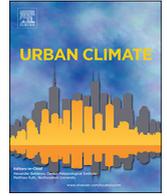




Contents lists available at ScienceDirect

## Urban Climate

journal homepage: [www.elsevier.com/locate/uclim](http://www.elsevier.com/locate/uclim)

## Sky View Factor footprints for urban climate modeling

Ariane Middel<sup>a,f,\*</sup>, Jonas Lukasczyk<sup>b</sup>, Ross Maciejewski<sup>c</sup>, Matthias Demuzere<sup>d</sup>, Matthias Roth<sup>e</sup><sup>a</sup> Department of Geography and Urban Studies, Temple University, 324 Gladfelter Hall, Philadelphia, PA 19122, USA<sup>b</sup> Department of Computer Science, Technische Universität Kaiserslautern, PO Box 3049, Kaiserslautern, Germany<sup>c</sup> School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, 699 S Mill Ave, Tempe, AZ 85281, USA<sup>d</sup> Laboratory of Hydrology and Water Management, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium<sup>e</sup> Department of Geography, National University of Singapore, 1 Arts Link, Kent Ridge, Singapore 117570<sup>f</sup> Urban Climate Research Center, Arizona State University, 975 S Myrtle Ave, Tempe, AZ 85281, USA

## 1. Introduction

Urban morphology, i.e. the physical form of cities, is a significant driver of urban climate at the micro- and local scale (Stewart and Oke, 2012; Middel et al., 2014). Resolving the spatial heterogeneity of cities at fine scales is of paramount interest to the weather and climate modeling community, as urban morphology impacts air temperature (Oke, 1981; Unger, 2004; Unger, 2009; Chen et al., 2012), air flow, ventilation, turbulence (Grimmond and Oke, 1999; Martilli et al., 2002; Ng et al., 2011), solar radiation (Johansson, 2006; Lindberg et al., 2008; Hofierka and Kaňuk, 2009; Middel et al., 2017), land surface temperature (Guo et al., 2016; Li et al., 2016), and air quality (Marquez and Smith, 1999; Stone, 2008) in the urban canopy layer. City morphology can be described through Urban Canopy Parameters (UCPs) that capture the 3-dimensional geometric form, arrangement, coverage, and material composition of urban features. Various UCPs have been defined to parameterize the heterogeneous urban environment for modeling purposes, ranging from fractional parameters such as plan fractions covered by buildings, impervious surfaces, water, bare soil, and vegetation to parameters that relate to buildings and streets, e.g., average building height, building volume density, building packing density, height-to-width ratio, and sky view factor (SVF). Recent improvements in mesoscale atmospheric modeling include advanced urban canopy parameterizations based on UCPs to treat complex atmospheric processes due to heterogeneous urban morphology (Ching et al., 2009; Chen et al., 2011; Ching, 2013; Brousse et al., 2016; Xu et al., 2017; Hammerberg et al., 2018). Yet, models are sensitive to how the urban canopy layer is parameterized, i.e. the choice of UCPs, and underlying data sources (Salamanca et al., 2010; Li et al., 2017; Demuzere et al., 2017).

The World Urban Database and Access Portal Tool (WUDAPT) aims to standardize, collect, and distribute data on the form and function of cities in a common framework by building a global UCP database (Mills et al., 2015; Ching et al., 2018). WUDAPT Level 0 data - ranges of spatially averaged UCPs characterized by the Local Climate Zone (LCZ) scheme (Stewart and Oke, 2012) - are gradually being processed using a crowd-sourcing approach (Bechtel et al., 2015) and disseminated through the portal (<https://wudapt.cs.purdue.edu/wudaptTools/>). WUDAPT Level 1 and 2 data - detailed 2.5-dimensional urban form information at the neighborhood scale and 3-dimensional data at the individual building scale including building typology and materials - are more difficult to retrieve due to insufficient data availability, quality, or processing capabilities. Previous studies have derived Level 1 UCPs from stereo and SAR satellite images (Xu et al., 2017), Landsat images (Li et al., 2017), DEMs (Gál et al., 2009; Lindberg and Grimmond, 2010), 3-dimensional city models (Chen et al., 2012; Gál and Unger, 2014), and LIDAR data (Carter et al., 2011; An et al., 2014).

This study presents an image-based methodology to calculate UCPs at fine spatial scales from high resolution ground level imagery, specifically focusing on the SVF. The SVF is a measure of how much sky is visible at a given location and can be used to

\* Corresponding author at: Department of Geography and Urban Studies, Temple University, 324 Gladfelter Hall, Philadelphia, PA 19122, USA.

E-mail addresses: [ariane.middel@temple.edu](mailto:ariane.middel@temple.edu) (A. Middel), [jl@jlu.de](mailto:jl@jlu.de) (J. Lukasczyk), [rmacieje@asu.edu](mailto:rmacieje@asu.edu) (R. Maciejewski), [demuzere@ugent.be](mailto:demuzere@ugent.be) (M. Demuzere), [geomr@nus.edu.sg](mailto:geomr@nus.edu.sg) (M. Roth).

easily describe the 3-dimensional form of the built environment as a 2-dimensional metric. The SVF has been strongly related to nocturnal Urban Heat Island (UHI) effects, intra-urban air temperature distribution, and thermal comfort (Svensson, 2004; Ali-Toudert and Mayer, 2007; Unger, 2009; Matzarakis et al., 2010; Chen et al., 2012; Middel et al., 2017). Various approaches have been developed to calculate SVFs from urban morphological information. Early studies used analytical methods based on geometric properties of urban canyons, such as angle measurements and height-to-width ratios (Oke, 1981; Johnson and Watson, 1984). With the availability of 3-dimensional city models, vector-based methods have been introduced that estimate SVFs from building envelopes in GIS databases (Compagnon, 2004; Gál et al., 2009; Chen et al., 2012; Gál & Unger, 2014). For areas where digital elevation models (DEMs) or digital surface models (DSMs) are available, the SVF can also be derived using a raster image-based approach by evaluating pixel height (Ratti et al., 2005; Lindberg et al., 2008; Lindberg and Grimmond, 2010; Zakšek et al., 2011). In the absence of urban morphological information, fisheye photography has frequently been employed to calculate the SVF for discrete locations (Grimmond et al., 2001; Chapman and Thornes, 2004; Ali-Toudert and Mayer, 2007). While photographic methods yield good results, they typically require manual acquisition and processing of fisheye images, which is time consuming and not feasible at the neighborhood or city-scale.

The recent availability of street view photos from Google, Baidu, and other providers enables researchers to derive information about the form, coverage, material, and function of urban features from a street canyon perspective that is more human-centric than planar or bird's eye view approaches. This study presents a methodology to calculate city-wide SVF footprints (histograms) from Google Street View (GSV) imagery for 15 cities around the world and demonstrates the application of GSV photos for calculating UPCs in two use cases: (1) evaluation of a LCZ map for Singapore by comparing average GSV-derived SVFs with recommended urban zonal ranges; and (2) hourly street-level sun duration maps for New York and San Francisco.

