpose of each study. Before detailed information is provided on how to obtain these parameters, it is important to stress that for some applications, it may not be necessary to invest, either time or money, into the construction of detailed maps of all parameters. The evaluation of the sensitivity of the model results on the accuracy of the input urban parameters is out of the scope of this paper. However, some guidance may still be given. Table 1 presents the type of parameters that are most crucial for several types of applications. From a broad point of view, the higher the spatial resolution, the more detailed parameters are required. However, the specific scope of a study, which could be, for example, the modelling of pollutant emissions from domestic heating or the quantification of human thermal comfort, will have a large influence on which input parameters are the most relevant.

Purpose \ Parameter	*Land use/cover classes* - meso-scale: at neighborhood scale (e.g. LCZ) - micro-scale: urban elemental objects (e.g. buildings, roads)	*Morphology* incl. height, building & (im)pervious fractions…	*Architecture*	*Socio-economy & uses*	*Vegetation description* (Type, LAI,…)
At meso-scale					
1km-res NWP & climate models	x (at neighborhood scale)				
100m res NWP	x	x			
Forcing of AQ models	x				
Forcing of & interactive emissions of buildings for AQ models	x	x	x	x	
At micro-scale					
1m-res Building resolving modelling	x (urban objects)	x			Leaf Area Density
Radiative effects, shadows	x	x	Albedo (incl. Windows)		Leaf Area Density
Flow modification	x	x			Leaf Area Density
Development of parameterizations for urban climate processes	x	x	Possibly (e.g. energy balance)	Possibly (e.g. traffic induced turbulence)	Possibly (e.g. pollutants dispersion)
At both micro and meso-scales					
Outdoor heat-stress quantification	x	x	Albedo (incl. Windows)		x
Indoor heat-stress and Energy consumption quantification	x	x	x	x	
CO2 fluxes in urban areas	x	x	x	x	x
Urban hydrology modelling	x	Coverage fractions			x

Table 1 : Parameters needed in order of priority depending on the purpose.

In current meso and microscale model applications, it is possible, and even common, to only use a map of land use/land cover to completely initialize all the other parameters. This is done through look-up tables that assign uniform values to each morphological (e.g., building height, building density) or other parameter depending on the land cover type (e.g., higher and denser buildings in “dense urban” compared to “suburban”). This is defined as the ‘indirect method’ in Figure 1. Numerical Weather Prediction (NWP) and its direct application to provide meteorological forcing to Air Quality (AQ) models applied at a kilometric scale typically relies on such look-up tables. This assumes that at a spatial scale of 1 km x 1km, details on the urban structure are of minor relevance for weather forecasting and air quality modelling. However, to fully benefit from an increased horizontal resolution (e.g. 100 m) of NWP and AQ models, and to be able to simulate fine-scale impacts (see Table 1), it is recommended that a fine scale description of the urban parameters is provided, with a priority to be given to the morphological parameters. This requires employing ‘direct methods’ using ancillary data (e.g., remote sensing or building cadasters) to compute these parameters. A summary of these direct methods is provided in Figure 1, and they are (briefly) presented throughout this article for each of the 5 types of parameters.

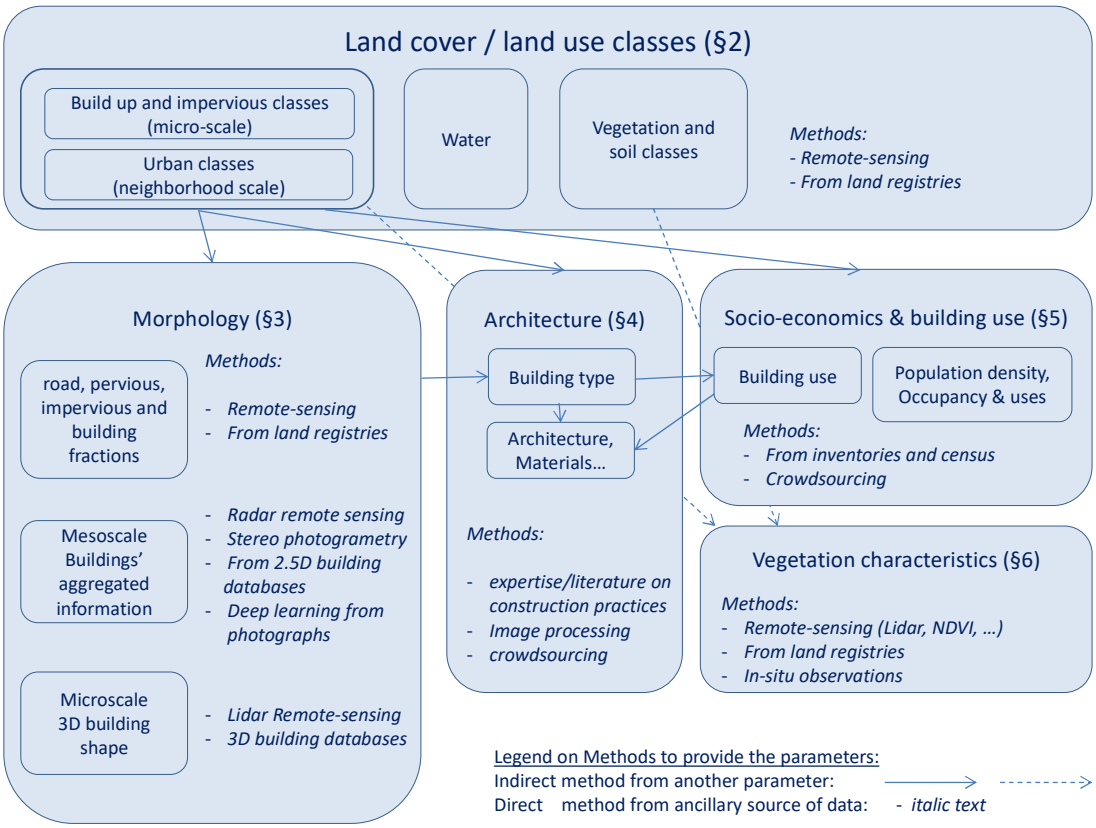

Figure 1: Overview of the methods to provide the parameters

2) Land use and land cover classes

2.1 Description of the parameters and their relevance

The physical and biophysical cover of the Earth strongly modifies the momentum, energy and water fluxes of the atmosphere above. The sensible heat flux is generally enhanced above impervious surfaces, while evapotranspiration by soil and plants will favor latent heat fluxes. Drastic spatial heterogeneity in land use and hence on land cover, like that which happens in coastal cities in sea breeze conditions, can significantly influence the entire boundary layer above a city when compared to a continental city. Therefore, the mapping of urban, vegetation and water cover is essential and can be further refined, depending on the model requirements. The class 'water' can, for example, be separated into sea, rivers, lakes, and ponds. Vegetation is often separated into high and low vegetation (see section 6 for further discussion).

Currently, there is no global data set on the urban tissue that can directly provide the most relevant parameters for UCMs (e.g., plan area, building density, building height, etc.). Instead, such parameters are derived from available land cover products using the relevant classes and a set of heuristics. There are two common ways to estimate the required parameters. The first option is to provide a map of each required parameter to the model. This approach is described in sections 3 to

6. However, it is generally impossible to have such a fine description for all of the UCM parameters for any given city of investigation. Therefore, the second option must be used for parameters that are unavailable, which is to employ land cover maps. Existing land cover products, e.g., produced automatically from remote sensing imagery, are of crucial interest since they usually identify urban areas. However, they can also be used to estimate all the other parameters related to morphology, architecture, the social and economic environment and vegetation, by using a lookup table, particularly when other sources of information are lacking (or not used).

At meso-scale, the absolute minimum information is to have at least one class called “urban” within the land cover map, and to assign uniform values to this class for all urban model parameters across the globe. Most databases used in atmospheric models, even recent ones, have only one urban class. Depending on the model requirements and the available data, the urban class can be refined in several ways. For example, it can be separated into different urban densities, which commonly refers to the density of the buildings although population density would be an alternative method. These classes typically use either the land use or land cover (and often both) to describe the city at the neighborhood scale.

Examples of global land cover maps with one single urban class are the satellite-based GlobCover (300 m, one class ‘artificial surfaces and associated areas’; Arino et al., 2008), MODIS Land Cover (500 m, one class ‘urban and built-up land’; Friedl et al., 2010), the ESA-CCI (300 m; one class ‘urban areas’; Bontemps et al., 2013) and GlobeLand30 (30 m, one class ‘artificial surfaces’; Chen et al., 2015). Several regional and global initiatives have been undertaken to gather a more detailed description of the urban tissue. An example is the Global Human Settlement Layer (GHSL) LABEL product (38 m resolution; Pesaresi et al., 2013, 2016). This data set distinguishes roads, built-up areas with different densities (very light/light/medium/strong) and for the strongly built-up areas, the building height (low rise/medium rise/high rise/very high rise). Congalton et al. (2014) and Grekousis et al. (2015) review a large number of global and regional land cover products. A (non-exhaustive) overview of global and regional land cover data sets that are suitable for urban climate modelling is provided in Appendix 1.

The deployment of the Local Climate Zone (LCZ) typology (Stewart and Oke, 2012, Figure 2) standardizes the way that UCM parameters are described. This typology aims to characterize the urban tissue and induce urban climate heterogeneity at the neighborhood scale (typically ≥ 1 km^2). LCZs represent urban areas that are relatively homogeneous in their type of urbanization. They represent either the rural landscape (in seven classes: dense or scattered trees, bushes, low plants, bare soil, bare rock, water) or the urban landscape, more specifically, in ten classes (compact or open high-rise buildings, compact or open mid-rise buildings, compact, open or sparse low-rise buildings, large low-rise buildings, heavy industry, lightweight low-rise buildings). LCZs are provided with ranges of values for a few UCM parameters, which are the sky view factor, the aspect ratio, the mean building height, the terrain roughness class, the building, pervious and impervious surface fractions, the overall thermal admittance and albedo, and the anthropogenic heat flux. This approach aims to be ‘universal’, i.e., with limited cultural bias, which means that it can be applied to cities worldwide. LCZs are intended to represent areas that have, independently of other geographic or topographic factors, a homogeneous local thermal climate. A few experimental campaigns, such as Houet and Pigeon (2011) using fieldwork, air temperature measurements and surface temperature satellite images, or Leconte et al. (2017) with an instrumented car, have confirmed that this assumption is correct. For ORMs, however, the above data sets and approaches are not sufficient as the urban canopy must be described explicitly.

At micro-scale, much finer spatial details are necessary for describing the urban surface. Contrary to the land use and land cover classes that represent the city at the neighborhood scale, here it is necessary to describe the elemental objects of the city individually. Another approach is to describe the urban land cover by its components, thus mapping urban objects such as buildings, paved surfaces (roads, etc.) and meadows, which is a typical approach for ORMs.

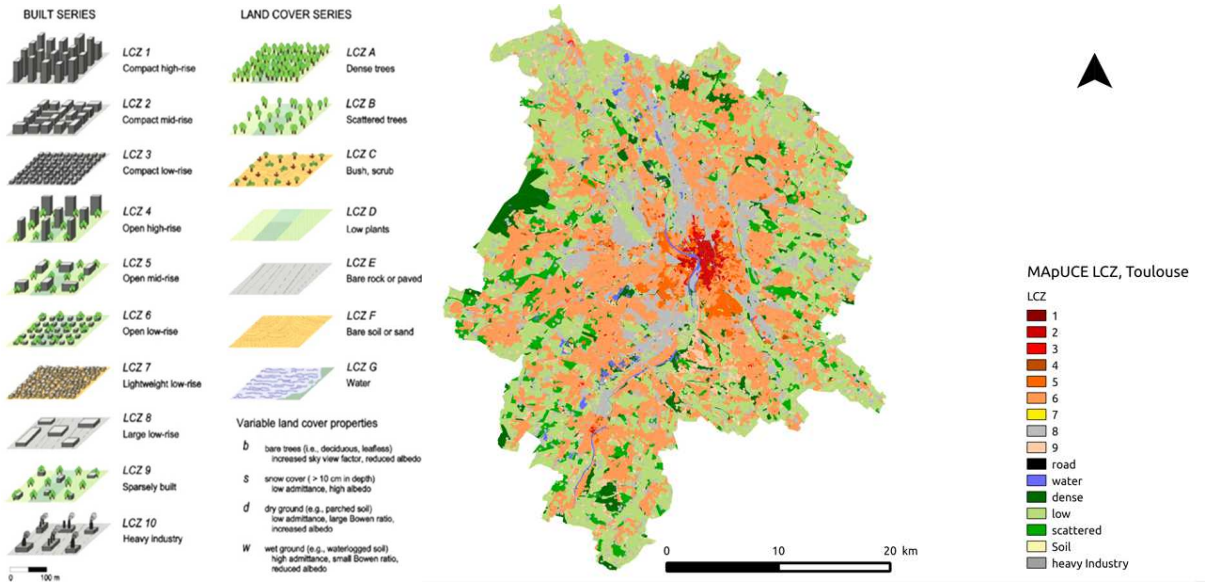

Figure 2: *LCZ classification scheme. Left: LCZ description (source: Stewart and Oke 2012); right: vector based LCZ map of Toulouse, France at urban block scale produced using a building registry.*

2.2 Methodologies to gather land cover data

2.2.1. Remote sensing methods

Remote sensing is a convenient tool for land cover mapping. It can cover large areas at once, allow comparable mapping across different cities, and the data sets can be updated quite easily. However, several aspects need to be considered when selecting a suitable sensor. The spatial resolution and the classification scheme need to be defined. When more classes need to be considered, then the remote sensing data will need more spectral and spatial detail. A suitable classification scheme for remote sensing applications for urban climate is the Vegetation, Impervious and Soil (V-I-S) approach by Ridd (1995), which also addresses the mixtures of land cover within one pixel. Often the impervious surface is separated into a high and a low albedo class (Lu and Weng, 2006), especially at higher spatial resolutions (Yang and He, 2017). The V-I-S approach has, for example, been applied to UHI mapping and surface temperature studies (e.g. Zhang et al., 2015; Wang et al., 2016).

With higher spatial and spectral resolution, more classes can be distinguished, such as buildings and pavements (e.g. Thomas et al., 2003) up to material mapping with high resolution hyperspectral remote sensing data (Heldens et al., 2017). Once the classification scheme has been determined, a suitable sensor can be selected. For spatial resolutions coarser than 10 m, many satellite systems are available. Depending on the sensor, the data can often be obtained for free. Spatial resolutions of less than 10 m require high resolution imagery, which is often not freely available. In addition to satellite data, airborne sensors can also be used. Based on the classification scheme and the remote

sensing data, the classification method is then selected. Generally, supervised and unsupervised methods are available. An overview of land cover mapping techniques is addressed by Lu and Weng (2007) or Ban et al. (2015).

For high resolution data, object-based classification techniques can be used (Ma et al., 2017). Considerable research efforts are being directed towards developing and improving machine learning techniques, and they are especially promising for complex classification tasks (Maxwell et al., 2018).

A land cover classification approach that has been developed specifically for urban climate applications is the World Urban Database and Access Portal Tools (WUDAPT) project. This is an initiative started by the community of urban climatologists (Bechtel et al., 2015; Ching et al., 2018), which proposes a common methodology for producing LCZ raster maps using supervised classification of satellite images. Some large areas have already been covered at a 100 m resolution, such as China and Europe.

2.2.2. From vector topographical databases and land registries

In addition to remotely sensed imagery, vector-formatted information about the urban built environment and human activities in cities is a useful resource for mapping land cover (and to derive parameters for UCMs). Typical sources of such vector data are local, national or European public agencies (e.g., planning agencies, mapping or statistical institutes, land and property registries, Eurogeographics, which is a consortium of national mapping agencies producing European products), private companies (e.g., real estate brokers, environmental services companies, transportation consultants), international organizations (e.g., the European Environment Agency, the European Commission's EUROSTAT agency and the thematic data banks of the World Bank, International Monetary Fund, or the United Nations), and crowdsourced initiatives such as OpenStreetMap[1] (OSM). OSM is an initiative in which volunteers contribute to a global map of the world (Mooney and Minghini, 2017).

One method to extract land use information from vector data sets is to combine cadaster plans of land plots and buildings with the information that is registered about these geographical entities in, e.g., taxation, planning permits, or ownership records. GIS (or AutoCad) software can be used to develop combined data sets that represent urban land patches and built structures in very high spatial resolution together with detailed information about their uses, which can then be transformed into land use maps. Another method is to use the vector polygons of building footprints and other man-made structures contained in the aforementioned data sets to "tag" those pixels in existing land cover maps as "urban" when these footprints and structures are encountered (see e.g. Jantz et al., 2004). This provides the possibility of creating more detailed urban classes since building footprints often contain information on their type and use. Similar to the above, a third method is to access spatial data services, e.g. the OSM query service or national portals to request semantic information about particular coordinate points or polygons. The semantic information can include the type of human activity reported for a specific structure, e.g., recreation, shopping, dining, etc., or the general category to which a structure is classified, e.g., residential, commercial, government, etc. This approach can be used to develop land use maps or enhance existing land cover maps (Olbricht, 2015; Boeing, 2017). Such methods using administrative cadasters have been used to map LCZs by Kotharkar and Bagade (2018) for Nagpur in India, Geletič and Lehnert (2016) for three cities in the Czech Republic, by Zheng et al. (2018) in Hong Kong and for French cities by Hidalgo et al. (2019)

[1] http://www.openstreetmap.org/

(Figure 2). UCMs with a grid resolution finer than 25 m require more detailed land cover/land use data. In Germany there are area-wide cadasters available as ALKIS (Official Real Estate Cadastre) and ATKIS (Official Topographic Information System), which provide land cover and land use data at different scales. These systems are updated and validated every year. Comparable products exist in every European country. The INSPIRE directive aims to reuse such national data for monitoring European environmental activities and to design European policies in order to have more consistency between the local and European levels. To achieve this goal, a federated workflow for such content has been designed, and every country must publish data produced by legally mandated organizations through web services according to interoperable specifications at the European level. OSM might also be used as a data source at this high resolution although the quality differs largely across different areas. However, the quality is sufficient as an input to ORMs and UCMs in densely populated areas, especially with regard to the building footprints.