## 2. Methods

To calculate a city's SVF footprint, all available GSV imagery for a given study area are retrieved through the GSV API. The images are then combined into fisheye photographs of the sky hemisphere using an equiangular fisheye projection. Next, to detect the amount of visible sky in each image, a combination of a modified Sobel filter and a flood-fill algorithm are used (Sobel and Feldman, 1968; Laungrunthip et al., 2008). The SVF for each GSV location is finally calculated employing an implementation by Middel et al. (2017). To test the accuracy of the sky detection approach, results are compared to SVFs that were derived from photographs using a deep learning approach.

### 2.1. Google Street View image retrieval and fisheye projection

First, an area of interest is defined by a latitude-longitude rectangle. Currently, the GSV API lacks functionality to retrieve a list of all available GSV image locations inside such a boundary. In fact, only image data for a specific latitude-longitude coordinate can be requested, and the API will return an image for the closest available location together with metadata that indicates the exact image coordinate and if the image was taken indoors or outdoors. To generate a uniform grid within the boundary, a spatial resolution of two meters is defined, and GSV locations are requested for each grid cell. Location tags are investigated to discard all indoor images, and redundant coordinates are removed. After deriving the grid, five images are requested per location with a 90-degree field of view: four facing the cardinal directions and one facing upwards. Next, the five images are used to generate a fisheye photograph.

Various types of fisheye lenses exist, and they correspond to different mathematical projections. A common fisheye lens for consumer use fits the equisolid angle projection (equal-area fisheye) that maintains surface relations, i.e. each pixel projected onto a unit sphere covers the same area. In contrast, a stereographic projection preserves angles and therefore does not compress marginal objects in the image, but few lenses are available for this type of projection. Most fisheye lenses manufactured today are of equiangular (equidistant) type, i.e. angular distances are maintained preserving approximately equal resolution across the whole image. Equiangular projection is often referred to as “ideal” fisheye projection and is the most widely used mathematical transformation for panoramic stitching and SVF applications.

This study calculates the position of pixels on the upper hemisphere using an equiangular projection (Steyn, 1980; Chapman et al., 2001; Middel et al., 2017):

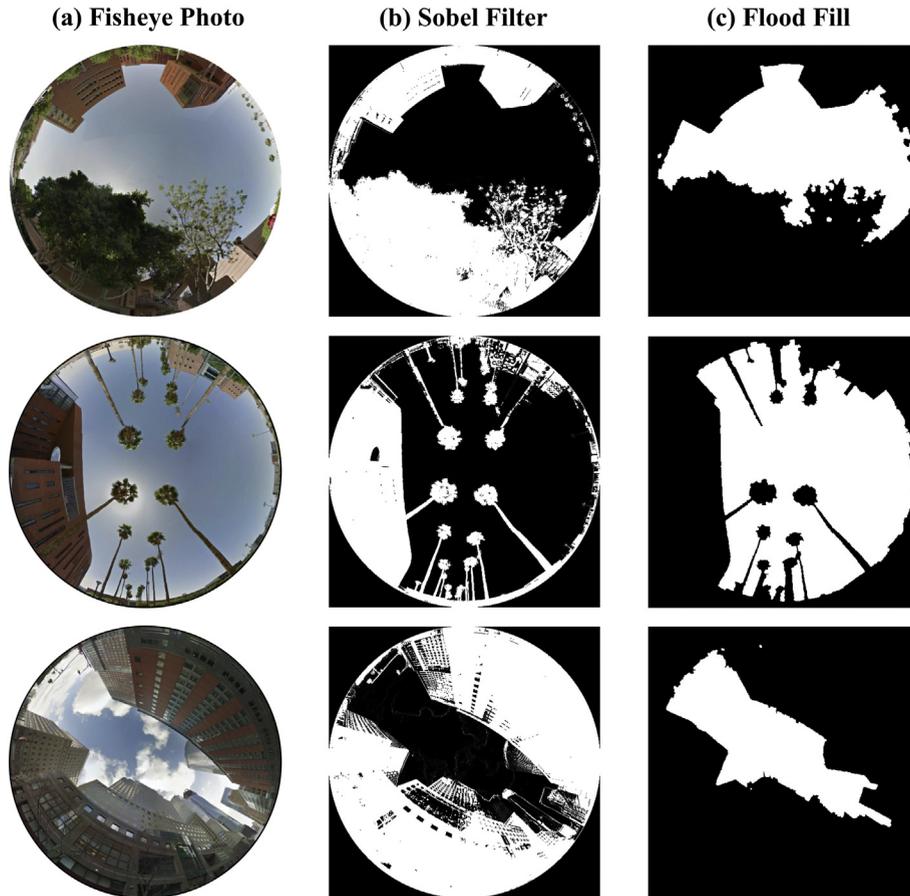
$$v = \begin{pmatrix} \sin(\theta)u_r \\ \sin(\theta)v_r \\ \cos(\theta) \end{pmatrix}$$

with latitude  $\theta = r\pi$  and normalized direction vector  $(u_r, v_r) = \begin{cases} \frac{p}{r} & \text{if } r > 0 \\ (0, 0) & \text{otherwise.} \end{cases}$

The fisheye image is treated as a unit square with center  $c = (0, 0)$  and  $p = (u, v)$  where length  $r = ||p|| \leq 0.5$  is the corresponding position of point  $p$  on the unit hemisphere. This yields a fisheye photograph of the upper hemisphere for each location within the study area.

### 2.2. Sky vs. non-sky pixel classification

To detect the sky in fisheye photographs, an algorithm presented by Laungrunthip et al. (2008) is implemented that divides the



**Fig. 1.** Three examples of sky detection in GSV fisheye photos using (a) a Sobel filter, (b) a flood-fill algorithm and (c) classification into sky and non-sky pixels.

photograph into non-overlapping areas classifying them as sky or non-sky (Fig. 1). Next, a Sobel edge detection algorithm (Sobel and Feldman, 1968) identifies the boundaries between sky and non-sky areas. This filter assigns a value between zero and one to each pixel by evaluating the contrast between pixel colors in its 8-neighborhood (adjacent pixels that share an edge or node). To make the edge detection more robust, the Sobel filter is modified to emphasize the edges between definite sky and non-sky pixel colors, i.e., if both color classes are detected within the 8-neighborhood, then the edge value for that pixel is set to one. Next, all edges with a value below a predefined threshold are removed, and the average color and color variance is calculated for each area. The remaining edges are subsequently used to flood-fill the obstructed areas (e.g., buildings, bridges, walls, and trees) to separate the image into sky and non-sky regions.

### 2.3. Sky View Factor calculation

A modified Steyn method (Steyn et al., 1986; Chapman et al., 2001; Middel et al., 2017) is implemented to compute the SVF for a fisheye photograph. Each fisheye is partitioned into  $n$  annular rings (default  $n = 36$ ) to calculate the SVF by summing up the contribution of each ring:

$$\text{SVF} = \frac{\pi}{2n} \sum_{i=1}^n \sin\left(\frac{\pi(2i-1)}{2n}\right) \left(\frac{p_i}{t_i}\right)$$

where  $p_i/t_i$  is the ratio between the number of sky pixels to the total number of pixels in ring  $i$ . An ensemble of different fisheyes and associated SVFs is shown in Fig. 2.

## 3. Sky View Factor results

### 3.1. Sky View Factor evaluation

The SVF algorithm used in the present study was previously compared to other SVF implementations (Middel et al., 2017) and

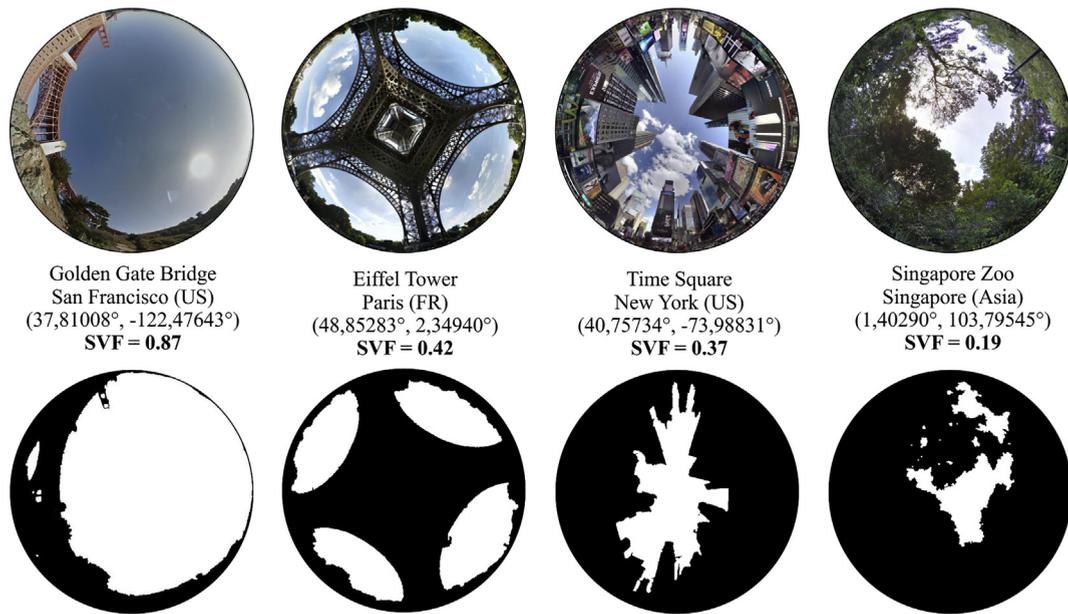


Fig. 2. Fisheye photographs generated from Google Street View (top), detected horizon limitations (bottom), and respective Sky View Factors (bolded text) for four locations around the globe.

showed excellent agreement with the traditional annulus approach by Steyn (1980). GSV-derived SVFs also compared well with SVFs from synthetic Google Earth fisheyes, yielding an average SVF difference of  $0.022 \pm 0.084$  between the two datasets. To assess the quality of the sky detection algorithm and further evaluate SVFs derived from GSV, results are compared to SVFs calculated from images that were segmented into various urban fractions using a deep learning approach. Lukasczyk et al. (under review) fine-tuned a convolutional neural network using 90-degree GSV images from 41 cities around the world to segment image cubes into five surface types and sky, respectively. The study achieved a sky classification accuracy of 98.5% for lateral and 99.6% for upward facing images.

For evaluation purposes, Lukasczyk et al.'s deep learning approach is used to segment GSV image cubes in the Phoenix metropolitan area into sky vs. non-sky pixels. The segmented image cubes are then projected onto the upper hemisphere as described in Section 2.1, and the SVF is calculated using the modified Steyn implementation (Section 2.3). Since Google regularly updates Street View imagery, the locations of the SVFs from the pixel-based classification (2015–2016 image data) and the SVFs from the deep learning approach (2017 image data) do not perfectly align. To assure reasonable proximity for comparison, a 2-m buffer is calculated around each location to pair up SVFs that fall within the intersection of two buffers. This results in 1,663,448 valid locations for evaluation with a mean SVF difference of 0.002,  $R^2 = 0.880$ , and a RMSE of 0.045 (Fig. 3). Deviations that manifest as outliers in the

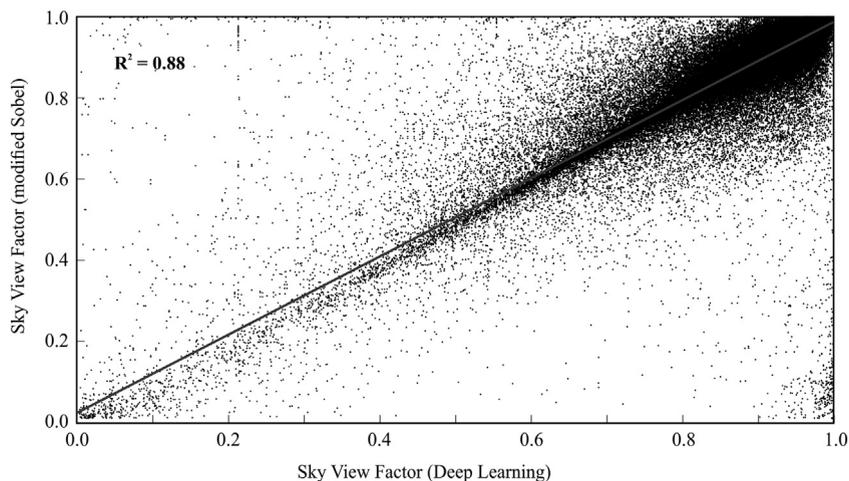


Fig. 3. Comparison of Sky View Factors calculated from Google Street View images that were classified into sky vs. non-sky pixels using deep learning (2017 image data, x-axis) vs. pixel-based Sobel filter and flood-fill approach (2015–2016 image data, y-axis) for 1,663,448 locations in the Phoenix metropolitan area.