2.2.3. Data fusion

The data required on land cover or land use can often not be represented with sufficient accuracy by one of the aforementioned data sets alone. A good alternative is to combine multiple data sets to improve the resulting information layer. Such a data fusion approach can, for example, rely on signal processing techniques to merge different spectral signatures to yield materials information, on photogrammetric techniques to merge data with different angles to obtain 3D information, or on advanced GIS operations to combine different spatial formats (i.e., raster, vector and tables). For example, GIS information can be combined with remotely sensed information as proxies for parameters. This approach can be used to derive highly detailed land use and building use information by associating temporal, spatial and spectral information patterns to particular human activities, often with the help of deep learning algorithms (Ebert et al., 2009; Ghaffarian et al., 2018). It is very important, however, to have good information about the quality of each data set, as they might contain the same information with different qualities or spatial resolutions.

3) Morphological parameters

3.1 Description of the parameters and their relevance

Morphological parameters allow the 3D aspect of the city to be described, and they influence the momentum as well as radiative exchange, with shadowing and multiple reflections between buildings, leading to infrared radiation trapping (especially during nighttime). Furthermore, the 3D character of buildings increases the volume covered by solid materials and the amount of surface area that is in direct contact with the atmosphere. The 3D character of cities, and the thermal properties of construction materials, favor the storage of heat inside the urban canopy. Both aspects are the main physical reasons leading to the development of the nocturnal UHI. At the microscale, these processes are simulated by the explicit interactions between individual buildings. In order to represent them in ORMs, detailed information on the 2D and 3D urban structure is needed at a decameter to meter scale. The relevant morphological parameters here are the 3D building configuration, i.e., the information regarding which grid volumes are covered by built-up structures, including thoroughfares and bridges, and 2D maps of the surface configuration regarding impervious (roads, footways, parking lots, etc.) and pervious surfaces (water, vegetation, soil). At the mesoscale, these parameters are statistically aggregated into indicator values. Many indicators exist, some of which may be retrieved by combination with others. For mesoscale UCMs, typical morphological parameters that should be produced are described in Table 2.

Parameters	Comments
building fraction λ_p	Surface of buildings S_{bld}, seen from above divided by the horizontal surface of the urban area under consideration, S_{hor}
pervious and road fraction	Idem, for pervious and road/impervious surfaces
wall density λ_w	Ratio between the surface of walls in contact with the atmosphere, S_w, and the horizontal urban surface, $\lambda_w=S_w/S_{hor}$
mean building height, h	This is, along with λ_p, the key parameter needed by all UCMs.
σ_h	Standard deviation of the building height
roughness length z_0	While this is fundamentally an aerodynamic parameter describing the air flow above a surface, it is classically estimated from the structure of the surface below. There are many ways to estimate the roughness length mathematically in cities (Grimmond and Oke, 1999). Strictly speaking, the roughness length is only applicable for uniform homogeneous surfaces, but due to lack of alternatives, it is often also applied to heterogeneous surfaces and buildings.
h/w and Ψ	Canyon aspect ratio (building height divided by the canyon width, h/w) and sky view factor Ψ, which describes the compactness of the city.
Frontal area density, λ_f	λ_f is the surface of walls facing the wind normalized by the horizontal surface. $\lambda_f < \lambda_w$. It depends on the wind direction.

Table 2: Typical morphological parameters used by UCMs

In multi-layer models with intersecting atmospheric layers (Martilli et al., 2002; Hamdi and Masson, 2008), the friction is simulated using a drag coefficient approach. For this purpose, λ_f which describes the vertical arrangement between buildings, is often coupled with λ_p, because together they represent the nature of the breathability between the canopy and the atmosphere.

The canyon aspect ratio (h/w) and sky view factor Ψ describe the ability of the city to trap the radiation. A large h/w (small Ψ) ratio favors the storage heat flux while a low sky view factor reduces longwave radiation loss at night, but also reduces shortwave gain during the day.

3.2 Links between morphological parameters

In reality, the variety in the urban tissue and the arrangement between buildings of diverse shapes can lead to various combinations of parameters for a given area (either a grid mesh of the model, or an LCZ or urban block). In other words, all of these parameters can be mapped independently.

However, when using morphological parameters in mesoscale UCMs, they assume a simplified geometry. This implies: (1) that the UCMs cannot simulate all of the internal variability of the urban tissue, and (2) that these morphological parameters are no longer independent. This means that one cannot specify many of them independently using fine scale data without violating the physics and energy conservation within the UCM. This is why model developers have to identify: (1) which parameters are absolutely necessary; (2) which can be deduced from others (but could be estimated independently if the data are available); and, (3) which must be, in order to conserve the consistency of the UCM, deduced from others.

Parameters entering in momentum flux parameterizations are generally in the second category. Building heights, or λ_f and λ_p, can be used to define the roughness length (Grimmond and Oke, 1999) and drag coefficients (Santiago and Martilli, 2010). All parameters can also be used freely for diagnostic purposes, e.g. fine-scale maps of Ψ can be used for human comfort evaluation at several places within the canyon (or more generally the city), even if they are not used directly as an input to the UCM for the calculation of the surface energy balance. However, when energy balances of individual surfaces (such as roofs, walls, roads, vegetation) are considered, some parameters can no longer be independently specified from the others. Most UCMs are based on the canyon hypothesis, which assumes that there is an infinitely long street canyon, bordered by two walls of identical height. This produces some relationships between morphological parameters that cannot be considered independently any more. Energy conservation is governed by the amount of each surface in contact with the atmosphere, and radiative exchanges by how the surfaces 'see' each other. In the infinite canyon geometry, this implies, for the normalized wall surface λ_w, that the canyon aspect ratio h/w, the plan area density (or building fraction) λ_p, and the sky view factor in the middle of the street Ψ (from Noilhan, 1981) can be expressed as follows:

$$\lambda_w = 2\,(1-\lambda_p)\,h/w \tag{1}$$

$$\Psi = [\,(h/w)^2 + 1\,]^{1/2} - h/w \tag{2}$$

This means that when maps of λ_p, λ_w, Ψ and h/w are available, only two of them should be used for the calculation of the UCM energy balance. Therefore, the two recomputed parameters would not be coherent with the actual city data. Hogan (2018) has explored another geometric hypothesis based on an exponential distribution of road widths, and has shown that the induced relationships between morphological parameters can also be computed using only λ_p, λ_w and the building height.

Given (1) the difficulty in defining what a canyon width (w) is for a heterogeneous real urban tissue, or where to define the middle of the road required to calculate Ψ, and (2) the importance of the quantity of the surfaces in contact with the atmosphere for energy exchange, we strongly recommend using the following as primary parameters: building height, the building fraction, λ_p, which is classical, and the normalized wall surface, λ_w, even though the latter parameter is not generally considered as an independent input parameter in UCMs.

3.3 Methodologies to gather morphological parameters

3.3.1 Data from remote sensing

Due to the capability of cost-efficient large-area data collection, remote sensing has become a key source for describing the 3D structure of the built environment. The most commonly used baseline product for urban structural analyses derived with remote sensing technologies is the digital surface model (DSM). A DSM provides a detailed picture of the terrain surface along with all other vertical features such as buildings and infrastructure elements or trees and hedgerows. To more effectively describe and analyze vertical objects not related to the terrain topography (e.g., buildings), a normalized DSM (nDSM) is usually calculated by subtracting the modelled terrain height (digital terrain model) from the DSM (Weidner and Förstner, 1995; Reinartz et al., 2017).

The traditional technique for DSM/nDSM generation and subsequent analysis of the urban morphology is photogrammetric processing of optical stereo imagery from sensors mounted on airplanes (Fradkin et al., 1999; Hirschmüller et al., 2005) or from high and very-high resolution satellite data (Toutin, 2006; Eckert and Hollands, 2010; Sirmacek et al., 2012; Aguilar et al., 2014). Recently, unmanned aerial vehicles (UAVs) have increasingly been deployed to measure the 3D urban morphology. Gevaert et al. (2017) describe how point-cloud and image-based morphologic features derived from UAV data can be used to classify specific built-up structures related to informal settlements. Esch et al. (2018a) describe new perspectives for surveying local urban morphology arising from the utilization of drones, UAVs or High-Altitude Pseudo Satellites (HAPS).

An established alternative to stereo photogrammetry is the analysis of point clouds from airborne light detection and ranging (LiDAR). LiDAR systems send laser pulses towards the ground whereas the runtime length can be used to generate precise DSMs (Vosselman and Maas, 2010). Usually, LiDARs collect a cloud of 4-50 points per m^2 and thereby allow for a detailed characterization of the small-scale urban morphology (Jutzi and Stilla, 2003; Goodwin et al., 2009; Yan et al., 2015). Bonczak and Kontokosta (2019) demonstrate the application of point-based voxelization techniques to extract design parameters in complex urban environments in unprecedented spatial detail.

Moreover, synthetic aperture radar (SAR) technology can be employed to generate DSMs for urban areas by means of spaceborne SAR interferometry (InSAR) or SAR tomography (Fornaro and Serafino, 2006; Frey et al., 2014; Marconcini et al., 2014; Geiß et al., 2015; Zhu et al., 2016). However, regarding InSAR, the spatial resolution of the resulting DSMs is usually lower than the products derived from optical satellite systems, whereas SAR tomography is still a rather experimental method due to the massive satellite data requirements.

The Urban Atlas land cover/land use product from Copernicus has recently added building height data for the capital cities of EU countries. Although covering only a large (700) but still limited number of cities, this can provide data for testing/validation. This information is also available in the INSPIRE data and in the European Location Service.

3.3.2 GIS treatment of 2.5D cadaster vector data of individual buildings

GIS software is often used to compute morphological indicators such as the density of buildings, the mean building height, the compactness ratio, the sky view factor, etc. These indicators are applied to study and monitor the urban structure at different geographic scales (e.g., urban parcels or districts). Techniques are based on spatial analysis chains that use geometric data models (also called vector data). Urban features are represented by a set of polygons (e.g., buildings), lines (e.g., roads) or

points (e.g., elevation) grouped in GIS layers restricted to a 2.5D dimension. A vector data model offers several advantages for urban climate studies:

- It is able to store many parameters in one layer. For example, if a building is represented by a set of polygons, additional attributes such as height, age, floor area and wall material can be specified to describe it.
- It allows an accurate representation of urban geometry allowing urban climate models to operate on a very fine scale (e.g. 1 m). Berghauser Pont and Haupt (2005) define a building block as an aggregation of buildings that are in contact. The building block concept allows architectural patterns in the urban fabric (mid-rise compact building, closed blocks, etc.) to be retrieved. This scale is particularly useful for studies dealing with the interaction of urban climate and building energy demand (Bouyer et al., 2011).

Most GIS approaches use a regular vector grid to represent the urban space and compute the morphological indicators inside the cells (Lindberg, 2007). Ching et al. (2009) compiled data on buildings and urban vegetation (mainly via airborne LiDAR detection), anthropogenic heat release due to buildings, traffic and human metabolism (via a top-down approach combining energy consumption data and the regional climatic conditions), as well as day- and night-time population density (via census data) for approximately 130 US cities.

Although a regular grid may be suitable for populating a climate model, it also represents an abstract feature that does not conform to real parcels or urban forms, which tend to have irregular shapes and their own distinct spatial boundaries that stem from socioeconomic processes. Cities are often characterized by complex spatial assemblages that are smoothed out when using regular grids; the challenge is to utilize a representation of space that fuses both options and is therefore suitable for the representation and modelling of both physical and socioeconomic processes.

Dealing with a detailed mapping of urban areas that describe their morphology and spatial relationships is an old but still very relevant issue. Barnsley and Barr (1997) proposed a GIS graph model to perform a spatial classification of a geographically referenced digital urban map. Steiniger (2006) extended their set of morphologic properties to classify urban structures for mapping purposes. Bocher et al. (2018) describe a GIS processing chain to compute a set of morphological indicators based on 2.5D vector data available for the French territory. 64 indicators were calculated at three scales: individual buildings, blocks of buildings and a specific unit area called a Reference Spatial Unit (Figure 3). These indicators are combined with socioeconomic data to categorize the urban fabric.

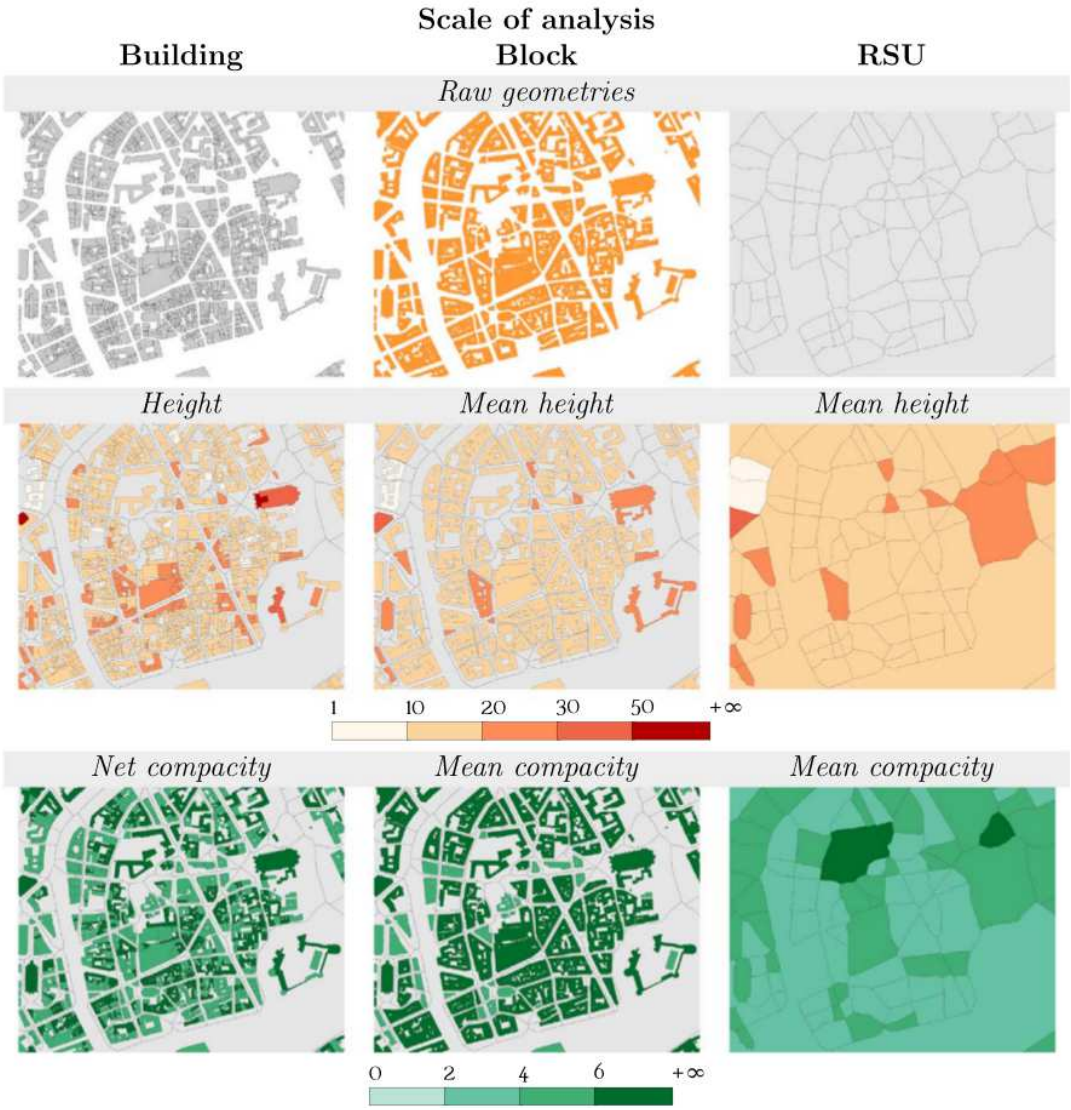

Figure 3. Maps for the three geographical scales (individual buildings, blocks of buildings and Reference Spatial Units) used to compute the morphological indicators. Source: Bocher et al (2018).

Samsonov et al. (2015) presented an advanced GIS technique to extract urban canyons from vector databases. The authors used a constrained Delaunay triangulation to delineate a hierarchy of the canyons based on the geometric properties of the triangles and their spatial relation with urban features, buildings and the street network. The morphological properties of the canyons are then combined with derived indicators (such as sky view factor and frontal area index) to feed the URB_MOS meteorological model. An application is discussed for the city of Moscow to model the spatial distribution of temperature and wind.

3.3.3 3D building data and CityGML for microscale modelling

The CityGML standard[2], developed by the *Open Geospatial Consortium* (*OGC*), defines three-dimensional geometry, topology, appearance and semantics of all relevant topographic urban objects at multiple well-defined level of details (LOD), ranging from simplified bounding boxes (LOD1: constant building height, LOD2: with roof shape) to explicit details like doors (LOD3) or indoor design (LOD4) (Gröger and Plümer, 2012). Because they represent the explicit form of

[2] https://www.opengeospatial.org/standards/citygml

buildings, they benefit ORMs more than UCMs in being able to use this finer scale data (e.g., LOD2, cf Figure 4). CityGML is very generic and does not only include all 3D objects of interest to ORMs, like buildings, trees, bridges, tunnels or even park benches, but it contains also basic classes such as relief, water bodies, roads or land use. Most available CityGML models have been automatically generated from a fusion of LiDAR and catastral data. Important landmarks are often manually added at a higher level of detail. Creation and improvement of CityGML data sets is an ongoing project for many municipalities. Germany, for example, plans to offer a comprehensive LOD2 building model by 2019 and to amend it with bridges and tunnels by 2020. The detailed CityGML data can then be rasterized and aggregated to the resolution and requirements of the ORM model. For ORMs at the meter resolution, information about textures can also be mined using machine learning approaches, e.g., to extract information like building materials and window or vegetation fractions for single facades.