**Table 1**  
Statistics for Sky View Factor footprint datasets from 15 cities around the globe.

Urban area	Country	Population (2016)	Land area [km <sup>2</sup> ]	# of GSV locations	SVF (average)	SVF (stdv)
Manhattan	US	1,645,000	59.1	143,183	0.545	0.243
Paris	FR	2,229,600	105.4	304,823	0.586	0.219
Singapore	SG	5,607,000	719.1	556,150	0.595	0.241
Seoul	KR	9,860,000	605.2	613,971	0.680	0.166
Tokyo	JP	9,273,000	2190	3,736,082	0.693	0.178
Vancouver	CA	647,540	115	307,856	0.713	0.234
Philadelphia	US	1,568,000	367	1,148,238	0.720	0.249
Bonn	DE	313,958	141.1	93,188	0.746	0.213
San Francisco	US	864,816	121.4	248,483	0.811	0.185
Dublin	IE	527,612	115	573,431	0.831	0.194
Los Angeles metro area	US	13,131,431	12,561	2,871,850	0.837	0.171
Be'er Sheva	IL	202,495	117.5	57,608	0.898	0.112
Dubai	IE	2,789,000	4114	609,301	0.921	0.118
Las Vegas	US	632,912	352	1,073,711	0.938	0.116
Phoenix metro area	US	4,574,531	23,490	3,600,297	0.947	0.097

scatterplot can be attributed to two reasons. First, base images for both SVF data sets were acquired in two different years leading to spatial sampling variations. Additional associated SVF differences are due to alterations in the built environment, e.g., new developments, and changes in vegetation cover during the 2-year period. Differences in acquisition years, acquisition time of day, and built structure lead to a wider spread of points along the regression line. Second, the Sobel filter-based sky detection algorithm underperforms if the contrast between the horizon limitation and the sky is too low and edge detection becomes difficult, e.g., for grey buildings in overcast conditions or glass buildings during clear sky conditions. In some cases, bright surfaces such as windows and white walls are incorrectly classified as sky, and dark clouds are identified as obstacles. In Fig. 3, the cluster of outliers to the bottom right (high SVF per deep learning, low SVF per modified Sobel and flood-fill) represents locations where clouds were wrongly classified as obstacles. In general, outliers that are not clustered around the regression line can be attributed to the second type of error.

On average, however, the approach yields reliable results with an RMSE of 0.045 and provides a viable alternative to hardware-intensive, more complex deep learning classifications. While segmenting 1000 images using the deep learning approach takes about 3 min on a single NVIDIA TITAN X Pascal graphics card, the Sobel filter and flood-fill approach runs in a regular web browser and takes about 30 s, making the presented approach much more efficient and accessible.

### 3.2. Sky View Factor footprints

Complete GSV image coverage is retrieved and SVFs are calculated for 15 urban areas in Asia, North America, and Europe (Table 1, Fig. 4). The study sites are located in various climatic zones (Köppen climate classification types Af, BSh, BWh, Cfa, Csb, Csc, Dwa) and exhibit a variety of urban forms, vegetation, and building materials. The areal extent ranges from 59.1 km<sup>2</sup> (New York City's borough Manhattan) to 23,490 km<sup>2</sup> (Phoenix metropolitan area), and the number of GSV locations varies from 57,608 (Be'er Sheva) to 3,736,082 (Tokyo) with a total of 15,938,172 GSV locations. All images were retrieved between December 2015 and September 2016, and image acquisition dates range from April 2015 to July 2016.

Fig. 5 illustrates the spatial distribution of SVFs in each urban area, arranged by average SVF magnitude, from largest to smallest mean SVF. Similarly, Fig. 6 shows histograms of the SVF distribution (bin size 0.025), in the following referred to as SVF footprints. The four urban areas with the highest average SVF are Phoenix, Las Vegas, Dubai, and Be'er Sheva. Their SVF footprints are negatively skewed with a long tail to the left and no visible normal distribution. Phoenix and Las Vegas, which show the highest average SVFs of 0.947 and 0.938, are very similar in urban form. Both metropolitan areas mostly consist of LCZ 6 (open lowrise) with detached residential single-family homes and wide arterial roads with up to 7 lanes. Los Angeles and Dublin exhibit very similar SVF footprints with mean SVF values of 0.837 and 0.831, respectively. Bonn and Philadelphia also compare well in terms of SVF footprints (0.746 vs. 0.720); both cities feature compact midrise (LCZ 2) neighborhoods with townhouses near the city center and residential suburbs with increased tree canopy cover. Tokyo's and Seoul's SVF footprints average just under 0.7, while Singapore, Paris, and Manhattan have the lowest SVFs with 0.595, 0.586, and 0.545. Paris is the only urban area to exhibit a bimodal SVF distribution, with a peak at 0.550 and another peak at 0.975. While the first peak is representative of the inner city urban form, the second peak mostly encompasses locations on highways and along the Seine river. Singapore's urban form will be discussed in more detail in Section 4.1.

## 4. Use cases

SVF footprints can serve a variety of urban climate applications from micro- to mesoscale. In this section, two use cases are provided. The first example demonstrates the potential of detailed city-wide SVF information for evaluating WUDAPT level 0 data. The second example outlines how SVF footprints can be used to calculate solar access (sun duration) at street level.