As an example of the processing of CityGML data to represent buildings in a microscale simulation of Berlin at a grid spacing of 1 m, Figure 4 shows the original LOD2 CityGML while Figure 5 is the voxeled building configuration of the Reichstag building in Berlin, Germany.

German municipalities and/or federal states usually order laser scan data collections to generate their high-resolution 3D city models. Such data sets are updated every 5 years. After data collection, georeferencing and further processing lead to a 3-dimensional point cloud, which is the product of the first processing step. The subsequent categorization into different object types like buildings or vegetation is either already done by the data supplier or must be done by the end user. If such a categorization is missing, a common method to obtain the different object types is the intersection of the height information with land cover data, e.g., building footprints or by a classified Normalized Difference Vegetation Index (NDVI) layer, which divides the whole domain into pervious and non-pervious areas to categorize the point cloud. Further classification steps may be necessary to obtain the required input data structures for the respective ORM.

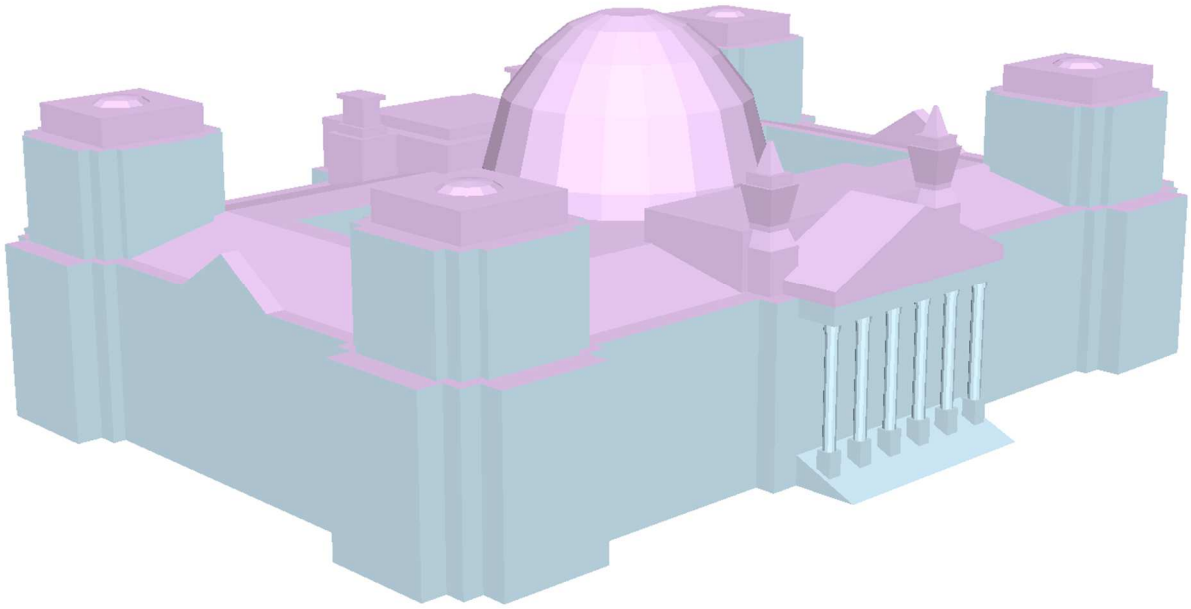

Figure 4: Original LOD2 3D CityGML representation of the German Reichstag Building, Berlin.

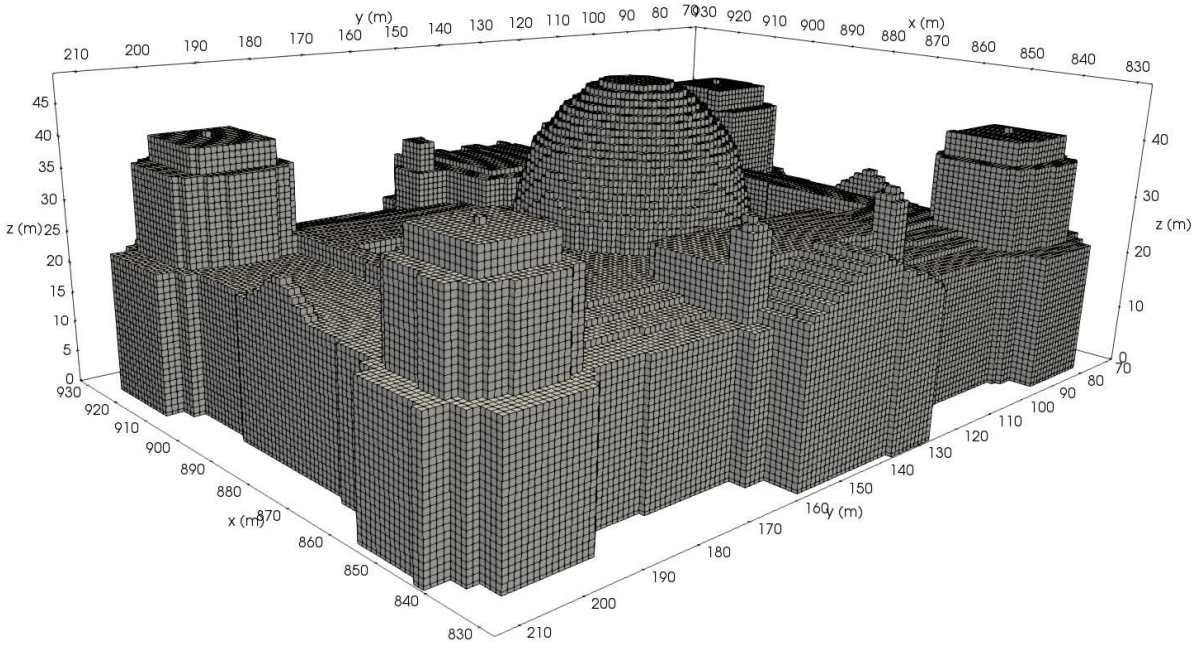

Figure 5: 3D visualization of a processed and rasterized CityGML representation of the Reichstag building in Berlin, Germany, at a grid spacing of 1 m.

The design of 3D data products at the national level has been considered for several years now (Stoter et al., 2015). Optical imagery with different angles is used to generate elevation models with photogrammetric technologies. This technique can be applied to airborne imagery using several images from the same flight but it can also be improved with UAVs or street imagery. LiDAR offers a

higher geometric accuracy but a smaller coverage. In some countries, nationwide LiDAR products are available (as over the USA, Ching et al., 2009). For some areas, simplified 3D databases can be built automatically out of existing topographic databases based on height attributes. For example, Brasebin et al. (2012) showed that the French national topographic product implementing INSPIRE can be used to analyze sky view factor when aggregated results are needed. The buildings are captured at the gutter level and have metric information about the height of the building.

3.3.4 Crowdsourcing or deep learning methods

Data on building heights can also be crowdsourced directly (Over et al., 2010; Fan and Zipf, 2016) through OSM. OSM has an agreed set of tags for 3D building models including 'height' for height of the building in meters and 'levels' for the number of floors. Although this information is not often added to buildings footprints in OSM, there are some demo cities with this information that can be visualized via the OSM Wiki[3], but building height is often missing (Lao et al., 2018). They report that less than 3% of the buildings globally have a height value and less than 4% have a level value. For the city of Paris, the values are 0.1% and 51.2%, respectively.

It would be possible to use crowdsourcing to manually extract building height or the number of floors from photographs in Google Street View, Mapillary or from other geo-tagged photo repositories such as Flickr. This would involve building a bespoke application for this data collection but it could provide a sample of building heights across a city that would be adequate enough for urban climate modelling purposes. Similarly, an application could be built for collecting this type of data on the ground.

Biljecki et al. (2017) developed a model to predict the building height based on parameters from the building cadaster, the geometry of the buildings from OSM and information from the census. Exploring the combination of different parameters, the mean absolute error of predicted building height varied from 0.8 to 3.1 m. Photograph or image analysis by deep learning is also being developed to extract some morphological parameters. For example, Liang et al. (2017), Gong et al (2018) and Middel et al (2018, 2019) showed that it is possible to analyze Google Street View photographs to map the sky view factor. Zhang et al. (2019) developed a procedural model using neural networks to produce a 3D model of buildings in cities from segmented satellite images, OSM street information, population density and terrain elevation.

4) Architectural parameters

4.1 Description of the parameters and their relevance

Architectural parameters describe the way that buildings are constructed. The choice of building materials and structure modify the heat conduction, and the roof cover and walls affect the radiative exchange with the atmosphere. Hence, they can strongly modify the UHI. Many adaptation strategies are based on modifications of these characteristics because they are relatively easy to implement, since this occurs at building scale. An example are white roofs or walls, which reflect the solar energy towards the sky (see Santamouris, 2014 for a review). Many traditional villages around the Mediterranean Sea are built this way. Another aspect relates to wall insulation, when present; depending if the insulation is on the inside or the outside, the total energy stored during daytime in the building fabric will be different, and hence so will the UHI. Architectural parameters are building

[3] https://wiki.openstreetmap.org/wiki/Simple_3D_buildings

materials, depth, thermal conductivity, and heat capacity (or thermal diffusivity) of all walls or for each layer of wall (e.g., for the layer of structural material and for the insulation layer inside or outside), and of the roofs. A description of the intermediate floors is required if a Building Energy Module (BEM, e.g. Salamanca et al., 2009; Pigeon et al., 2014) is included in the UCM, as is the case, for example, for BEP (Building Energy Performance) and TEB (Town Energy Balance) models. The presence of windows is also important, as it modifies the internal energy balance of the buildings and the subsequent anthropogenic heat emissions due to domestic heating or air-conditioning. Furthermore, windows have different thermal and radiative properties than walls. In particular, the knowledge of both the location and the window fraction per surface element (or per square meter) become important for microscale simulations where the spatial variations of windows can be explicitly represented (Resler et al., 2017). Also, green roofs and facades are a known strategy for the cooling of cities and should thus be taken into account in modelling studies. This is not only important at the urban scale, but particularly at the microscale, as green elements are altering the thermodynamic conditions, especially in the vicinity of their location.

To summarize, the most important parameters include:

- albedo and emissivity of walls and roofs;
- thermal properties (thermal conductivity and heat capacity) and thickness of the layers of materials constituting the roofs and walls;
- the fraction of windows on the external facades;
- the thermal characteristics of the windows;
- the presence of shelters on windows

4.2 Developing comprehensive architectural databases

There exists for no large city an exhaustive database on building architectural practices or building material characteristics at the scale of the building. Even with the increasing use of Building Information Modelling (BIM), such information is, at best, limited to a very small part of buildings in a city, which are mostly recently constructed large buildings. Due to the specific lack of spatial data about architectural structures at the city scale, how can such parameters be defined? The main problem is that all of these parameters depend on the materials used to construct the buildings, and this varies considerably across time and space. Almost no generic information exists about these architectural features and certainly not globally. Even at the local scale, the only comprehensive way to gather data on building materials is limited to roof material identification using hyper-spectral remote-sensing (Heiden et al., 2007). Architectural knowledge is mostly available at the building scale, but not extensively. The issue arises of how to upscale such sparse information. Some attempts have been made to derive surface parameters like wall albedo and emissivity as well as window fractions by manual observations within case studies (Resler et al., 2017).

As databases of the architectural parameters needed by UCMs and ORMs are not available, an approximate approach is possible, i.e., the development of comprehensive architectural databases. The objective is to describe, for several archetypes of buildings, their typical architectural characteristics. For example, the material of the main structural wall is documented. To link this information to the models, the thermal characteristics of these materials are recorded. While these parameters still only represent an educated guess, they allow the variability in building architectural characteristics to be taken into account. Jackson et al. (2010) have described the thermal and radiative properties of buildings for 4 building types in 33 regions in the world (defined by socio-economics, architectural practices and climate). Tornay et al. (2017), using an architect's expertise

approach, have defined building archetypes in each of the 95 administrative regions of France. This has allowed the spatial variability between the different cities in France to be described, but not within each city. Such building architectural databases can be completed based on expertise or crowdsourcing. One persistent issue is the fact that there is generally no cadaster for monitoring changes in the state of buildings (e.g., due to improved insulation during reconstruction works). The actual building properties are thus likely to be very different from those at construction.

However, the architectural database would describe elements that are at the building scale (but without spatial information, since they are only referencing archetypes). Therefore, in a second step, once such a comprehensive architectural database is completed, a link must be made between the building archetypes documented within the database and the buildings in the city. This can be accomplished by observing the characteristics of the buildings in the field. Four building parameters that are relevant for how the building has been built are:

- the period of construction of the building;
- the use of the building (i.e., residential, commercial, offices);
- the building type (i.e., house, mid-rise building, high-rise building, industrial building); and
- the geographic area (this can be at the scale of neighborhoods, the city, the region, or the country), but can also be related to the climate type.

Mapping this information is possible. For example, the building type can be linked directly to the LCZ (see section 2). The building use and period of construction can come from socio-economic databases and a census (see section 5). We now concentrate on how to gather the architectural information itself.

4.3 Methodologies to gather architectural information

4.3.1 Identification of representative archetypes

Taking into account the four inputs of urban typology, building use, construction period and location, it is possible to define representative building archetypes with their building materials using an expert approach. Since 1990, there have been research projects that have developed expert approaches to identify representative buildings on a regional and national scale. The first step is to identify representative buildings to determine the characteristics of their construction. Such work has been undertaken in Germany with the German building stock (Ebel et al., 1990), in Canada by Canada's National Research Council Institute for Research in Construction (Jackson et al., 2010), and in France, with a detailed analysis of existing residential real estate (RAGE, 2012). More recently, initiatives on a larger scale have been undertaken, e.g., the "Tabula" project (Loga et al., 2016), which has developed a catalog of national building typologies representing the residential building stock in 20 European countries. These studies only deal with residential building use; however, the proposed methodology can be applied to all building uses, and to a larger geographical scale.

The construction of the comprehensive architectural database consists of describing the construction materials used for the building archetypes. A literature review on construction practices in France by Tornay et al. (2017) revealed that several construction systems might coexist for one combination of urban typology, building use, construction period and geographical location. As a consequence, the French database on building archetypes provides three options for the construction of the wall, roof and floor.

These expert approaches require an exhaustive collection of information (Figure 6), knowledge to generate reference building databases (step 1) and characterize their materials (step 2). To extract results (step 3) that are exploitable and realistic, they can be cross checked with an extended literature review and statistical data in order to be generalized. At present, existing databases are still heterogeneous, and the main factor to be further developed is the different uses of buildings (not just residential buildings).

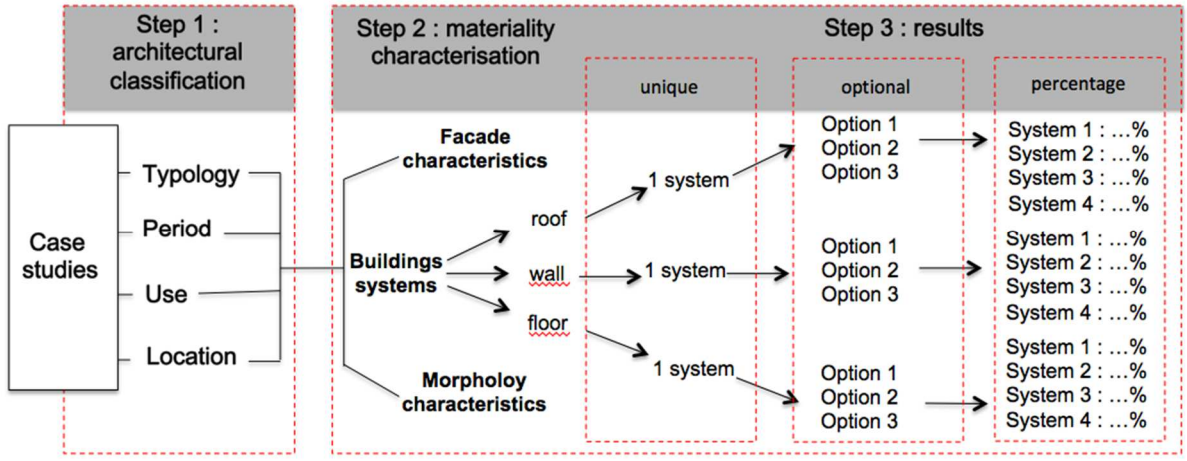

Figure 6. Overview of the expert methodology to define the architectural parameters

Known for its compact high-rise urban environment, which is distinctly different from typical European or North American settings, Hong Kong provides an exemplary testing ground for the approach described here. Although databases detailing Hong Kong's heterogeneous urban morphology are available, they are not sufficient for characterizing building architectural parameters. Several unique features in building characteristics have been identified through reviewing the urban development history and building ordinances in Hong Kong (Shelton et al., 2011; Wong, 2014), from interviews with practitioners, and from field surveys. First, there is a small variation in construction materials across buildings of different type, use and age. Reinforced concrete, often in the form of prefabricated component blocks, is the predominant structural material for all building types. A notable exception is tall commercial skyscrapers with curtain walls. Wall insulation is not of major concern for the subtropical climate in Hong Kong. Secondly, the range of building types defined by Tornay et al. (2017) and LCZs cannot adequately represent bulky building structures prevalent in Hong Kong. Examples include massive mid-rise shopping complexes and connected podiums at the base of multiple tall towers built in order to maximize the use of limited land resources. To consider these extensive mid-rise buildings as detached mid-rise buildings (from the point of view of construction materials) seems, however, an acceptable approximation, and is reconciled with the LCZ building description. Thirdly, an interesting pattern is observed among residential buildings. Prior to the emergence of high-rise private housing estates in the 1960s, people lived in uniform rows of Tong Lau or tenement houses. Private residential buildings then evolved in height, podium scale, and window designs through the years (Figure 7). The architectural characteristics of public rental housing estates, despite the various block types, have remained largely similar since their introduction in the 1950s.

Figure 7: Temporal evolution of public and private residential buildings in Hong Kong.

4.3.2 Remote sensing and image processing

Building type can be evaluated with photogrammetric software, like the open source Micmac software, using available images, provided that several perspectives are available for the same point. The detection of windows on facades has been investigated using photographs at the street level (Burochin et al., 2009, 2010). Facade characterization from aerial imagery has also been considered (Burochin et al., 2014).

The building period of construction can also be derived from the recent digitization of archival photogrammetric aerial images captured by geographic agencies in many countries, which often occurs at least every 5 years and has taken place since the middle of the 20th century (Giordano et al., 2017). Building use evaluation, and more generally land use evaluation, based on image processing technologies and aerial or terrestrial images is a hot research topic (Li et al., 2017). For example, the spatial arrangement and composition of the extracted features has been analyzed to yield building use and land use data.

4.3.3 Crowdsourcing

The description of a building, in contrast to morphological or land cover parameters, is easily understandable by people. This is because the scale is well identified (while the spatial scale of land cover elements is larger and only directly visible from above) and it is thematically comprehensive (contrary to most morphological indicators such as λ_p, λ_w, roughness length, etc.). This makes crowdsourcing a suitable method for gathering such architectural information. Within the WUDAPT initiative (Ching et al., 2019), a smart-phone application has been developed to gather 1) building use; 2) building age; 3) building material; 4) whether the buildings are painted (and the color if so); 5) window-to-wall ratio coverage; 6) existence and location of heating or air conditioning systems; 7) roof covering materials; and 8) number of floors. The latter is a proxy for building height (and hence building type). Mhedhbi et al. (2019) adapted this questionnaire as a Google form accessible through social media (Facebook), and gathered information on more than 100 buildings in the city of Tunis. For example, half of the houses in LCZ 3 and low historical core buildings in LCZ 6 are described as

having ‘a few windows’, while the other half were described as having ‘many regularly spaced windows’ plus ‘glass windows’, while more modern mid-rise buildings were mostly reported with ‘many regularly spaced windows’. The buildings in the sample were spread across the city and across various LCZs and hence provide valuable information for describing typical buildings in the various city neighborhoods.

Crowdsourcing of building age can be difficult, i.e., it requires some expert knowledge or training of the crowd to identify features associated with different age categories, where these can also vary by country. As with height, building age can be collected using a crowdsourcing application like that described previously. An alternative approach is to build a model that predicts building age if a sample of building age data exists. For example, Rosser et al. (2019) built a predictive model of building age using parameters such as area of the building footprint, the perimeter, the number of buildings in the block, etc. By aggregating the results from different methods and using some simple rules for filtering the data, the overall accuracy achieved was 77%. A similar approach was undertaken by Biljecki and Sindram (2017), who fitted a random forest regression model to predict building age in Rotterdam. Their results indicated that the model was limited in predicting the exact age of a building but that it could be used to predict the approximate period of construction. For the purpose of urban climate modelling, prediction of the age category would be sufficient.

5) Socio-economic data and building use

5.1 Description of the parameters and their relevance

Human activities in cities produce direct releases of heat and water vapor to the atmosphere. Domestic heating releases heat directly through chimneys and within the buildings (but is transferred later to the atmosphere through heat conduction, air leakages, or venting). Air-conditioning extracts the heat inside and releases it outside. Combustion heat from cars is usually smaller than the contribution from buildings, but it can be significant in the immediate vicinity of major roads (Pigeon et al., 2007). For US cities, Sailor et al. (2015) report that combustion heat from vehicles accounts for 40% of the anthropogenic heat flux during the summer months when building heating does not play a major role. Heavy industries and power plants also release a considerable amount of heat. All these fluxes are known as anthropogenic fluxes.

Anthropogenic heat fluxes due to traffic and industrial activities need to be specified, typically through emission inventories, even though these can provide only first-guess estimates of the true emissions. Traffic can also influence road temperature (due to the pneumatic contact of the tires with the road during car movement) and radiative exchanges (by the presence of the cars), or the air turbulence produced by vehicles (Kastner-Klein et al. 2001). Such processes can be parameterized into UCMs (Khalifa et al 2016) or studied in micro-scale models. They need specific information on traffic, such as traffic density or mean vehicle speed.

Building related anthropogenic fluxes need to be specified if no BEM is implemented in the urban canopy model. If a BEM is used, anthropogenic heat fluxes due to building energy consumption may be computed as a function of the prevailing meteorological conditions. This requires knowledge of how (and eventually by whom) the buildings are used and inhabited (Schoetter et al., 2017). Important parameters include:

- Population density;
- Fraction of each use in the building (e.g., to describe complex patterns such as commercial use on the ground floor and then offices and residential apartments above);

- Schedules of building occupancy during daytime/nighttime/holidays;
- Internal heat release describing how much energy per square meter comes from electric appliances, cooking, etc.;
- The type of domestic heating;
- Domestic heating target temperatures (day/night); and
- Air conditioning use and, if available, target temperatures (day/night).

5.2 Methodologies to gather uses, socio-economic and anthropogenic heat parameters

5.2.1 From inventories

The anthropogenic heat flux is mainly due to traffic, industrial activities, the building sector, and human and animal metabolism. It can be quantified using three strategies (Sailor, 2011): observing the urban surface energy balance; inventories (top-down approach); and building energy consumption modelling (bottom-up approach).

The observation-based method can only be employed for locations with available urban surface energy balance observations (e.g. Pigeon et al., 2007 for Toulouse). This method is also subject to uncertainty due to observation errors, especially since the anthropogenic heat flux is determined as the residual of all other fluxes.

The top-down inventory method consists of disaggregating large-scale energy consumption data in space and time. It is the state-of-the-art for the construction of global anthropogenic heat flux databases. Country-scale data on primary energy consumption can be taken from the United States Energy Information Administration. Flanner (2009) and Allen et al. (2011) spatially disaggregated these data using population density data as a weight.

However, population density is not easy to obtain. For example, in Figure 8, we compare two data sets of residential population density for the year 2010 in the Paris Metropolitan Region. The Gridded Population of the World Version 4 (GPWv4; CIESIN, 2017) combines tabular counts of population by administrative area with georeferenced data on administrative boundaries and land cover (e.g., water bodies). The spatial resolution is 30'' (~1 km). Another estimation of residential population was constructed during the French research project MAPUCE. It is based on gridded population density (200 m resolution) compiled by the French Institute for Economics and Statistics[4] (INSEE). Population is disaggregated in space using the total residential floor area as a spatial weight. The residential floor area is determined using the building's geometric properties (i.e., footprint, height) and use, which is available from administrative data and was processed following Bocher et al. (2018). The GPWv4 captures the population density pattern well in the Paris Metropolitan Region. However, it misses sharp heterogeneities at the kilometric scale compared to the MAPUCE data set, which uses the fine building database. We, therefore, conclude that the effective resolution of the GPWv4 population density, which has the advantage of global availability, is in reality coarser than 1 km² for the Paris Metropolitan Region.

[4] https://www.data.gouv.fr/fr/datasets/donnees-carroyees-a-200-m-sur-la-population/

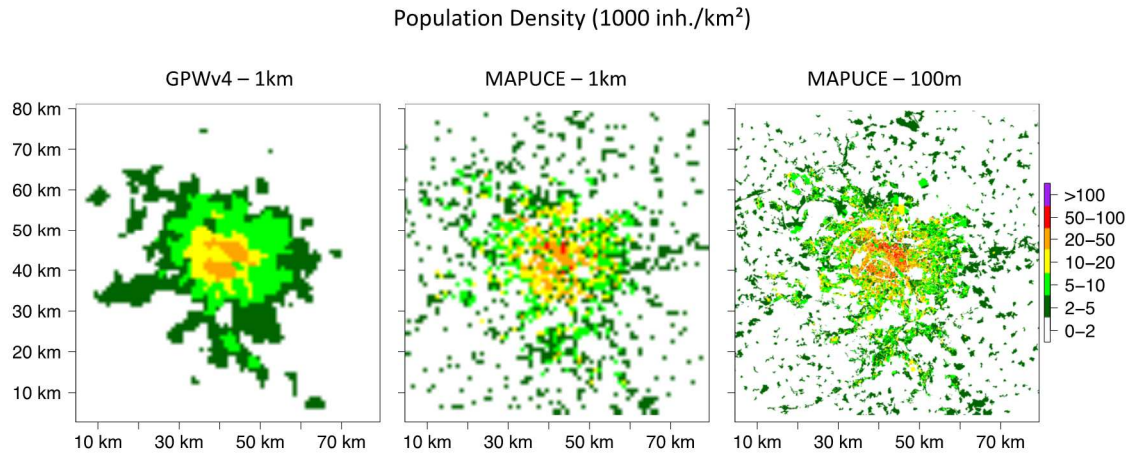

Figure 8: Comparison, for the agglomeration of Paris (France), between residential population density in 2010 of (left) the World Version 4 (GPWv4) at a 1 km of resolution, and (middle and right) from the French MAPUCE project at 1 km and 100 m resolution.

Dong et al. (2017) refined the top-down methodology by also considering, in addition to population density, the sector-specific energy consumption included in the International Energy Agency Energy Balances. They disaggregated only the final energy consumed in the commercial, residential and transportation sectors as a function of population density. The final energy consumed by agriculture and industry was distributed uniformly in the populated area, and the difference between the primary and final energy consumption was distributed evenly across the country. The exact location of power plants, electricity grid lines, and agricultural and industrial facilities would need to be known for a more precise spatial disaggregation of primary energy consumption. Disaggregation in time is typically done by specifying energy consumption schedules (day of year, day of week, hour of day). Allen et al. (2011) took information on normal business hours, weekend days and fixed public holidays from the Lonely Planet and the World Travel guide. They further assumed that the daily cycle of building energy consumption is the same as that for traffic during the weekend, which was taken from an inventory. Dong et al. (2017) applied a daily cycle of building energy consumption derived for Tokyo to all cities in the world. Allen et al. (2011) and Dong et al. (2017) assumed that due to heating and air conditioning, building energy consumption depends on the prevailing air temperature. Anthropogenic heat flux due to metabolic heat was specified by Allen et al. (2011) and Dong et al. (2017) as a function of the population density, considering the daily cycle of people's activities. Uncertainties arise if temporal variability of population density is not taken into account.

The bottom-up method consists of simulating building energy consumption as a function of the building type, use and the prevailing meteorological conditions. Heiple and Sailor (2008) simulated the energy budget of prototype buildings in the United States and weighted the results with their spatial frequency of occurrence. They found that their bottom-up method led to a larger spatial and temporal variability in building energy consumption than a top-down method. The main reason is that the bottom-up method allows the temporal variability of meteorological conditions, human behaviour and the spatial variability of building type and use to be considered. As an alternative to the simulation of prototype buildings, the building energy consumption can be obtained from a building energy module implemented in an urban canopy model. Both building-scale and district-scale bottom-up approaches rely on detailed knowledge of building use, occupation, heating and air conditioning practices, and internal loads. To our knowledge, no global scale data sets exist that could directly provide such parameters. Instead, urban climate modellers specify these parameters

for a specific city/country based on expertise or national data sets. Kikegawa et al. (2003) simulated a central business district in Tokyo and initialized the relevant parameters (especially the design temperature for air conditioning) based on their expertise for typical office buildings in Tokyo. Oleson et al. (2008) assumed a heating and air conditioning design temperature of 18°C and 24°C for Mexico City and Vancouver, respectively. Salamanca and Martilli (2009) performed simulations for the BUBBLE observation campaign in Basel, Switzerland. They derived information on the indoor design temperature for air conditioning from indoor air temperature sensors and used the population density and inventories to estimate the internal load due to metabolism and electrical appliances. Schoetter et al. (2017) estimated the fractional building use in French cities based on administrative data sets and the ratio between the number of inhabitants and the total floor area. They initialized the parameters related to heating design temperature and internal heat release based on a combination of surveys and statistical models developed by Bourgeois et al. (2017). A sensitivity study for Toulouse, France, showed that taking the variety of building use and behavior into account is crucial for an accurate simulation of the spatial-temporal variability of building energy consumption.

5.2.2 Crowdsourcing

It is also possible to obtain building use from crowdsourcing, e.g. OSM has specific tags for users to indicate the building function. In a study by Fonte et al. (2018), OSM data were extracted for a section of the city of Milan. By analyzing the tags associated with buildings as well as the points of interest layer, more than 80% of the buildings in Milan could be assigned a building function. Moreover, the analysis could also help to identify mixed function building types, e.g. buildings that might be commercial on the ground floor but contain residences above. A similar study was undertaken by Kunze and Hecht (2015), who specifically focused on using OSM to determine the amount of non-residential use in residential buildings in order to calculate the non-residential floor area (Kunze and Hecht, 2015). Other crowdsourced information can also be used to infer building use. In Fonte et al. (2018), crowdsourced data from Facebook and Foursquare were additionally used to fill in some of the gaps from OSM regarding building use. Using only Foursquare, Spyratos et al. (2017) classified buildings in Amsterdam according to building use types. The best results were obtained for hotels, restaurants, cafes, and retail establishments while such an approach is less suited for identification of industrial use.

Finally, as outlined above, a bespoke application could be built to crowdsource building use from photographs or collected by volunteers on the ground using a mobile app. An application specifically developed for crowdsourcing building use (and other land cover/land use attributes) has been developed by IIASA and IGN France called PAYSAGES, which sends volunteers to specific buildings within a city and asks them to provide information on building function along with a photograph of the building (Olteanu-Raimond et al., 2018). The purpose is to enhance IGN's building database, which does not currently contain this type of building use information.

6) Urban vegetation

6.1 Description of the parameters and their relevance

Cities are very heterogeneous environments, composed of artificial and natural surfaces. The proportion of urban natural surfaces varies not only from one city to another (Fuller and Gaston, 2009; MIT Senseable City Lab, 2018) but varies also strongly within cities (Lu et al., 2017). In addition

to these spatial variations, there is a wide variety of vegetation systems (lawns, street trees, gardens and urban parks, green walls and roofs, etc.), and plant species, whose physiological characteristics evolve over time, according to their own vegetative cycles.

It is known that the bare soil and vegetation present in cities influence the urban micro-climate through various physical processes: air cooling by water retention and consequent evaporation from bare soil and evapotranspiration from plants; modification of the radiative balance and the airflow in urban canyons in the case of urban trees; decreased heat flux emitted by buildings covered with green envelopes; and horizontal advection of fresh air from larger scale structures (urban parks, green belts, etc.). High vegetation (i.e., trees) can also have an adverse effect since they can reduce the ventilation at street level and thus lead to increased pollutant concentrations near the surface (Salmond et al., 2013, Santiago et al 2017). Therefore, taking into account natural surfaces and their spatial and temporal heterogeneities in UCMs allows for a more realistic simulation of micro-climatic conditions (Shashua-Bar and Hoffman, 2000; Grimmond et al., 2011). In fact, over the past decade, more and more models have incorporated the modelling of physical processes associated with the presence of urban vegetation. These models require a detailed description of vegetation systems, which involves knowing at least the following parameters:

- the respective proportions of ground covered by bare soil and by different vegetation strata (herbaceous, shrub and tree layers) as well as the building surfaces covered by green envelopes (facades and roofs);
- the depths and textures of structural layers for urban soils;
- the physiological characteristics of the vegetation, which are species-dependent, namely leaf area index, stomatal resistance, albedo, emissivity, root depths and densities; and
- for the tree strata, tree type (deciduous/evergreen) and geometric characteristics, namely tree height as well as crown size, shape and trunk diameter.

However, the level of detail of the physical processes accounted for, and therefore the required descriptive parameters, will vary depending on the models. For microscale ORMs, for example, trees are usually described by means of three-dimensional leaf (or plant) area density fields.

6.2 Methodologies to collect vegetation parameters at mesoscale

Currently, no global data are available to describe urban vegetation in terms of coverage and characteristics. At best, through local administrative databases or open data, it is possible to build land use maps for vegetated surfaces, often with an incomplete coverage, and without the distinction of strata (even if street trees are sometimes listed). Furthermore, the geometric and physiological parameters inherent to vegetation species are unknown. Bare soils are rarely mapped. The properties of urban soils, which are very heterogeneous because they have been significantly reworked, are also not documented in such administrative databases.

However, the land cover map obtained by this approach may show an incomplete representation of vegetated surfaces (Figure 9b) compared to reality (Figure 9a) and without distinction of vegetation strata. This results in a significant proportion of urban surfaces that remain undefined, and on which the modeller will have to make assumptions to simulate the urban micro-climate. For street trees, data may be available for some cities via local authorities or through new initiatives based on open data, such as that of Treepedia developed by the MIT Senseable City Lab (2018), which offers a Green Vegetation Index (coverage of urban surfaces by tree crowns) based on Google Street View panoramas (Li et al., 2015; Seiferling et al., 2017). For some European cities, data on urban tree coverage is also available through the Copernicus Urban Atlas portal (EEA, 2018). Moreover, some

individual cities have decided to provide various city-descriptive data sets freely, including trees. For example, the city of Berlin, Germany, has developed the online data browser FIS-Broker.