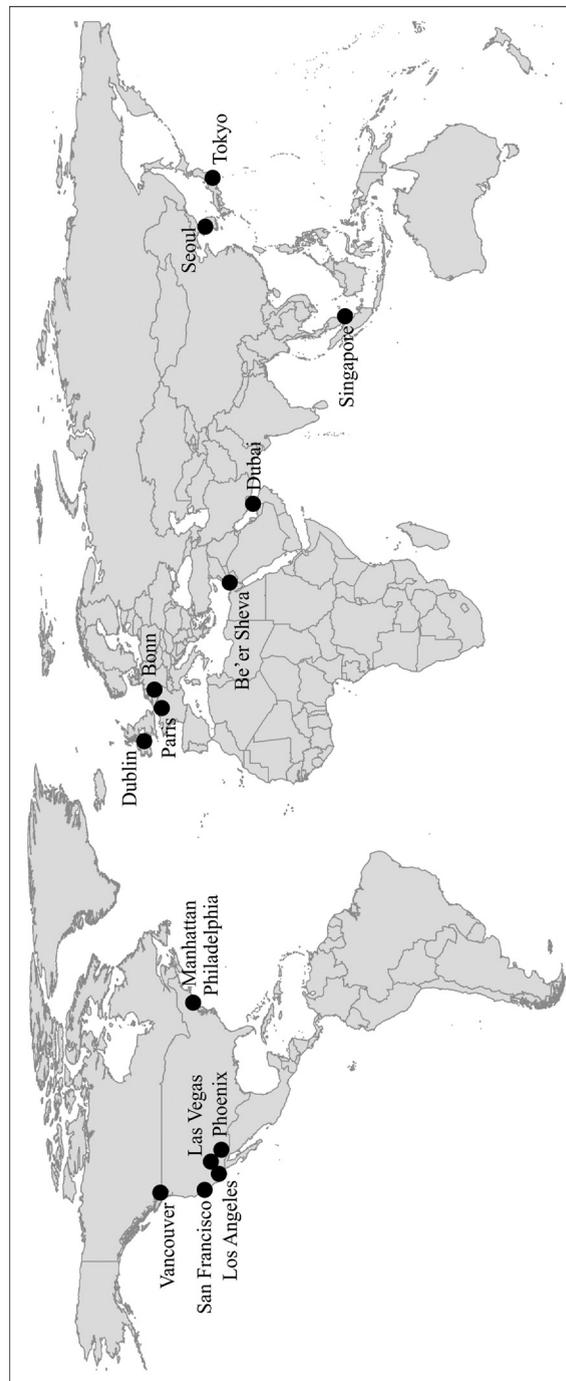
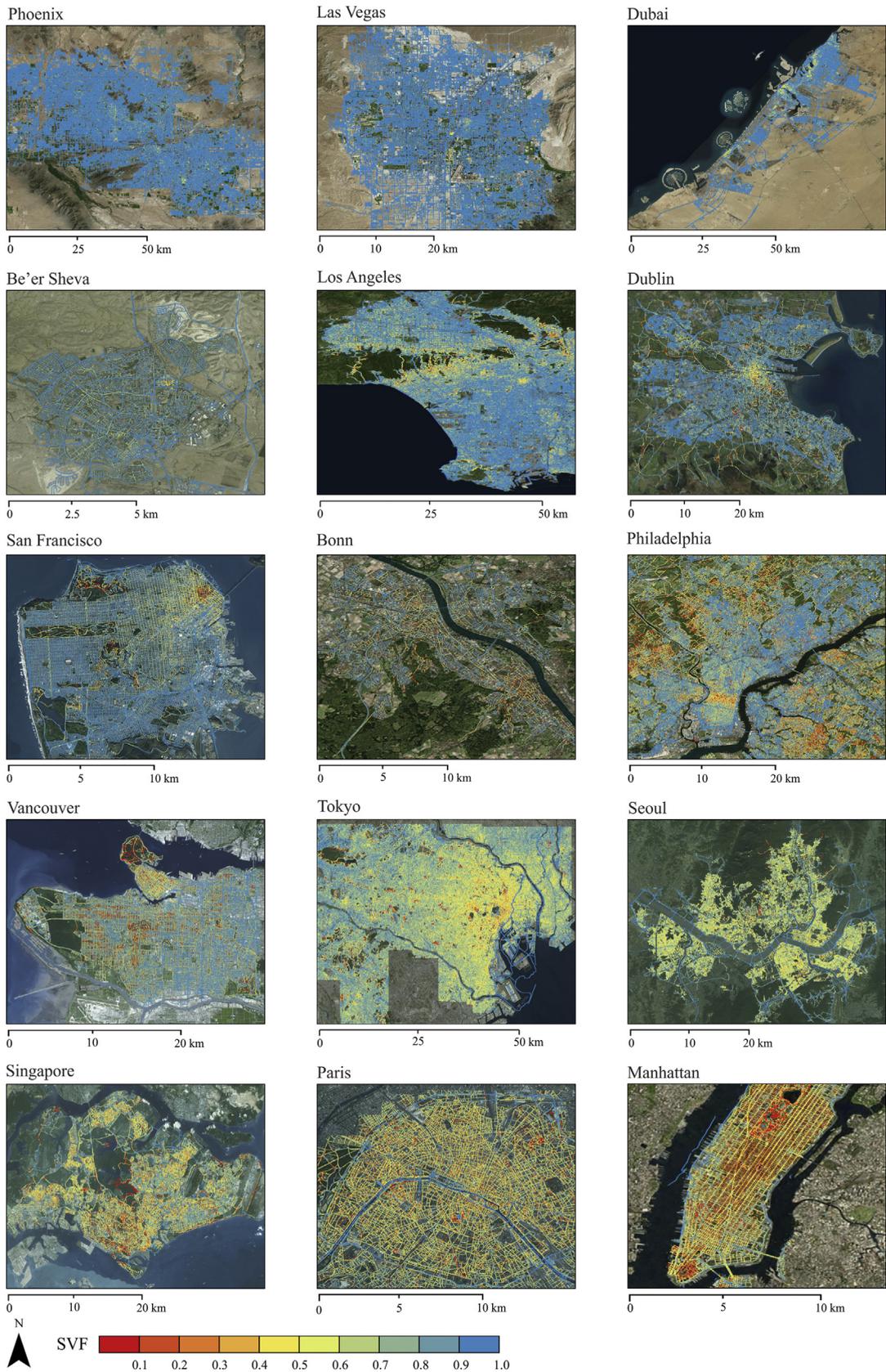


Fig. 4. Location of urban areas under investigation.



(caption on next page)

Fig. 5. Sky View Factor distribution for 15 cities around the world, sorted by average Sky View Factor from highest (Phoenix) to lowest (Manhattan).

#### 4.1. Local climate zone evaluation

The LCZ scheme, developed by [Stewart and Oke \(2012\)](#), aims to help standardize methods of UHI observations and documentation by characterizing urban morphology and function at the local scale. WUDAPT is a crowdsourcing effort to establish a census of cities globally, starting with L0 level information (LCZs) that provides ranges of spatially averaged, generic UCPs. Aggregated to LCZ spatial units, higher level data (L1 and L2) can be used to assess the quality of a LCZ classification. Here, WUDAPT L0 data are evaluated for Singapore using SVFs from GSV.

The default protocol for generating WUDAPT L0 data is based on a supervised random forest classification, implemented in SAGA GIS, using freely available cloud-free Landsat 8 images and training areas developed in Google Earth ([Bechtel et al., 2015](#), [Conrad et al., 2015](#)). Here, the Singapore LCZ map is developed using an alternative approach based on the Google Earth Engine (EE), a cloud-based platform for planetary-scale geospatial analysis ([Gorelick et al., 2017](#), [Demuzere et al., under review](#)). Taking advantage of EE's public data catalogue of widely used geospatial datasets, the following input features for the classifier are used (covering the period 2014–2016):

- two Landsat 8 composites (May to September and November to March) including bands 2 to 7, 10, and 11;
- four spectral indices derived from the Landsat 8 bands: minimum and maximum Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and the Normalized Difference Built Index (NDBI);
- Normalized Difference Urban Index (NDUI) after [Zhang et al. \(2015\)](#), which combines the maximum of the spectral indices with the coarse resolution Defense Meteorological Satellite Program/Operational Linescan System (DMSP/OLS) nighttime light imagery;
- Sentinel-1 composite based on all Sentinel-1 data filtered by VV single co-polarization, the Interferometric Wide swath acquisition mode and both ascending and descending orbits.

The 274 training areas developed for Singapore are randomly split into training (70%) and testing (30%) data and classified into LCZ categories; this action is performed 25 times, and the final LCZ map is comprised of the modal category ([Fig. 7](#)). [Verdonck et al. \(2017\)](#) stressed the importance of a quantitative accuracy assessment, and therefore the indices suggested in [Bechtel et al. \(2017\)](#) are used: OA (overall accuracy of all pixels),  $OA_{bu}$  (overall accuracy of urban pixels only),  $OA_{bu}$  (overall accuracy of built vs. natural types only, ignoring their internal differentiation) and  $OA_w$  (weighted accuracy taking into account similarity between classes). Finally, the F1 metric provides the class-wise accuracy ([Verdonck et al., 2017](#)).