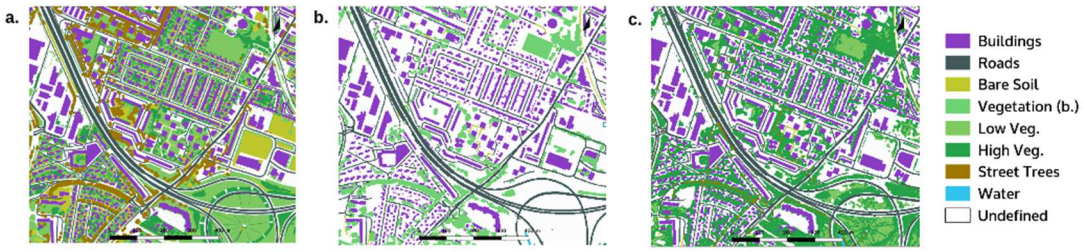

Figure 9. Land cover maps of a neighborhood in Toulouse (France) constructed by different approaches (de Munck et al., 2018): a) manual scanning of vegetation strata from orthophotos (closest to reality), b) from the BD-TOPO national database, and c) vegetation extracted from Pleiades imagery.

In this context of a lack of data, remote sensing (satellite or airborne images) or LiDAR can fill some of these gaps at high resolution, in particular for the coverage and spatial heterogeneity of the various vegetation strata. In the case of satellites, revisits also allow for the monitoring of urban vegetation temporal dynamics, which is highly relevant in terms of modelling. Methodologically, this approach passes, at least, by the spectral or multispectral treatment of the NDVI and possibly the Normalized Difference Water Index (NDWI) to determine the vegetation cover in the city. Thresholding of the NDVI is generally not sufficient to distinguish vegetation strata. For this, a more robust approach is to retrieve the height of vegetation from the difference between a Digital Terrain Model (DTM) of high resolution (or a DSM) and a Digital Elevation Model (DEM). This approach can also be refined by taking into account the contextual information of images such as granulometry or texture (Herold et al., 2003). This type of approach makes it possible to more accurately characterize the coverage of high and low vegetation in cities.

Remote sensing approaches, even if they show good results, have certain challenges and limitations, such as the need for very fine spatial resolutions to allow for the classification of vegetation into different layers, the complex treatment of shadows (from buildings and trees), and the impossibility of characterizing surfaces hidden by tree crowns. Also, even if a key biophysical parameter such as the leaf area index of urban vegetation can technically be retrieved from remote sensing images (Jensen and Hardin, 2005; Alonzo et al., 2015), very little data are currently readily available for modelers. Finally, these approaches are not designed to answer questions related to the nature of the vegetation species present (crown forms, albedo, etc.). To determine these parameters, modellers generally use vegetation characteristics derived for rural areas from the scientific literature (assuming they are valid in urban areas) or use *in-situ* surveys at the microscale.

6.3 Methodologies to collect vegetation parameters at microscale

At the microscale, city trees, for example, are registered in tree cadasters, but the registered details differ from city to city. In Germany, the cadasters only cover municipal trees. Urban woods and trees on private areas are not included, which requires special methods to estimate this information from these areas.

ORMs with a very high spatial resolution (down to 1 meter) require even more detailed information concerning high vegetation like groves and trees than models with coarser spatial resolutions. Vegetation is usually represented in building-resolving UCMs as three-dimensional arrays of the leaf area density (LAD). There are no common approaches to estimating three-dimensional LAD-

information of high vegetation. The discrimination between high and low vegetation, as well as individual tree detection, is generally done using very high resolution aerial or satellite imagery and associated 3D data (Iovan et al., 2008; Lefebvre et al., 2017). Species extraction can also rely on new sensors and hyperspectral imagery (Ouerghemmi et al., 2018).

As part of the German joint project "Urban Climate under Change"[5], the Leibniz University of Hannover has developed an approach to estimate three-dimensional LAD for a given single tree based on one-dimensional LAD-profiles after Lalic and Mihailovic (2004) and for tree patches after the approach of Markkanen et al. (2003). However, this method requires a large set of tree-physiological input parameters, which are difficult or rare to obtain. Therefore, look-up tables have been developed, which at least provide estimates of these parameters for common tree species. Thus, as a minimum requirement for the simulations, the location and tree height are needed. Furthermore, currently available methods for estimating the leaf area density profiles have all been developed for closed homogeneous canopies. Their application to representing individual trees is, thus, rather questionable.

Another very straightforward method for mapping vegetation parameters is direct observation. Although this cannot be carried out for large areas, direct observation yields valuable information at microscale, which might even be interpolated to larger scales. During a field campaign, a trained person can classify the type of vegetation and trees directly and with high accuracy. Also basic tree parameters such as height or diameter at breast height can be measured relatively easily. Much more effort is needed to directly measure the quantity of leaves. Breda (2003) has reviewed different methods for LAI estimation. Direct measurements require harvesting or litter collection of the vegetation. As this is destructive and cannot be applied to larger areas or trees, indirect methods have been developed; these methods measure the radiation through the canopy or make use of hemispherical image analysis (e.g., fish-eye photography). For describing the shape of the tree or the leaf area density, 3D portable (ground-based) LiDAR imaging can be used (Hosoi and Omasa, 2006), which yields more accurate results than airborne LiDAR imaging.

7) Discussion

Once a UCM or ORM for an urban climate study has been defined, a remaining challenge is to select the relevant sources for the city descriptive data and to process them for the required spatial and temporal coverage. In the previous sections, many methods for deriving the main parameters have been described. This section discusses a set of key issues associated with accessing the corresponding data.

Often a combination of methods and data sets is needed to retrieve the required information for all five types of parameter, i.e., land use and land cover; morphological parameters; architectural parameters; socio-economic parameters and building use; and urban vegetation. They often require expert knowledge, and the process of mapping parameters from raw data can be time consuming. Remote sensing techniques are useful for deriving UCM and ORM parameters regarding land cover, urban morphology and vegetation. However, they provide insufficient information on architectural and socio-economic characteristics. This becomes problematic for studies dealing with vulnerability and exposure to weather and climate risks. The main disadvantages of land-based administrative information are the slow rate at which the information is updated and the sensitive, non-public

[5] http://uc2-program.org

nature of such information. Crowdsourced data have addressed the former issue, and although not as complete as official government data in some areas, the literature reports sufficient data quality (Haklay, 2010). In the future, deep learning methods will likely yield relevant results in some domains where large amounts of available data can be used to train automatic classifiers.

7.1 Licensing issues

Urban data from authorities, i.e., public organizations mandated to organize the coproduction and maintenance of data required for public action, are usually available for free, at least for research but not always under free licenses. Some countries, regions, or cities already provide data for free, accessible from data platforms (e.g. France, Berlin (Germany), Hamburg (Germany), Federal State of North-Rhine-Westfalia (Germany), Helsinki (Finland)). Other cities provide the data only under certain conditions and restrictions. In France, the BD-TOPO national database (IGN, 2018) is available for free for research, education and public missions. It is generally used to construct urban land use maps. Administrations of large cities generally possess more precise cadaster data than information covering national or international territories. Such administrative data can also cover many aspects including morphological, socio-economic, architectural, financial and administrative features. To use such data sets can lead to a large workload because of the city-specific data form and content. Even if GIS-based vector maps can be easily transformed into other data formats, their adaptation for use in a UCM requires effort and programming skills. Furthermore, the exploitation of some data, especially for ORM models that require data at the building scale, can require formal authorization for use due to data protection of people's privacy (e.g., as required by the European General Data Protection Regulation (GDPR) enforced on 25 May 2018).

The area of data licenses is a complex domain. The Open ELS project, which is led by several mapping agencies across Europe and brought together through the EuroGeographics organization, has even proposed a new specific open license. In Europe, after a public consultation on the Directive for Public Sector Information, a proposal for revision has been adopted by the Commission and is being discussed by the Parliament and Council. This revision would lead to greater availability of geographic data under an open license for microscale analysis. The licensing problems associated with the use of public national data also underline the interest in using open databases such as OSM.

7.2 Cataloguing issues

It is currently still hard to search and find suitable data sets, even if the situation is improving. Existing catalogues and portals are dedicated either to a territory like the French géoportail[6] or data.gouv portals, to a data provider like the EuroGeographics portal[7], to a technology like the Copernicus[8] portal, dedicated to distributing products derived from Earth Observation data, and funded by the European Union program Copernicus, or to a community like the enes[9] portal for climate models (output data).

Recently Google has launched a prototype search engine that is dedicated to earth science raster data sets called Google Dataset Search. This is still rather empty but it illustrates a key issue: the availability of metadata (see section 7.3). An interesting category of national portals has emerged over the last decade to distribute and view old maps and old aerial photographs in order to compare them with current databases.

---

[6] http://geoportail.gouv.fr, http://www.geocatalogue.fr/ , https://www.data.gouv.fr/en/

[7] http://www.eurogeographics.org/products-and-services

[8] http://www.copernicus.eu/main/climate-change , http://land.copernicus.eu/

[9] https://portal.enes.org/

7.3 Data quality

Standards have been defined for portals to exchange and describe information (e.g., DCAT, CSW). However, there is too little metadata available that documents existing products. Although the INSPIRE directive has made it mandatory to document authoritative data sets with standard metadata, some metadata are still missing and a dashboard has been set up to monitor the progress of metadata documentation.

Metadata are also important for documenting possible sources of uncertainties (in terms of precision and of accuracy) in UCMs and ORMs. Metadata also allow the temporal coverage to be specified for which the description of a given city is valid.

The increasing availability, diversity and amount of urban data, together with the increasing availability and affordability of computing resources, has made it possible to increase the spatial resolution of urban modelling for climate studies, but also to be able to select a preferred spatial resolution. This flexibility has highlighted the issue of the trade-off between spatial and semantic (thematic) resolution. In urban models that contain both physical and socioeconomic dimensions, pushing the spatial resolution above or below certain levels is not always advisable, because some socio-economic phenomena do not occur at very fine or at very coarse scales; forcing them to very large or very small scales would misrepresent human dynamics (see e.g., Votsis, 2017). ORMs using a grid mesh of less than 10 m can take many characteristics of the urban surface into account so reaching this amount of detail is almost infeasible at the city level, e.g. the shape of the trees. Until detailed data sets such as CityGML or LiDAR point clouds become more easily accessible, generalizations and assumptions will need to be applied to these models.

7.4 Open data

Accessing city-descriptive data from different providers to design an urban climate model experiment that is replicable in Europe typically requires accessing different data repositories, which could have different licensing schemes, and to become familiar with different implementations. As this increases the costs (in terms of time and effort) of using these data sets, a trade-off needs to be made between accuracy and detail.

This has been addressed by the OSM project in which a unified framework has been proposed to collaboratively build an open database, largely from integrating existing open data in different countries. The geographic features represented in this database (e.g., buildings, highways, land use, etc.) are organized in a relational database model by a set of elements (NODES, WAYS and RELATIONS). Each element is described by tags (key=value). For example, a building is described by a WAY element (with at least one “building=yes” key-value[10]), which refers to a NODE collection (each node represents a coordinate in latitude and longitude). Since its creation in 2004, OSM has gained popularity and grown to over 1 million map contributors and more than 1 million registered users. OSM's success lies in the flexibility of its data model and project operation. The community is voluntary so adherence to semantic tagging is also voluntary. These lack of restrictions, the absence of procedures to capture data and quality control, plus the use of a non-standard data model, have created some difficulties when using OSM data in GIS applications. Extracting OSM data and adapting them to user needs is a current topic of research interest as it requires consideration of data matching, web semantics, data fusion, data quality, etc. The approach proposed by the ELF project, set up by national mapping and cadastral agencies, is to provide a mediation layer above the authoritative data served by national legal organizations, which could be a way to facilitate the use

[10] https://wiki.openstreetmap.org/wiki/Key:building

of OSM data (Jakobsson, 2012). An important benefit from this approach and the follow up Open European Location Service started by EuroGeographics is the connection to the INSPIRE framework, and hence to authoritative data producers, the European environment directorate and policy makers.

7.5 Research challenges for the next decade

As described above, some of the methods to produce the input parameters for the UCMs and ORMs are now mature. One should note that both improvements were necessary in the gathering of the urban input parameters and the way that they are taken into account by urban climate models.

These mature developments may still not yet be widely spread among modellers. This may be due to human or monetary costs or licensing issues. Such mature methods encompass, for example, the LCZ classification (Stewart and Oke 2012, section 2.1) or the retrieval of planar and volumetric morphological parameters using remote sensing. Even deep-learning techniques have recently been used to retrieve sky-view factors for entire cities from photographs at street level.

However, several scientific challenges arise, either related to model development or the production of the urban parameters. Four main challenges can be identified:

1. One challenge is the ability to produce parameters at global scale, especially for climate modelling using urbanized earth system models. Such climate models are expected to reach kilometric resolution in the near future.
2. Another challenge is to extend the processes (e.g., the building energy budget, more detailed representation of the urban radiative exchanges) modelled by UCMs and ORMs, in order to be able to provide finer information and integrated urban services. Albeit, these new processes will require new parameters that are difficult to acquire, such as socio-economic ones.
3. Many other research or operational communities gather and use urban data, such as Building Information Modelling. To incorporate such data, the atmospheric modelling community will need to use data standards from other scientific fields. However, this raises the issue of being able to use information not initially built by or for urban climatologists and modelers. The use of open databases, and more generally the development of open data policies at state and world levels, should be encouraged.
4. Finally, exploration of new techniques such as crowdsourcing and deep-learning should be encouraged for the production of any type of urban parameter. Such methods should ideally be conceived in order to be transferable where data are lacking, eventually going as far as reconstructing plausible city information if necessary. This links directly with challenge 1.

An interdisciplinary approach will be necessary to tackle these scientific challenges.

7.6 From data of various origins to Urban Climate Services

The trade-off between spatial and thematic resolution is complemented by the trade-off between regular and irregular geometry when representing the urban area. In particular, some types of socioeconomic information are not well-represented in a pixel-based format, but require irregular polygons for representation, e.g., postcode polygons. Such information could be rasterized and harmonized with other pixel-based information, but this nevertheless creates an artificially high resolution and misrepresents values along the boundaries between polygons.

Studies within the FP7 project ToPDAd and the H2020 project EU-MACS have shown that, in the context of climate services and of climate adaptation at regional scales (in this context, metropolitan

regions), in which accurate representation of socioeconomic information is crucial, the uncertainty introduced by harmonizing climate and socioeconomic data was problematic. First, climate data may be less sensitive to post-processing alteration compared to socioeconomic data; secondly, the chosen geographic unit of analysis is a defining factor in the amount of error introduced (Larosa and Perrels, 2018). This study has highlighted the need for a better conceptualization of the geographic unit of analysis. While the size of grid cells or administrative units is one issue, the main problem is the sensible representation of clusters of human activities in space: a neatly organized gridded representation of the world may be inappropriate altogether, raising the discussion of zonal versus grid-based models and the possibility of urban models that operate at variable spatial resolution and data geometry. The European project URCLIM currently studies the propagation of uncertainties from urban parameters and regional climate models projections at continental scale in the construction of urban climate services at city and infra-city scales.

8 Conclusions

Many different data sets and processing techniques are needed to create a full data set with all of the main parameters required for urban climate modelling. This is time consuming and complex. The parameters needed for UCMs and ORMs have been classified into 5 main categories: land use/land cover classes, morphological information, architectural information, socio-economic parameters (such as the building use or demographic density), and a description of the urban vegetation.

The simplest way to initialize these parameters follows the concept of using only land use and land cover classes as primary data (such as LCZs). This is widely used, and still promising for most applications of mesoscale UCMs, reducing the number of data sets by using (approximate) parameter values from look-up tables for the parameters of the other categories (e.g., for mean building height and density, albedo of the different surfaces and facets, etc.). Joint efforts by the urban climate community will help to improve these tables. This land-cover based approach may be sufficient for numerical weather prediction models, which have a current resolution of 1 km or larger.

However, for providing urban climate services, where stakeholders demand both fine-scale information and applications for specific cities, a more precise description of the city is required for urban climate modelling at mesoscale (100 m) and microscale (1 m) resolutions. Although remote sensing has the advantage of being applicable anywhere, particularly when satellite images are used, it has limitations in providing such high-resolution data sets. Methods that combine 2.5D or 3D buildings and infrastructure cadasters to build the parameters for the model are, however, promising. Unfortunately, such data are not yet available everywhere since many high-resolution data sets are required.

This advocates for removing obstacles to the reuse of available data about public spaces, especially data provided with a well-documented quality by legally mandated organizations, by encouraging user communities to contribute information relevant to climate change studies such as the OSM crowdsourcing community. Such data should also be as homogeneous in their content and quality as possible. This would greatly ease the production of generic urban data for future regional and even global very-high resolution urbanized climate models as well as tailored urban climate services.

Appendix 1: Overview of several global land cover data sets with an urban description

Table A1 provides a list of recently produced data sets that have a focus on land cover classification of urban areas. Contrary to most of the land use/land cover maps that are used in atmospheric modelling, these have, in general, several urban classes, in order to better describe the urban structure. An exception to this is GUF, with only 1 urban class. A further description of these data sets is provided below.

Table A1: overview of several global and continental datasets with urban description

	number of urban classes	spatial extension	spatial resolution
Global Human Settlement Layer	9	Global	100m
Global Urban Footprint	1	Global	12m (public 75m)
Ecoclimap-SG	10 (LCZ)	Global	300m
Jackson et al. (2010)	4	Global	1km
WUDAPT	10 (LCZ)	100s of cities around the world, China, Europe	100m, maps not yet available from a central repository
CORINE	9	Europe	25 ha
Urban Atlas	17	700 large cities in Europe	urban block

- The Global Human Settlement Layer (GHSL) LABEL product (38 m resolution; Pesaresi et al., 2013, 2016) is mainly based on Landsat 8 and uses the Normalized Difference Vegetation Index (NDVI) combined with the SRTM and ASTER-GDEM digital surface models to distinguish roads, built-up areas with different densities (very light/light/medium/strong), and for the strongly built-up areas, the building height (low rise/medium rise/high rise/very high rise).
- The Global Urban Footprint (GUF®) shows the world's human settlement patterns in urban and rural environments at a so far unprecedented spatial resolution of about 12 m (Esch et al., 2017). The GUF® data set reflects the distribution of vertical built-up structures that were derived from radar imagery (SAR) from the German satellites TerraSAR-X and TanDEM-X. Most of the data were collected in 2012 (93% of images recorded in 2011-2012, 7% in 2013-2014). The German Aerospace Center (DLR) is currently adapting the GUF methodology to the use of fully open and free satellite data provided by the European Sentinel-1 (SAR) and Sentinel-2 (multispectral) satellites as well as the US Landsat (multispectral) missions. This activity will lead to the provision of a new suite of global layers under the label World Settlement Footprint (WSF), starting in 2018, with a release of the WSF 2015 (equivalent to the binary GUF, based on a joint analysis of multi-temporal Sentinel

1 and Landsat 8 data for the year 2015) and followed by a WSF Evolution product in 2019 (Esch et al., 2018b). The WSF Evolution provides detailed information about the spatiotemporal development from 1985-2015 for each human settlement identified in the WSF 2015. The corresponding analysis is based on processing of multi-temporal mass data collections of the Landsat archive using Google Earth Engine.

- Masson et al. (2003) and Faroux et al. (2013) built a 1 km resolution global database suitable for vegetation and urban surface models for atmospheric modelling, with the associated database. The urban description is based on CORINE Land Cover over Europe (Bossard et al., 2000). An updated version, Ecoclimap-SG, has been built globally at a 300 m resolution from the ESA-CCI for natural covers and the Global Human Settlement Layer (see above) for urban areas. In Ecoclimap-SG, the GHSL classes have been translated into the 10 LCZ urban classes globally.
- For Europe, the satellite-based CORINE Land Cover (100 m resolution; Bossard et al., 2000) is constructed in a bottom-up approach by national teams coordinated by the European Environment Agency (EEA). It includes 9 land cover types related to the morphology and use of urban areas (e.g. ''Continuous urban fabric'', ''Industrial or commercial units'', ''airports'') and is available as yearly snapshots for 1990, 2000, 2006 and 2012.
- The Urban Atlas, compiled by the EEA (https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-urban-atlas), includes information on 17 urban land cover types for ~700 urban areas (version 2012) in the EU28 and EFTA countries. Cities larger than 100 000 inhabitants and their surroundings (more than 50 000 inhabitants) are covered. It has been constructed by combining image classification and visual interpretation of very high-resolution satellite imagery (SPOT5/6, Formosat), city maps and online map services. The land cover classes describe the density of the urban fabric and its use (e.g. “roads”, “airports”). In addition, the Urban Atlas provides data on building height and street trees for a sub-sample of cities.
- Jackson et al. (2010) define 33 regions in the world with different climates, socio-economic characteristics and architectural practices for four classes of urbanization, which are characterized by their morphology (low/medium/high density and tall building district), thermal and radiative properties.
- The World Urban Database and Access Portal Tools (WUDAPT; Ching et al., 2018) aims to construct a global database on urban form and function. Landsat satellite data and local expertise are used to create spatial maps of Local Climate Zones (LCZ; Stewart and Oke, 2012). The LCZs are strongly linked to urban morphology but to a lesser degree to construction materials, building use and energy consumption. In the follow-up to WUDAPT, there are plans to enrich the data set using crowdsourcing techniques and advanced high-resolution satellite data.

Acknowledgements

Coauthors Burmeister, Esch, Heldens, Kanani-Sühring, Maronga, Pavlik and Zeidler express their gratitude to the German Federal Ministry of Education and Research (BMBF) for funding grant 01LP1601 within the framework of Research for Sustainable Development (FONA; www.fona.de). Coauthor See would like to acknowledge the support of the FP7-funded ERC project CrowdLand (Grant n° 617754). Coauthors Masson, Bocher, de Munck, Lemonsu, Lévy, Schoetter, Tornay, Bonhomme thank the French National Agency of Research for their support through the project applied Modelling and urbAn Planning laws: Urban Climate and Energy (MApUCE) with reference ANR-13-VBDU-0004. Coauthors Masson, Bocher, de Munck, Lemonsu, Schoetter, Votsis and Bucher

express their gratitude to ERA4CS, an ERA-NET initiated by JPI Climate with co-funding from the European Union (Grant n° 690462) for the URCLIM project (www.urclim.eu).

References

1. Aguilar, M.A., del Mar Saldana, M., Aguilar, F.J., 2014. Generation and quality assessment of stereo-extracted DSM From GeoEye-1 and WorldView-2 imagery. IEEE Trans. Geosci. Remote Sens. 52, 1259–1271. https://doi.org/10.1109/TGRS.2013.2249521
2. Allen, L., Lindberg, F., Grimmond, C.S.B., 2011. Global to city scale urban anthropogenic heat flux: model and variability. Int. J. Climatol. 31, 1990–2005. https://doi.org/10.1002/joc.2210
3. Alonzo, M., Bookhagen, B., McFadden, J.P., Sun, A., Roberts, D.A., 2015. Mapping urban forest leaf area index with airborne lidar using penetration metrics and allometry. Remote Sens. Environ. 162, 141–153. https://doi.org/10.1016/j.rse.2015.02.025
4. Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J.-L., 2008. GlobCover the most detailed portrait of Earth. ESA Bull. 136, 25–31.
5. Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23, 1–26. https://doi.org/10.1002/joc.859
6. Ban, Y., Gong, P., Giri, C., 2015. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. ISPRS J. Photogramm. Remote Sens. 103, 1–6. https://doi.org/10.1016/j.isprsjprs.2015.01.001
7. Barnsley, M.J., Barr, S.L., 1997. Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system. Comput. Environ. Urban Syst., Remote Sensing of Urban Systems 21, 209–225. https://doi.org/10.1016/S0198-9715(97)10001-1
8. Bechtel, B., Alexander, P.J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L., Stewart, I., 2015. Mapping Local Climate Zones for a worldwide database of the form and function of cities. ISPRS Int. J. Geo-Inf. 4, 199–219. https://doi.org/10.3390/ijgi4010199
9. Berghauser Pont, M., Haupt, P., 2005. The Spacemate: Density and the typomorphology of the urban fabric. Nord. J. Archit. Res. 4, 55–68.
10. Biljecki, F., Ledoux, H., Stoter, J., 2017. Generating 3D city models without elevation data. Comput. Environ. Urban Syst. 64, 1–18. https://doi.org/10.1016/j.compenvurbsys.2017.01.001
11. Biljecki, F., Sindram, M., 2017. Estimating building age with 3D GIS, in: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the 12th 3D Geoinfo Conference 2017, Copernicus GmbH, pp. 17–24. https://doi.org/10.5194/isprs-annals-IV-4-W5-17-2017
12. Bocher, E., Petit, G., Bernard, J., Palominos, S., 2018. A geoprocessing framework to compute urban indicators: The MApUCE tools chain. Urban Clim. 24, 153–174. https://doi.org/10.1016/j.uclim.2018.01.008
13. Boeing, G., 2017. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. 65, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004
14. Bonczak, B., Kontokosta, C.E., 2019. Large-scale parameterization of 3D building morphology in complex urban landscapes using aerial LiDAR and city administrative data. Comput. Environ. Urban Syst. 73, 126–142. https://doi.org/10.1016/j.compenvurbsys.2018.09.004
15. Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., Zülkhe, M., Kalogirou, V., Seifert, F.M., Arino, O., 2013. Consistent global land cover maps for climate modelling communities: Current

achievements of the ESA LAND COVER CCI. Presented at the ESA Living Planet Symposium 2013, Edinburgh, UK.
16. Bossard, M., Feranec, J., Otahel, J., 2000. CORINE land cover technical guide – Addendum 2000. European Environment Agency, Technical report No 40. (Technical report No 40.). European Environment Agency.
17. Bourgeois, A., Pellegrino, M., Lévy, J.-P., 2017. Modeling and mapping domestic energy behavior: Insights from a consumer survey in France. Energy Res. Soc. Sci., Energy Consumption in Buildings: 32, 180–192. https://doi.org/10.1016/j.erss.2017.06.021
18. Bouyer, J., Inard, C., Musy, M., 2011. Microclimatic coupling as a solution to improve building energy simulation in an urban context. Energy Build. 43, 1549–1559. https://doi.org/10.1016/j.enbuild.2011.02.010
19. Brasebin, M., Perret, J., Mustière, S., Weber, C., 2012. Measuring the impact of 3D data geometric modeling on spatial analysis: Illustration with Skyview factor, in: Leduc, T., Moreau, G., Billen, R. (Eds.), Usage, Usability, and Utility of 3D City Models – European COST Action TU0801. Presented at the Usage, Usability, and Utility of 3D City Models – European COST Action TU0801, EDP Sciences, Nantes, France, p. 02001. https://doi.org/10.1051/3u3d/201202001
20. Breda, N.J.J., 2003. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. J. Exp. Bot. 54, 2403–2417. https://doi.org/10.1093/jxb/erg263
21. Burochin, J.-P., Tournaire, O., Paparoditis., N., 2009. An unsupervised hierarchical segmentation of a facade building image in elementary 2D-models. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 38, 223–228.
22. Burochin, J.-P., Vallet, B., Brédif, M., Mallet, C., Brosset, T., Paparoditis, N., 2014. Detecting blind building façades from highly overlapping wide angle aerial imagery. ISPRS J. Photogramm. Remote Sens. 96, 193–209. https://doi.org/10.1016/j.isprsjprs.2014.07.011
23. Burochin, J.-P., Vallet, B., Tournaire, O., Paparoditis., N., 2010. A formulation for unsupervised hierarchical segmentation of façade images with periodic models. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. IAPRS 38, 227–232.
24. Chen, Jun, Chen, Jin, Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27. https://doi.org/10.1016/j.isprsjprs.2014.09.002
25. Ching, J., Aliaga, D., Mills, G., Masson, V., See, L., Neophytou, M., Middel, A., Baklanov, A., Ren, C., Ng, E., Huang, Y., Stewart, I., Fung, J., Wong, M., Zhang, X., Shehata, A., Martilli, A., Miao, S., Wang, X., Yamagata, Y., Duarte, D., Schwander, L., Wang, W., Li, Y., Hanna, A., Feddema, J., Bechtel, B., Gonzales, J., Hidalgo, J., Roustan, Y., Kim, Y.S., Bornstein, R., Tzu-Ping, Brousse, O., 2019. Pathway using WUDAPT's Digital Synthetic City tool towards generating urban canopy parameters for multi-scale urban atmospheric modeling. Urban Clim. In review.
26. Ching, J., Brown, M., Burian, S., Chen, F., Cionco, R., Hanna, A., Hultgren, T., McPherson, T., Sailor, D., Taha, H., Williams, D., 2009. National Urban Database and Access Portal Tool. Bull. Am. Meteorol. Soc. 90, 1157–1168. https://doi.org/10.1175/2009BAMS2675.1
27. 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., 2018. WUDAPT: An urban weather, climate, and environmental modeling infrastructure for the Anthropocene. Bull. Am. Meteorol. Soc. 99, 1907–1924. https://doi.org/10.1175/BAMS-D-16-0236.1

28. CIESIN, 2017. Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 10 Data Sets. https://doi.org/10.7927/H4B56GPT
29. Congalton, R., Gu, J., Yadav, K., Thenkabail, P., Ozdogan, M., 2014. Global land cover mapping: A review and uncertainty analysis. Remote Sens. 6, 12070–12093. https://doi.org/10.3390/rs61212070
30. de Munck, C., Bernard, E., Lemonsu, A., Hidalgo, J., Touati, N., Bouyer, J., 2018. Impact of modelling vegetation at high resolution on urban climate variability. Presented at the 10th Inernational Conference on Urban Climate, New York, USA.
31. de Munck, C., Pigeon, G., Masson, V., Meunier, F., Bousquet, P., Tréméac, B., Merchat, M., Poeuf, P., Marchadier, C., 2013. How much can air conditioning increase air temperatures for a city like Paris, France? Int. J. Climatol. 33, 210–227. https://doi.org/10.1002/joc.3415
32. Dong, Y., Varquez, A.C.G., Kanda, M., 2017. Global anthropogenic heat flux database with high spatial resolution. Atmos. Environ. 150, 276–294. https://doi.org/10.1016/j.atmosenv.2016.11.040
33. Ebel, W., Eicke-Hennig, W., Feist, W., 1990. Energiesparpotential im Gebäudebestand. IWU, Darmstadt, Germany.
34. Ebert, A., Kerle, N., Stein, A., 2009. Urban social vulnerability assessment with physical proxies and spatial metrics derived from air- and spaceborne imagery and GIS data. Nat. Hazards 48, 275–294. https://doi.org/10.1007/s11069-008-9264-0
35. Eckert, S., Hollands, T., 2010. Comparison of automatic DSM generation modules by processing IKONOS stereo data of an urban area. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 162–167. https://doi.org/10.1109/JSTARS.2010.2047096
36. EEA, 2018. Street Tree Layer (STL). URL: https://land.copernicus.eu/local/urban-atlas/street-tree-layer-stl.
37. Esch, T., Asamer, H., Bachofer, F., Balhar, J., Boettcher, M., Boissier, E., d' Angelo, P., Gevaert, C.M., Hirner, A., Jupova, K., Kurz, F., Kwarteng, A.Y., Mathot, E., Marconcini, M., Marin, A., Metz-Marconcini, A., Pacini, F., Paganini, M., Permana, H., Soukup, T., Uereyen, S., Small, C., Svaton, V., Zeidler, J.N., 2018a. Digital world meets urban planet – new prospects for evidence-based urban studies arising from joint exploitation of big earth data, information technology and shared knowledge. Int. J. Digit. Earth 1–22. https://doi.org/10.1080/17538947.2018.1548655
38. Esch, T., Bachofer, F., Heldens, W., Hirner, A., Marconcini, M., Palacios-Lopez, D., Roth, A., Üreyen, S., Zeidler, J., Dech, S., Gorelick, N., 2018b. Where we live—A summary of the achievements and planned evolution of the Global Urban Footprint. Remote Sens. 10, 895. https://doi.org/10.3390/rs10060895
39. Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., Strano, E., 2017. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. ISPRS J. Photogramm. Remote Sens. 134, 30–42. https://doi.org/10.1016/j.isprsjprs.2017.10.012
40. Fan, H., Zipf, A., 2016. Modelling the world in 3D from VGI/crowdsourced data, in: Capineri, C., Haklay, M., Huang, H., Antoniou, V., Kettunen, J., Ostermann, F., Purves, R. (Eds.), European Handbook of Crowdsourced Geographic Information. Ubiquity Press, pp. 435–446. https://doi.org/10.5334/bax
41. Faroux, S., Kaptué Tchuenté, A.T., Roujean, J.-L., Masson, V., Martin, E., Le Moigne, P., 2013. ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. Geosci. Model Dev. 6, 563–582. https://doi.org/10.5194/gmd-6-563-2013
42. Flanner, M.G., 2009. Integrating anthropogenic heat flux with global climate models. Geophys. Res. Lett. 36. https://doi.org/10.1029/2008GL036465

43. Fonte, C.C., Minghini, M., Antoniou, V., Patriarca, J., See, L., 2018. Classification of building function using available sources of VGI. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII–4, 209–215. https://doi.org/10.5194/isprs-archives-XLII-4-209-2018
44. Fornaro, G., Serafino, F., 2006. Imaging of single and double scatterers in urban areas via SAR tomography. IEEE Trans. Geosci. Remote Sens. 44, 3497–3505. https://doi.org/10.1109/TGRS.2006.881748
45. Fouillet, A., Rey, G., Laurent, F., Pavillon, G., Bellec, S., Guihenneuc-Jouyaux, C., Clavel, J., Jougla, E., Hémon, D., 2006. Excess mortality related to the August 2003 heat wave in France. Int. Arch. Occup. Environ. Health 80, 16–24. https://doi.org/10.1007/s00420-006-0089-4
46. Fradkin, M., Roux, M., Maitre, H., Leloglu, U.M., 1999. Surface reconstruction from multiple aerial images in dense urban areas, in: Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149). Presented at the Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), pp. 262-267 Vol. 2. https://doi.org/10.1109/CVPR.1999.784639
47. Freitas, E.D., Rozoff, C.M., Cotton, W.R., Dias, P.L.S., 2007. Interactions of an urban heat island and sea-breeze circulations during winter over the metropolitan area of São Paulo, Brazil. Bound.-Layer Meteorol. 122, 43–65. https://doi.org/10.1007/s10546-006-9091-3
48. Frey, O., Hajnsek, I., Wegmuller, U., Werner, C.L., 2014. SAR tomography based 3-D point cloud extraction of point-like scatterers in urban areas, in: 2014 IEEE Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2014 - 2014 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Quebec City, QC, pp. 1313–1316. https://doi.org/10.1109/IGARSS.2014.6946675
49. Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182. https://doi.org/10.1016/j.rse.2009.08.016
50. Fuller, R.A., Gaston, K.J., 2009. The scaling of green space coverage in European cities. Biol. Lett. 5, 352–355. https://doi.org/10.1098/rsbl.2009.0010
51. Geiß, C., Wurm, M., Breunig, M., Felbier, A., Taubenböck, H., 2015. Normalization of TanDEM-X DSM data in urban environments with morphological filters. IEEE Trans. Geosci. Remote Sens. 53, 4348–4362. https://doi.org/10.1109/TGRS.2015.2396195
52. Geletič, J., Lehnert, M., 2016. GIS-based delineation of local climate zones: The case of medium-sized Central European cities. Morav. Geogr. Rep. 24, 2–12. https://doi.org/10.1515/mgr-2016-0012
53. Gevaert, C.M., Persello, C., Sliuzas, R., Vosselman, G., 2017. Informal settlement classification using point-cloud and image-based features from UAV data. ISPRS J. Photogramm. Remote Sens. 125, 225–236. https://doi.org/10.1016/j.isprsjprs.2017.01.017
54. Ghaffarian, S., Kerle, N., Filatova, T., 2018. A review of remote sensing-based proxies and data processing methods for urban disaster risk management. Presented at the EGU General Assembly 2018.
55. Giordano, S., Bris, A.L., Mallet, C., 2017. Fully automatic analysis of archival aerial images current status and challenges, in: 2017 Joint Urban Remote Sensing Event (JURSE). Presented at the 2017 Joint Urban Remote Sensing Event (JURSE), pp. 1–4. https://doi.org/10.1109/JURSE.2017.7924620
56. Gong Fang-Ying, Zhao-Cheng Zeng, Fan Zhang, Xiaojiang Li, Edward Ng, Leslie K. Norford. Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. Building and Environment 134 (2018) 155–167. doi.org/10.1016/j.buildenv.2018.02.042