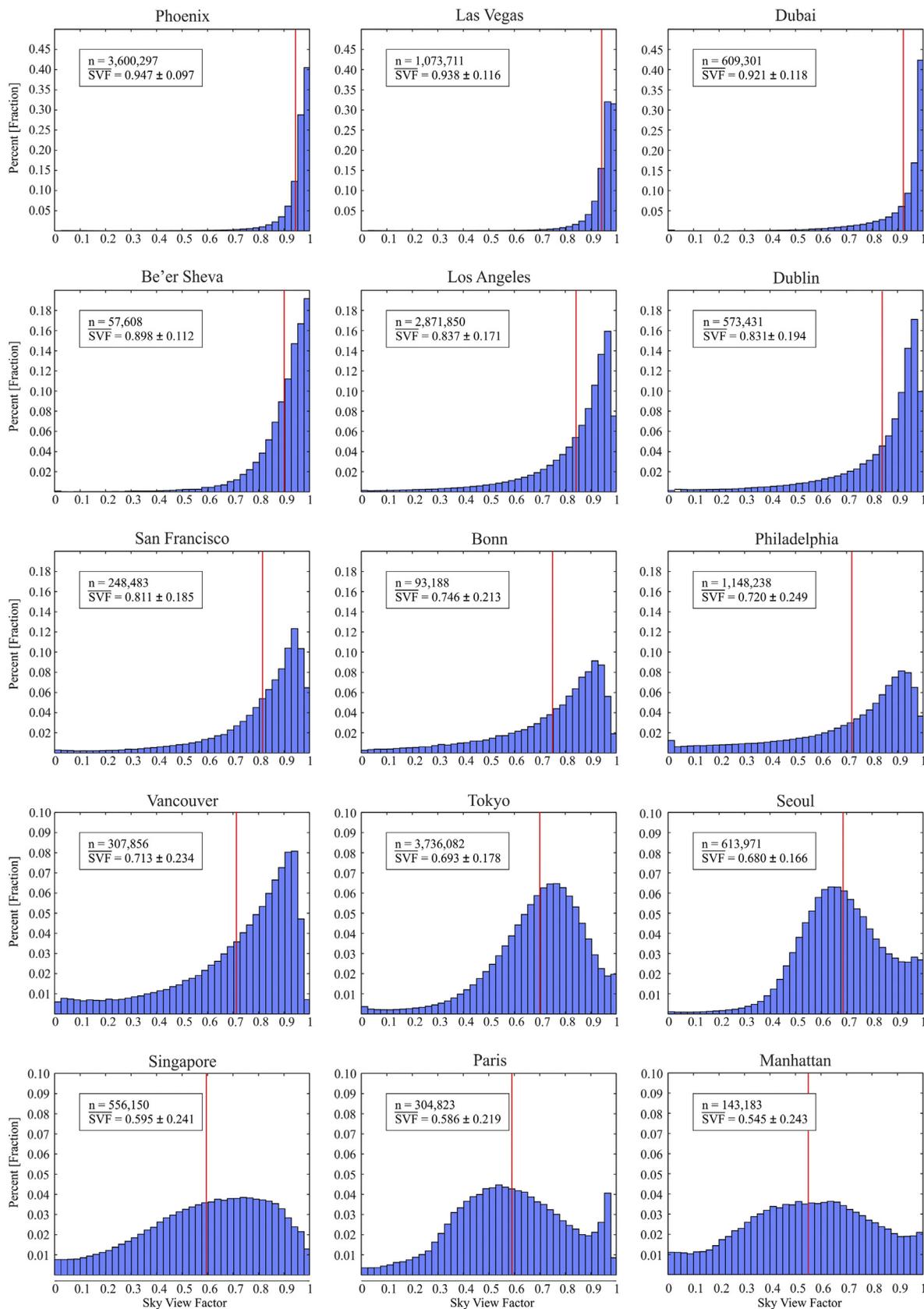
The modal LCZ map and the corresponding accuracy measures are depicted in [Fig. 7](#). Considering the built-up classes only, the region of interest is largely dominated by LCZ 3 (compact low-rise, 9%), 4 (open high-rise, 16.9%) and 9 (sparsely built, 22.6%). Industrial areas occupying large parts of the Southwest of Singapore are also clearly visible, represented by LCZs 8 (large low-rise), 10 (heavy industry) and E (bare rock or paved, in this case oil depots or container port infrastructure). The median overall accuracy of the LCZ map is 70.4%, with similar or higher values for  $OA_{bu}$ ,  $OA_{bu}$  and  $OA_w$ . The classification therefore meets the accuracy requirements (> 50% for all accuracy metrics) for review and subsequent dissemination via the WUDAPT portal ([Bechtel et al., under review](#)). The class-wise F1 metric, however, indicates a strong variability in accuracy: LCZs 1 (compact high-rise), 3, 4, 8, A (dense trees), F (bare soil or sand) and G (water) are generally correctly classified. On the other hand, low skill is obtained for LCZs 2 (compact mid-rise), 6 (open low-rise) and D (low plants).

[Fig. 8](#) provides a LCZ class-wise comparison of SVF ranges derived from the training areas, the full LCZ map, and those suggested by [Stewart and Oke \(2012\)](#). The values derived from the full LCZ map generally match those directly derived from the training area polygons only. This already demonstrates the potential value of the SVF dataset in evaluating WUDAPT L0 data across a larger region of interest. For this use case, the classifier is able to retain the relevant information from the training areas and transfer it to the remaining out-of-sample pixels.

The median SVF values for the 'compact' LCZ classes 1, 2, and 3 are at the upper end of the ranges reported by [Stewart and Oke \(2012\)](#), while the opposite is true for the 'open' LCZ classes 4, 5, and 6. This is counterintuitive, as one would expect higher SVF values for open mid- and low-rise compared to their compact counterparts. The large low-rise (LCZ 8) and especially the sparsely built (LCZ 9) classes have values below the ranges reported by [Stewart and Oke \(2012\)](#), while the median of the heavy industry class (LCZ 10) does fall within the reported range. Except for LCZ B (scattered trees), the natural classes have lower SVF values than the [Stewart and Oke \(2012\)](#) reference ranges. This is probably due to a lack of GSV image samples in such areas, which is further discussed below.

#### 4.2. Sun duration

Determining solar access in street canyons, i.e. the ability to receive sunlight without obstruction, is essential for heat mitigation in hot urban areas ([Middel et al., 2016](#)) and to estimate the potential for solar heating and photovoltaic electricity production ([Compagnon, 2004](#)). Solar access varies by geographic location (latitude and longitude), diurnally, and by day of year, but it also depends on the horizon limitation imposed by the built environment. It can be quantified by overlaying a sun path polar diagram over a hemispherical view such as the processed fisheye photos.