57. Goodwin, N.R., Coops, N.C., Tooke, T.R., Christen, A., Voogt, J.A., 2009. Characterizing urban surface cover and structure with airborne lidar technology. Can. J. Remote Sens. 35, 297–309. https://doi.org/10.5589/m09-015
58. Grekousis, G., Mountrakis, G., Kavouras, M., 2015. An overview of 21 global and 43 regional land-cover mapping products. Int. J. Remote Sens. 36, 5309–5335. https://doi.org/10.1080/01431161.2015.1093195
59. Grimmond, C.S.B., Blackett, M., Best, M.J., Baik, J.-J., Belcher, S.E., Beringer, J., Bohnenstengel, S.I., Calmet, I., Chen, F., Coutts, A., Dandou, A., Fortuniak, K., Gouvea, M.L., Hamdi, R., Hendry, M., Kanda, M., Kawai, T., Kawamoto, Y., Kondo, H., Krayenhoff, E.S., Lee, S.-H., Loridan, T., Martilli, A., Masson, V., Miao, S., Oleson, K., Ooka, R., Pigeon, G., Porson, A., Ryu, Y.-H., Salamanca, F., Steeneveld, G.J., Tombrou, M., Voogt, J.A., Young, D.T., Zhang, N., 2011. Initial results from Phase 2 of the international urban energy balance model comparison. Int. J. Climatol. 31, 244–272. https://doi.org/10.1002/joc.2227
60. Grimmond, C.S.B., Blackett, M., Best, M.J., Barlow, J., Baik, J.-J., Belcher, S.E., Bohnenstengel, S.I., Calmet, I., Chen, F., Dandou, A., Fortuniak, K., Gouvea, M.L., Hamdi, R., Hendry, M., Kawai, T., Kawamoto, Y., Kondo, H., Krayenhoff, E.S., Lee, S.-H., Loridan, T., Martilli, A., Masson, V., Miao, S., Oleson, K., Pigeon, G., Porson, A., Ryu, Y.-H., Salamanca, F., Shashua-Bar, L., Steeneveld, G.-J., Tombrou, M., Voogt, J., Young, D., Zhang, N., 2010. The international urban energy balance models comparison project: First results from phase 1. J. Appl. Meteorol. Climatol. 49, 1268–1292. https://doi.org/10.1175/2010JAMC2354.1
61. Grimmond, C.S.B., Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. J. Appl. Meteorol. 38, 1262–1292. https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2
62. Gröger, G., Plümer, L., 2012. CityGML – Interoperable semantic 3D city models. ISPRS J. Photogramm. Remote Sens. 71, 12–33. https://doi.org/10.1016/j.isprsjprs.2012.04.004
63. Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environ. Plan. B Plan. Des. 37, 682–703. https://doi.org/10.1068/b35097
64. Hamdi, R., Masson, V., 2008. Inclusion of a drag approach in the Town Energy Balance (TEB) scheme: offline 1D evaluation in a street canyon. J. Appl. Meteorol. Climatol. 47, 2627–2644. https://doi.org/10.1175/2008JAMC1865.1
65. Heiden, U., Segl, K., Roessner, S., Kaufmann, H., 2007. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. Remote Sens. Environ. 111, 537–552. https://doi.org/10.1016/j.rse.2007.04.008
66. Heiple, S., Sailor, D.J., 2008. Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles. Energy Build. 40, 1426–1436. https://doi.org/10.1016/j.enbuild.2008.01.005
67. Heldens, W., Heiden, U., Esch, T., Mueller, A., Dech, S., 2017. Integration of remote sensing based surface information into a three-dimensional microclimate model. ISPRS J. Photogramm. Remote Sens. 125, 106–124. https://doi.org/10.1016/j.isprsjprs.2017.01.009
68. Herold, M., Liu, X., Clarke, K.C., 2003. Spatial metrics and image texture for mapping urban land use. Photogramm. Eng. Remote Sens. 69, 991–1001.
69. Hidalgo, J., Dumas, G., Masson, V., Petit, G., Betchtel, B., Bocher, E., Foley, M., Schoetter, R., Mills, G., 2019. Comparison between Local Climate Zones maps derived from administrative datasets and satellite observations. Urban Clim. In press.
70. Hidalgo, J., Lemonsu, A., Masson, V., 2018. Between progress and obstacles in urban climate interdisciplinary studies and knowledge transfer to society: Urban climate interdisciplinary studies. Ann. N. Y. Acad. Sci. https://doi.org/10.1111/nyas.13986
71. Hirschmüller, H., Scholten, F., Hirzinger, G., 2005. Stereo vision based reconstruction of huge urban areas from an airborne pushbroom camera (HRSC), in: Kropatsch, W.G., Sablatnig, R.,

Hanbury, A. (Eds.), Pattern Recognition, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 58–66.
72. Hogan, R.J., 2018. An exponential model of urban geometry for use in radiative transfer applications. Bound.-Layer Meteorol. https://doi.org/10.1007/s10546-018-0409-8
73. Hosoi, F., Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar. IEEE Trans. Geosci. Remote Sens. 44, 3610–3618. https://doi.org/10.1109/TGRS.2006.881743
74. Houet, T., Pigeon, G., 2011. Mapping urban climate zones and quantifying climate behaviors--an application on Toulouse urban area (France). Environ. Pollut. Barking Essex 1987 159, 2180–2192. https://doi.org/10.1016/j.envpol.2010.12.027
75. IGN, 2018. BD-TOPO. Version 2.1. URL: http://professionnels.ign.fr/bdtopo/.
76. Iovan, C., Boldo, D., Cord, M., 2008. Detection, characterization, and modeling vegetation in urban areas from high-resolution aerial imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 1, 206–213. https://doi.org/10.1109/JSTARS.2008.2007514
77. Jackson, T.L., Feddema, J.J., Oleson, K.W., Bonan, G.B., Bauer, J.T., 2010. Parameterization of urban characteristics for global climate modeling. Ann. Assoc. Am. Geogr. 100, 848–865. https://doi.org/10.1080/00045608.2010.497328
78. Jakobsson, A., 2012. European Location Framework White Paper v1.0. URL: http://www.elfproject.eu/sites/default/files/ELF%20White%20Paper.pdf.
79. Jantz, C.A., Goetz, S.J., Shelley, M.K., 2004. Using the Sleuth urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington Metropolitan Area. Environ. Plan. B Plan. Des. 31, 251–271. https://doi.org/10.1068/b2983
80. Jensen, R.R., Hardin, P.J., 2005. Estimating urban leaf area index using field measurements and satellite remote sensing data. J. Arboric. 31, 21–27.
81. Joe, P., Belair, S., Bernier, N. b., Bouchet, V., Brook, J.R., Brunet, D., Burrows, W., Charland, J.-P., Dehghan, A., Driedger, N., Duhaime, C., Evans, G., Filion, A.-B., Frenette, R., de Grandpré, J., Gultepe, I., Henderson, D., Herdt, A., Hilker, N., Huang, L., Hung, E., Isaac, G., Jeong, C.-H., Johnston, D., Klaassen, J., Leroyer, S., Lin, H., MacDonald, M., MacPhee, J., Mariani, Z., Munoz, T., Reid, J., Robichaud, A., Rochon, Y., Shairsingh, K., Sills, D., Spacek, L., Stroud, C., Su, Y., Taylor, N., Vanos, J., Voogt, J., Wang, J.M., Wiechers, T., Wren, S., Yang, H., Yip, T., 2017. The Environment Canada Pan and Parapan American Science Showcase Project. Bull. Am. Meteorol. Soc. 99, 921–953. https://doi.org/10.1175/BAMS-D-16-0162.1
82. Jutzi, B., Stilla, U., 2003. Laser pulse analysis for reconstruction and classification of urban objects. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 34, 151–156.
83. Kastner-Klein, P., Fedorovich, E., Rotach, M.W., 2001: A wind tunnel study of organised and turbulent air motions in urban street canyons, Journal of Wind Engineering and Industrial Aerodynamics, 89 (9), 849-861, https://doi.org/10.1016/S0167-6105(01)00074-5.
84. Khalifa A., M. Marchetti, L. Bouilloud, E. Martin, M. Bues, and K. Chancibaut, 2016: Accounting for anthropic energy flux of traffic in winter urban road surface temperature simulations with TEB model, Geosci. Model Dev., 9, 547–565, doi:10.5194/gmd-9-547-2016
85. Kikegawa, Y., Genchi, Y., Yoshikado, H., Kondo, H., 2003. Development of a numerical simulation system toward comprehensive assessments of urban warming countermeasures including their impacts upon the urban buildings' energy-demands. Appl. Energy 76, 449–466. https://doi.org/10.1016/S0306-2619(03)00009-6
86. Kohler, M., Blond, N., Clappier, A., 2016. A city scale degree-day method to assess building space heating energy demands in Strasbourg Eurometropolis (France). Appl. Energy 184, 40–54. https://doi.org/10.1016/j.apenergy.2016.09.075
87. Kotharkar, R., Bagade, A., 2018. Local Climate Zone classification for Indian cities: A case study of Nagpur. Urban Clim. 24, 369–392. https://doi.org/10.1016/j.uclim.2017.03.003

88. Krayenhoff, E.S., Santiago, J.-L., Martilli, A., Christen, A., Oke, T.R., 2015. Parametrization of drag and turbulence for urban neighbourhoods with trees. Bound.-Layer Meteorol. 156, 157–189. https://doi.org/10.1007/s10546-015-0028-6
89. Kunze, C., Hecht, R., 2015. Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population. Comput. Environ. Urban Syst., Special Issue on Volunteered Geographic Information 53, 4–18. https://doi.org/10.1016/j.compenvurbsys.2015.04.002
90. Kusaka, H., Kondo, H., Kikegawa, Y., Kimura, F., 2001. A simple single-layer urban canopy model for atmospheric models: Comparison With multi-layer and slab models. Bound.-Layer Meteorol. 101, 329–358. https://doi.org/10.1023/A:1019207923078
91. Lalic, B., Mihailovic, D.T., 2004. An empirical relation describing leaf-area density inside the forest for environmental modeling. J. Appl. Meteorol. 43, 641–645. https://doi.org/10.1175/1520-0450(2004)043<0641:AERDLD>2.0.CO;2
92. Lao, J., Bocher, E., Petit, G., Palominos, S., Le Saux, E., Masson, V., 2018. Is OpenStreetMap suitable for urban climate studies? Presented at the OGRS2018 (Open Source Geospatial Research & Education Symposium 2018), Lugano, Switzerland.
93. Larosa, F., Perrels, A., 2018. Assessment of the Existing Resourcing and Quality Assurance ofCurrent Climate Services, H2020 project EU-MACS Deliverable 1.2. URL: http://eu-macs.eu/wp-content/uploads/2017/07/EUMACS_D12_revision.pdf.
94. Leconte, F., Bouyer, J., Claverie, R., Pétrissans, M., 2017. Analysis of nocturnal air temperature in districts using mobile measurements and a cooling indicator. Theor. Appl. Climatol. 130, 365–376. https://doi.org/10.1007/s00704-016-1886-7
95. Lee, S.-H., Park, S.-U., 2008. A vegetated urban canopy model for meteorological and environmental modelling. Bound.-Layer Meteorol. 126, 73–102. https://doi.org/10.1007/s10546-007-9221-6
96. Lefebvre, A., Corpetti, T., Nabucet, J., Hubert-Moy, L., 2017. Urban vegetation extraction with multi-angular Pléiades images, in: 2017 Joint Urban Remote Sensing Event (JURSE). Presented at the 2017 Joint Urban Remote Sensing Event (JURSE), pp. 1–4. https://doi.org/10.1109/JURSE.2017.7924624
97. Lemonsu, A., Masson, V., Shashua-Bar, L., Erell, E., Pearlmutter, D., 2012. Inclusion of vegetation in the Town Energy Balance model for modelling urban green areas. Geosci. Model Dev. 5, 1377–1393. https://doi.org/10.5194/gmd-5-1377-2012
98. Li, M., Beurs, K.M. de, Stein, A., Bijker, W., 2017. Incorporating open source data for Bayesian classification of urban land use from VHR stereo images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4930–4943. https://doi.org/10.1109/JSTARS.2017.2737702
99. Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For. Urban Green. 14, 675–685. https://doi.org/10.1016/j.ufug.2015.06.006
100. Liang, J., Gong, J., Sun, J., Zhou, J., Li, W., Li, Y., Liu, J., Shen, S., 2017. Automatic sky view factor estimation from Street View photographs—A big data approach. Remote Sens. 9, 411. https://doi.org/10.3390/rs9050411
101. Lindberg, F., 2007. Modelling the urban climate using a local governmental geo-database. Meteorol. Appl. 14, 263–273. https://doi.org/10.1002/met.29
102. Loga, T., Stein, B., Diefenbach, N., 2016. TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable. Energy Build., Towards an energy efficient European housing stock: monitoring, mapping and modelling retrofitting processes 132, 4–12. https://doi.org/10.1016/j.enbuild.2016.06.094
103. Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28, 823–870. https://doi.org/10.1080/01431160600746456

104. Lu, D., Weng, Q., 2006. Use of impervious surface in urban land-use classification. Remote Sens. Environ. 102, 146–160. https://doi.org/10.1016/j.rse.2006.02.010
105. Lu, Y., Coops, N.C., Hermosilla, T., 2017. Estimating urban vegetation fraction across 25 cities in pan-Pacific using Landsat time series data. ISPRS J. Photogramm. Remote Sens. 126, 11–23. https://doi.org/10.1016/j.isprsjprs.2016.12.014
106. Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y., 2017. A review of supervised object-based land-cover image classification. ISPRS J. Photogramm. Remote Sens. 130, 277–293. https://doi.org/10.1016/j.isprsjprs.2017.06.001
107. Marconcini, M., Marmanis, D., Esch, T., Felbier, A., 2014. A novel method for building height estimation using TanDEM-X Data - The case of Dongying, China. Presented at the IGARSS 2014 / 35th CSRS, Quebec, Canada.
108. Markkanen, T., Rannik, Ü., Marcolla, B., Cescatti, A., Vesala, T., 2003. Footprints and fetches for fluxes over forest canopies with varying structure and density. Bound.-Layer Meteorol. 106, 437–459. https://doi.org/10.1023/A:1021261606719
109. Maronga, B., Groß, G., Raasch, S., Banzhaf, S., Forkel, R., Heldens, W., Matzarakis, A., Mauder, M., Pavlik, D., Pfafferott, J., Schubert, S., Seckmeyer, G., Sieker, H., Trusilova, K., 2019a. Development of a new urban climate model based on the model PALM - Project overview, planned work, and first achievements. Meteorol. Z., 28(2)105 - 119. https://doi.org/10.1127/metz/2019/0909.
110. Maronga, B., Gryschka, M., Heinze, R., Hoffmann, F., Kanani-Sühring, F., Keck, M., Ketelsen, K., Letzel, M.O., Sühring, M., Raasch, S., 2015. The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. Geosci. Model Dev. 8, 2515–2551. https://doi.org/10.5194/gmd-8-2515-2015
111. Maronga, B., Banzhaf, S., Burmeister, C., Esch, T., Forkel, R., Fröhlich, D., Fuka, V., Gehrke, K. F., Geletič, J., Giersch, S., Gronemeier, T., Groß, G., Heldens, W., Hellsten, A., Hoffmann, F., Inagaki, A., Kadasch, E., Kanani-Sühring, F., Ketelsen, K., Khan, B. A., Knigge, C., Knoop, H., Krč, P., Kurppa, M., Maamari, H., Matzarakis, A., Mauder, M., Pallasch, M., Pavlik, D., Pfafferott, J., Resler, J., Rissmann, S., Russo, E., Salim, M., Schrempf, M., Schwenkel, J., Seckmeyer, G., Schubert, S., Sühring, M., von Tils, R., Vollmer, L., Ward, S., Witha, B., Wurps, H., Zeidler, J., and Raasch, S., 2019b: Overview of the PALM model system 6.0, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-103
112. Martilli, A., Clappier, A., Rotach, M.W., 2002. An urban surface exchange parameterisation for mesoscale models. Bound.-Layer Meteorol. 104, 261–304. https://doi.org/10.1023/A:1016099921195
113. Masson, V., 2006. Urban surface modeling and the meso-scale impact of cities. Theor. Appl. Climatol. 84, 35–45. https://doi.org/10.1007/s00704-005-0142-3
114. Masson, V., 2000. A physically-based scheme for the urban energy budget in atmospheric models. Bound.-Layer Meteorol. 94, 357–397. https://doi.org/10.1023/A:1002463829265
115. Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C., Lacaze, R., 2003. A global database of land surface parameters at 1-km resolution in meteorological and climate models. J. Clim. 16, 1261–1282. https://doi.org/10.1175/1520-0442-16.9.1261
116. Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: an applied review. Int. J. Remote Sens. 39, 2784–2817. https://doi.org/10.1080/01431161.2018.1433343
117. Mhedhbi, Z., Masson V., Haoues-Jouve, S., Hidalgo, J., 2019, Collection of refined architectural parameters by crowdsourcing using Facebook social network: case of Greater Tunis, Urban Climate, 29, Article 100468, doi:10.1016/j.uclim.2019.100499

118. Middel A., J. Lukasczykb, S. Zakrzewskib, M. Arnoldc, R. Maciejewskid 2019, Urban form and composition of street canyons: A human-centric big data and deep learning approach, Landscape and Urban Planning 183, 122–132
119. Middel A., J. Lukasczykb, R. Maciejewskic, M. Demuzered, M. Rothe, 2018, Sky View Factor footprints for urban climate modeling, Urban Climate 25, 120–13
120. MIT Senseable City Lab, 2018. Treepedia. URL: http://senseable.mit.edu/treepedia.
121. Mooney, P., Minghini, M., 2017. A review of OpenStreetMap data, in: Foody, G.M., See, L., Fritz, S., Fonte, C.C., Mooney, P., Olteanu-Raimond, A.-M., Antoniou, V. (Eds.), Mapping and the Citizen Sensor. Ubiquity Press, London, UK, pp. 37–59.
122. Noilhan, J., 1981. A model for the net total radiation flux at the surfaces of a building. Build. Environ. 16, 259–266. https://doi.org/10.1016/0360-1323(81)90004-4
123. Ohashi, Y., Genchi, Y., Kondo, H., Kikegawa, Y., Yoshikado, H., Hirano, Y., 2007. Influence of air-conditioning waste heat on air temperature in Tokyo during summer: Numerical experiments using an urban canopy model coupled with a building energy model. J. Appl. Meteorol. Climatol. 46, 66–81. https://doi.org/10.1175/JAM2441.1
124. Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/9781139016476
125. Olbricht, R., 2015. Data retrieval for small spatial regions in OpenStreetMap, in: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, Cham, pp. 101–122.
126. Oleson, K.W., Bonan, G.B., Feddema, J., Jackson, T., 2011. An examination of urban heat island characteristics in a global climate model. Int. J. Climatol. 31, 1848–1865. https://doi.org/10.1002/joc.2201
127. Oleson, K.W., Bonan, G.B., Feddema, J., Vertenstein, M., 2008. An Urban Parameterization for a Global Climate Model. Part II: Sensitivity to Input Parameters and the Simulated Urban Heat Island in Offline Simulations. J. Appl. Meteorol. Climatol. 47, 1061–1076. https://doi.org/10.1175/2007JAMC1598.1
128. Olteanu-Raimond, A.-M., Jolivet, L., Van Damme, M.-D., Royer, T., Fraval, L., See, L., Sturn, T., Karner, M., Moorthy, I., Fritz, S., 2018. An experimental framework for integrating citizen and community science into land cover, land use, and land change detection processes in a National Mapping Agency. Land 7, 103. https://doi.org/10.3390/land7030103
129. Ouerghemmi, W., Gadal, S., Mozgeris, G., 2018. Urban vegetation mapping using hyperspectral imagery and spectral library, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Valencia, Spain, pp. 1632–1635. https://doi.org/10.1109/IGARSS.2018.8518893
130. Over, M., Schilling, A., Neubauer, S., Zipf, A., 2010. Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. Comput. Environ. Urban Syst. 34, 496–507. https://doi.org/10.1016/j.compenvurbsys.2010.05.001
131. Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Carneiro Freire Sergio, M., Halkia, S., Julea Andreea, M., Kemper, T., Soille, P., Syrris, V., 2016. Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014 (EUR - Scientific and Technical Research Reports). Publications Office of the European Union. https://doi.org/10.2788/253582 (online)
132. Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L., Marin-Herrera, M.A., Ouzounis, G.K., Scavazzon, M., Soille, P., Syrris, V., Zanchetta, L., 2013. A Global Human Settlement Layer from optical HR/VHR RS data: Concept and first results. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 2102–2131. https://doi.org/10.1109/JSTARS.2013.2271445

133. Pigeon, G., Legain, D., Durand, P., Masson, V., 2007. Anthropogenic heat release in an old European agglomeration (Toulouse, France). Int. J. Climatol. 27, 1969–1981. https://doi.org/10.1002/joc.1530
134. Pigeon, G., Zibouche, K., Bueno, B., Le Bras, J., Masson, V., 2014. Improving the capabilities of the Town Energy Balance model with up-to-date building energy simulation algorithms: an application to a set of representative buildings in Paris. Energy Build. 76, 1–14. https://doi.org/10.1016/j.enbuild.2013.10.038
135. RAGE, 2012. Analyse détaillée du parc résidentiel existant. Report in French. URL: http://www.reglesdelart-grenelle-environnement-2012.fr/regles-de-lart/detail/rapport-2012-analyse-detaillee-du-parc-residentiel-existant.html.
136. Reinartz, P., D'Angelo, P., Krauss, T., Poli, D., Jacobsen, K., Buyuksalih, G., 2017. Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data. https://doi.org/10.15488/1121
137. Resler, J., Krč, P., Belda, M., Juruš, P., Benešová, N., Lopata, J., Vlček, O., Damašková, D., Eben, K., Derbek, P., Maronga, B., Kanani-Sühring, F., 2017. PALM-USM v1.0: A new urban surface model integrated into the PALM large-eddy simulation model. Geosci. Model Dev. 10, 3635–3659. https://doi.org/10.5194/gmd-10-3635-2017
138. Ridd, M.K., 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities†. Int. J. Remote Sens. 16, 2165–2185. https://doi.org/10.1080/01431169508954549
139. Rosser, J.F., Boyd, D.S., Long, G., Zakhary, S., Mao, Y., Robinson, D., 2019. Predicting residential building age from map data. Comput. Environ. Urban Syst. 73, 56–67. https://doi.org/10.1016/j.compenvurbsys.2018.08.004
140. Sailor, D.J., Georgescu, M., Milne, J.M., Hart, M.A., 2015. Development of a national anthropogenic heating database with an extrapolation for international cities. Atmospheric Environment, 118, 7-18.
141. Sailor, D.J., 2011. A review of methods for estimating anthropogenic heat and moisture emissions in the urban environment. Int. J. Climatol. 31, 189–199. https://doi.org/10.1002/joc.2106
142. Salamanca, F., Krpo, A., Martilli, A., Clappier, A., 2009. A new building energy model coupled with an urban canopy parameterization for urban climate simulations—part I. formulation, verification, and sensitivity analysis of the model. Theor. Appl. Climatol. 99, 331. https://doi.org/10.1007/s00704-009-0142-9
143. Salim, M.H., Schlünzen, K.H., Grawe, D., Boettcher, M., Gierisch, A.M.U., Fock, B.H., 2018. The microscale obstacle-resolving meteorological model MITRAS v2.0: model theory. Geosci. Model Dev. 11, 3427–3445. https://doi.org/10.5194/gmd-11-3427-2018
144. Salmond, J.A., Williams, D.E., Laing, G., Kingham, S., Dirks, K., Longley, I., Henshaw, G.S., 2013. The influence of vegetation on the horizontal and vertical distribution of pollutants in a street canyon. Sci. Total Environ. 443, 287–298. https://doi.org/10.1016/j.scitotenv.2012.10.101
145. Samsonov, T.E., Konstantinov, P.I., Varentsov, M.I., 2015. Object-oriented approach to urban canyon analysis and its applications in meteorological modeling. Urban Clim. 13, 122–139. https://doi.org/10.1016/j.uclim.2015.07.007
146. Santamouris, M., 2014. Cooling the cities – A review of reflective and green roof mitigation technologies to fight heat island and improve comfort in urban environments. Sol. Energy 103, 682–703. https://doi.org/10.1016/j.solener.2012.07.003
147. Santamouris, M., Cartalis, C., Synnefa, A., Kolokotsa, D., 2015. On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—A review. Energy Build., Renewable Energy Sources and Healthy Buildings 98, 119–124. https://doi.org/10.1016/j.enbuild.2014.09.052

148. Santiago, J.L., Martilli, A., 2010. A dynamic urban canopy parameterization for mesoscale models based on computational fluid dynamics Reynolds-average Navier–Stokes microscale simulations. Bound.-Layer Meteorol. 137, 417–439. https://doi.org/10.1007/s10546-010-9538-4
149. Santiago, J.L., Martilli, A.,F. Martin, 2017: On Dry Deposition Modelling of Atmospheric Pollutants on Vegetation at the Microscale: Application to the Impact of Street Vegetation on Air Quality, Bound.-Layer Meteorol., 162(3), 451-474
150. 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 the Town Energy Balance (SURFEX-TEB v. 8.2). Geosci. Model Dev. 10, 2801–2831. https://doi.org/10.5194/gmd-10-2801-2017
151. Schubert, S., Grossman-Clarke, S., Martilli, A., 2012. A double-canyon radiation scheme for multi-layer urban canopy models. Bound.-Layer Meteorol. 145, 439–468. https://doi.org/10.1007/s10546-012-9728-3
152. Seiferling, I., Naik, N., Ratti, C., Proulx, R., 2017. Green streets − Quantifying and mapping urban trees with street-level imagery and computer vision. Landsc. Urban Plan. 165, 93–101. https://doi.org/10.1016/j.landurbplan.2017.05.010
153. Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., Masson, V., 2011. The AROME-France convective-scale operational model. Mon. Weather Rev. 139, 976–991. https://doi.org/10.1175/2010MWR3425.1
154. Shashua-Bar, L., Hoffman, M.E., 2000. Vegetation as a climatic component in the design of an urban street: An empirical model for predicting the cooling effect of urban green areas with trees. Energy Build. 31, 221–235. https://doi.org/10.1016/S0378-7788(99)00018-3
155. Shelton, B., Karakiewicz, J., Kvan, T., 2011. The Making of Hong Kong: From Vertical to Volumetric, 1 edition. ed. Routledge, London; New York.
156. Shepherd, J.M., 2005. A review of current investigations of urban-induced rainfall and recommendations for the future. Earth Interact. 9, 1–27. https://doi.org/10.1175/EI156.1
157. Sirmacek, B., Taubenbock, H., Reinartz, P., Ehlers, M., 2012. Performance evaluation for 3-D city model generation of six different DSMs from air- and spaceborne sensors. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5, 59–70. https://doi.org/10.1109/JSTARS.2011.2178399
158. Spyratos, S., Stathakis, D., Lutz, M., Tsinaraki, C., 2017. Using Foursquare place data for estimating building block use. Environ. Plan. B Urban Anal. City Sci. 44, 693–717. https://doi.org/10.1177/0265813516637607
159. Steiniger, S., 2006. Classifying urban structures for mapping purposes using discriminant analysis, in: Steiniger, Stefan (2006). Classifying Urban Structures for Mapping Purposes Using Discriminant Analysis. In: GISRUK 2006, Nottingham (UK), 5 April 2006 - 7 April 2006, Online. Presented at the GISRUK 2006, Nottingham, UK, p. online. https://doi.org/info:doi/10.5167/uzh-77840
160. Stewart, I.D., Oke, T.R., 2012. Local Climate Zones for Urban Temperature Studies. Bull. Am. Meteorol. Soc. 93, 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1
161. Stoter, J., Roensdorf, C., Home, R., Capstick, D., Streilein, A., Kellenberger, T., Bayers, E., Kane, P., Dorsch, J., Woźniak, P., Lysell, G., Lithen, T., Bucher, B., Paparoditis, N., Ilves, R., 2015. 3D modelling with national coverage: Bridging the gap between research and practice, in: Breunig, M., Al-Doori, M., Butwilowski, E., Kuper, P.V., Benner, J., Haefele, K.H. (Eds.), 3D Geoinformation Science: The Selected Papers of the 3D GeoInfo 2014, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, Cham, pp. 207–225. https://doi.org/10.1007/978-3-319-12181-9_13

162. Tan, J., Zheng, Y., Tang, X., Guo, C., Li, L., Song, G., Zhen, X., Yuan, D., Kalkstein, A.J., Li, F., 2010. The urban heat island and its impact on heat waves and human health in Shanghai. Int. J. Biometeorol. 54, 75–84. https://doi.org/10.1007/s00484-009-0256-x
163. Thomas, N., Hendrix, C., Congalton, R.G., 2003. A comparison of urban mapping methods using high-resolution digital imagery. Photogramm. Eng. Remote Sens. 69, 963–972. https://doi.org/10.14358/PERS.69.9.963
164. Tornay, N., Schoetter, R., Bonhomme, M., Faraut, S., Masson, V., 2017. GENIUS: A methodology to define a detailed description of buildings for urban climate and building energy consumption simulations. Urban Clim. 20, 75–93. https://doi.org/10.1016/j.uclim.2017.03.002
165. Toutin, T., 2006. Generation of DSMs from SPOT-5 in-track HRS and across-track HRG stereo data using spatiotriangulation and autocalibration. ISPRS J. Photogramm. Remote Sens. 60, 170–181. https://doi.org/10.1016/j.isprsjprs.2006.02.003
166. von Glasow, R., Jickells, T.D., Baklanov, A., Carmichael, G.R., Church, T.M., Gallardo, L., Hughes, C., Kanakidou, M., Liss, P.S., Mee, L., Raine, R., Ramachandran, P., Ramesh, R., Sundseth, K., Tsunogai, U., Uematsu, M., Zhu, T., 2013. Megacities and large urban agglomerations in the coastal zone: interactions between atmosphere, land, and marine ecosystems. Ambio 42, 13–28. https://doi.org/10.1007/s13280-012-0343-9
167. Vosselman, G., Maas, H.-G., 2010. Airborne and terrestrial laser scanning. Whittles ; Distributed in North America by CRC, Dunbeath, Scotland; Boca Raton, FL.
168. Votsis, A., 2017. Utilizing a cellular automaton model to explore the influence of coastal flood adaptation strategies on Helsinki's urbanization patterns. Comput. Environ. Urban Syst. 64, 344–355. https://doi.org/10.1016/j.compenvurbsys.2017.04.005
169. Wang, J., Qingming, Z., Guo, H., Jin, Z., 2016. Characterizing the spatial dynamics of land surface temperature–impervious surface fraction relationship. Int. J. Appl. Earth Obs. Geoinformation 45, 55–65. https://doi.org/10.1016/j.jag.2015.11.006
170. Wang, Y., Li, Y., Sabatino, S.D., Martilli, A., Chan, P.W., 2018. Effects of anthropogenic heat due to air-conditioning systems on an extreme high temperature event in Hong Kong. Environ. Res. Lett. 13, 034015. https://doi.org/10.1088/1748-9326/aaa848
171. Weidner, U., Förstner, W., 1995. Towards automatic building extraction from high-resolution digital elevation models. ISPRS J. Photogramm. Remote Sens. 50, 38–49. https://doi.org/10.1016/0924-2716(95)98236-S
172. Wong, W.S., 2014. Architectural phenomena following law—Review of residential buildings in Hong Kong's colonial era. J. Civ. Eng. Archit. Res. 1, 215–229.
173. Wouters, H., Demuzere, M., Blahak, U., Fortuniak, K., Maiheu, B., Camps, J., Tielemans, D., van Lipzig, N.P.M., 2016. The efficient urban canopy dependency parametrization (SURY) v1.0 for atmospheric modelling: description and application with the COSMO-CLM model for a Belgian summer. Geosci. Model Dev. 9, 3027–3054. https://doi.org/10.5194/gmd-9-3027-2016
174. Yan, W.Y., Shaker, A., El-Ashmawy, N., 2015. Urban land cover classification using airborne LiDAR data: A review. Remote Sens. Environ. 158, 295–310. https://doi.org/10.1016/j.rse.2014.11.001
175. Yang, J., He, Y., 2017. Automated mapping of impervious surfaces in urban and suburban areas: Linear spectral unmixing of high spatial resolution imagery. Int. J. Appl. Earth Obs. Geoinformation 54, 53–64. https://doi.org/10.1016/j.jag.2016.09.006
176. Zhang, J., Gao, S., Chen, H., Yu, J., Tang, Q., 2015. Retrieval of the land surface-air temperature difference from high spatial resolution satellite observations over complex surfaces in the Tibetan Plateau. J. Geophys. Res. Atmospheres 120, 8065–8079. https://doi.org/10.1002/2015JD023395
177. Zhang, X., Shehata, A., Benes, B., Aliaga, D., 2019. Inferring 3D Urban Models, submitted for publication, 12 pages.

178. Zheng, Y., Ren, C., Xu, Y., Wang, R., Ho, J., Lau, K., Ng, E., 2018. GIS-based mapping of Local Climate Zone in the high-density city of Hong Kong. Urban Clim. 24, 419–448. https://doi.org/10.1016/j.uclim.2017.05.008
179. Zhu, X.X., Montazeri, S., Gisinger, C., Hanssen, R.F., Bamler, R., 2016. Geodetic SAR tomography. IEEE Trans. Geosci. Remote Sens. 54, 18–35. https://doi.org/10.1109/TGRS.2015.2448686