(caption on next page)

**Fig. 6.** Sky View Factor footprints, sorted from largest (Phoenix, top left) to smallest footprint (Manhattan, bottom right) for  $n$  Google Street View locations in each city; histograms show distribution of Sky View Factors in each city, average Sky View Factor (SVF), and standard deviation (see also Table 1).

An algorithm by Blanco-Muriel et al. (2001) is implemented to calculate the sun position in the form of a solar vector for any time of the day and any day of the year. The diurnal sun trajectory is approximated by calculating the sun's position every 15 min for a chosen date and location. The resulting sun path, given as azimuth and zenith angle pairs, is projected onto the black and white sky/non-sky GSV fisheye generated during the SVF calculation process. Finally, fisheye pixels in the 8-neighborhood of each sun position along the path are tested for sun occlusion, and the unobstructed time periods (at least 5 pixels in the 8-neighborhood are sky pixels) are summarized into total daily sun duration, i.e. the number of hours of direct incoming solar radiation for a specific location and date. Based on hemispherical views for Manhattan and San Francisco from the SVF dataset, street network maps of SVFs and daily sun duration are created for the two cities for July 21, 2015 and November 2, 2016, respectively (Fig. 9). These maps can be used to estimate the range of solar-powered cars in street canyons; calculate the energy use for air-conditioning in buildings during the summer months; predict snow melting in the winter; and estimate pedestrian exposure to UV for skin cancer prevention.

## 5. Discussion

### 5.1. Sky View Factor footprints

Our approach to retrieve SVFs from GSV works at fine spatial scales (distance between GSV locations is typically 5–10 m), includes all available GSV locations (not a stratified sample), and can be used to achieve close to global urban coverage. GSV images are usually updated annually, which enables long-term monitoring of urban form. The presented approach has several limitations. The implemented sky detection algorithm has difficulties discovering edges if the contrast between two features, e.g. building and sky, is too low. Also, bright surfaces such as windows and white walls might be falsely classified as sky, while dark clouds could be identified as obstacles. Although trees are reliably detected as sky obstructions, the flood-fill algorithm overestimates the density of tree canopy cover of sparsely-leaved trees (see vegetation in Figs. 2 and 3). This is advantageous for comparing SVFs independent of image acquisition dates, as SVF results for leaf-off and leaf-on seasons will be similar, but limits the ability to investigate seasonal SVF differences.

Current state-of-the-art approaches use convolutional neural networks to segment images more accurately into sky and non-sky regions (Gebu et al., 2017; Iannelli and Dell'Acqua, 2017; Lukaszczuk et al., under review). Although these methods yield more accurate results, they require an elaborate training process, specialized hardware, and increased computation time. As demonstrated in the evaluation section, the Sobel filter and flood-fill approach yields reliable results (RMSE = 0.045) and provides a viable and more accessible alternative to hardware-intensive machine learning solutions. The presented sky detection and SVF calculation can be performed in a regular web-browser and would therefore be suitable for crowdsourcing SVFs.

GSV does not offer complete coverage around the world. Although GSV data exist for the southern hemisphere of the globe (e.g. São Paulo in Brazil and Kampala in Africa), coverage is less dense in those areas (see <https://www.google.com/streetview/understand/#where>). GSV is also not available in some Asian countries, although Baidu provides an alternative for China. With regards to intra-city coverage, GSV images are inherently biased towards street locations and therefore do not provide continuous spatial coverage of the urban environment. Under-sampled areas include parks, golf courses, backyards, and natural areas in general. This sampling bias becomes particularly evident in the WUDAPT L0 evaluation use case where natural classes exhibit consistently lower SVFs than the recommended LCZ ranges. Google is constantly improving Street View coverage, with new additions of hiking trails and pedestrian areas captured by 360° cameras mounted on backpacks in recent months. More complete coverage is to be expected, potentially incorporating stereo-photography as employed in Google Earth. Until continuous coverage is available, the SVF approach could be complemented through Lidar data or digital surface models.

### 5.2. Local climate zone evaluation

Care should be taken when generalizing the LCZ results from Singapore (e.g. Fig. 8) to other cities. A number of local features impose limitations on the quality of the LCZ classification, the representativeness of GSV-derived SVF results, and finally the use of SVF to infer a specific LCZ. The particular nature of Singapore's building regulations pose challenges to the accurate determination of a specific LCZ. Similar to many other Asian cities, low- and high-rise developments may occur side-by-side due to changes in zoning laws regulating the density of developments and/or height of buildings. Consequently, the downtown core often contains a mix of very tall (> 100 m) and short (10–15 m) buildings, and many residential areas are in a transition phase where various 2-storey high buildings are converted into 5-storey, small apartment blocks. The results in Fig. 8 are promising, because the derived SVF values show the expected increase with decreasing height of roughness elements (in particular for the 'compact' built zones), however, the range across LCZs 1–6 is only about 0.3 with little difference between the 'compact' and 'open' built zones.

Not all LCZs are equally well represented through GSV imagery, which may result in erroneous SVFs. While many of the commercial and residential zones are crisscrossed by several roads, other zones suffer from under-sampling or no sampling at all. For example, no GSV images are available for LCZ E (mostly off-limits) in which case the SVF is taken from scattered pixels in the LCZ map that have been confused by the classifier, or that are under-sampled such as LCZ 10 (partly off-limits). For many of the natural

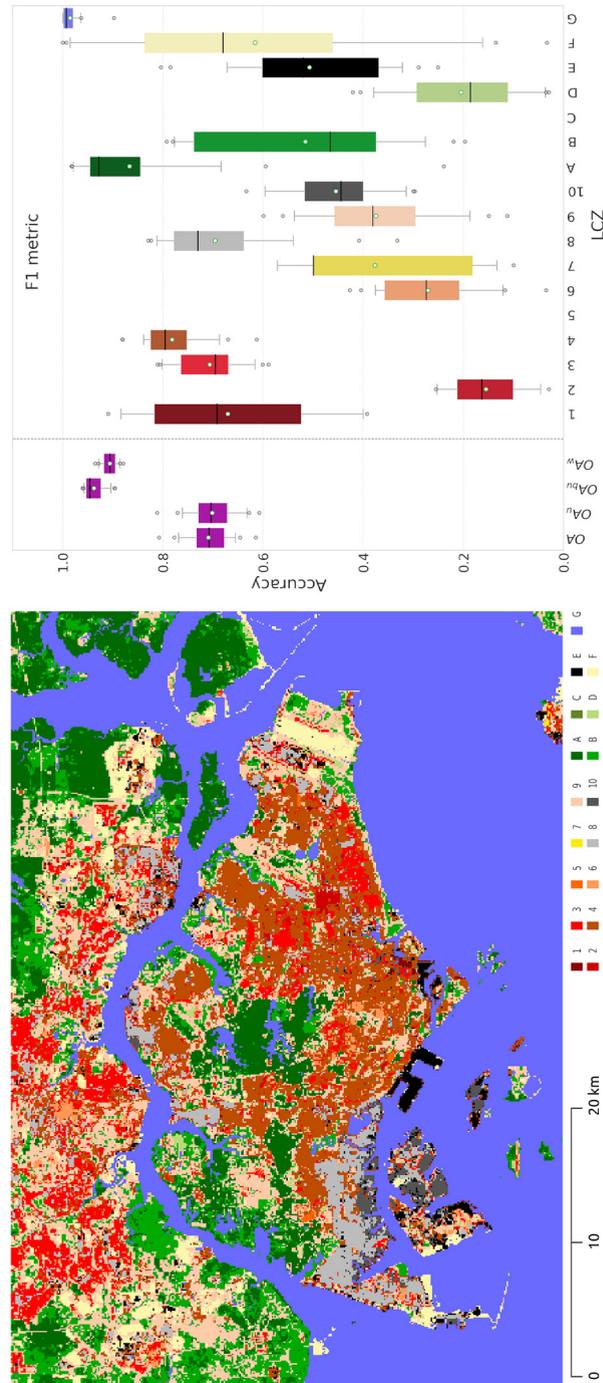
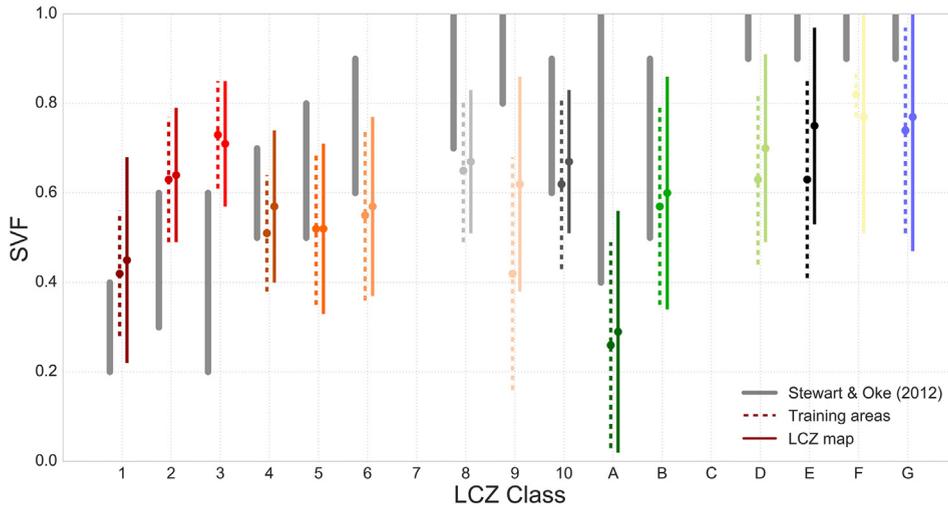


Fig. 7. Singapore Local Climate Zone map (left) and accuracies (right). See Stewart and Oke (2012) for a detailed description of Local Climate Zones.



**Fig. 8.** Sky View Factor comparison for Singapore. The thick grey bars are Sky View Factor ranges from Stewart & Oke (2012), colored filled circles and lines are the median  $\pm$  1 standard deviation derived from the training areas (dashed lines) and for the full Local Climate Zone map (full lines), respectively. The colors refer to the Local Climate Zone class, as in Fig. 7.

areas (e.g. LCZs B-D), GSV images are only available along roads, which will likely result in an underprediction of the LCZ-wide SVF given the frequent presence of trees planted along roads (Fig. 8). The similar low bias observed for LCZs 4–6 may also be due to often mature street trees, sometimes forming closed canopies above the road surface, which are abundant along many residential roads. In the case of LCZ A, a value of around 0.3 is reasonable considering the local dense tropical forests. While these measurement issues contribute to less extreme SVF values, the small range observed in Fig. 8 is also real and a reflection of the fact that there are very few expansive, really densely built or truly open spaces in Singapore.

Finally, as shown above, evaluation of LCZs using SVF values should be approached with caution given the various issues associated with both the intrinsic difficulty to accurately determine a LCZ and the estimation of representative SVFs from GSV imagery. The SVF is one of many parameters defining a LCZ and the ranges given in Stewart and Oke (2012) are probably best interpreted as approximations of real values based on past studies and expectations for a certain roughness element geometry. A SVF of 0.6 for example is indicative of six of the ten LCZs of the ‘Built Zone’ series. Yet, the GSV-derived SVFs provide an accessible, widely-available resource to analyze LCZ parameters across multiple cities, because the methodology does not depend on the availability of fine-scale information such as 3-dimensional models or LIDAR data. Efforts are currently underway to compare GSV SVFs to LCZ ranges for various WUDAPT datasets including Dublin (Bechtel et al., under review), Phoenix, and Las Vegas (Wang et al., 2018).

### 5.3. Sun duration

The precision of the calculated sun duration mainly depends on the sky detection accuracy but is also sensitive to seasonal influences and image acquisition dates. GSV images are collected during all seasons, but deciduous trees provide more solar access at street level during leaf-off months and increased shade for the rest of the year, generating different sun duration outcomes depending on the season. Yet, the impact of deciduous trees on sun duration is limited to morning and afternoon. It is also minimized at the center of the road and in urban areas where streets are lined by buildings that dominate shading over trees. Tree pruning further impacts shade patterns in street canyons, and so do urban infrastructure changes such as the addition of new buildings. Processing the most recent GSV dataset to retrieve SVFs is crucial to yield reliable sun duration results.

## 6. Conclusions

UCPs are crucial descriptors of urban form that resolve the spatial heterogeneity for urban climate applications. This study introduces an innovative approach to calculate UCPs from Google imagery, specifically focusing on the SVF as a dimension-reduced metric representative of 3-dimensional urban form. This methodology enables the retrieval of urban morphology from a human-centric street canyon perspective at low computational costs with minimum hardware requirements. Calculating typical SVF footprints for 15 cities around the world, this study demonstrates how SVFs from GSV can be used to estimate sun duration and to assess the quality of WUDAPT L0 data. Scaled up to other geographic locations, the growing SVF dataset can inform various climate modeling efforts around the world, from microscale to mesoscale, and is a first step towards building a database of L1 information for the WUDAPT portal, complementary to other UCPs from GSV such as building and tree fractions that are currently being processed (Lukaszczuk et al., under review). Fine-scale standardized and globally available UCPs are paramount in observation and modeling efforts to deepen the understanding of urban climate dynamics as they relate to the form and function of cities. Big-Data approaches such as the GSV application presented in this study highlight new innovative paths for the urban climate community to achieve this

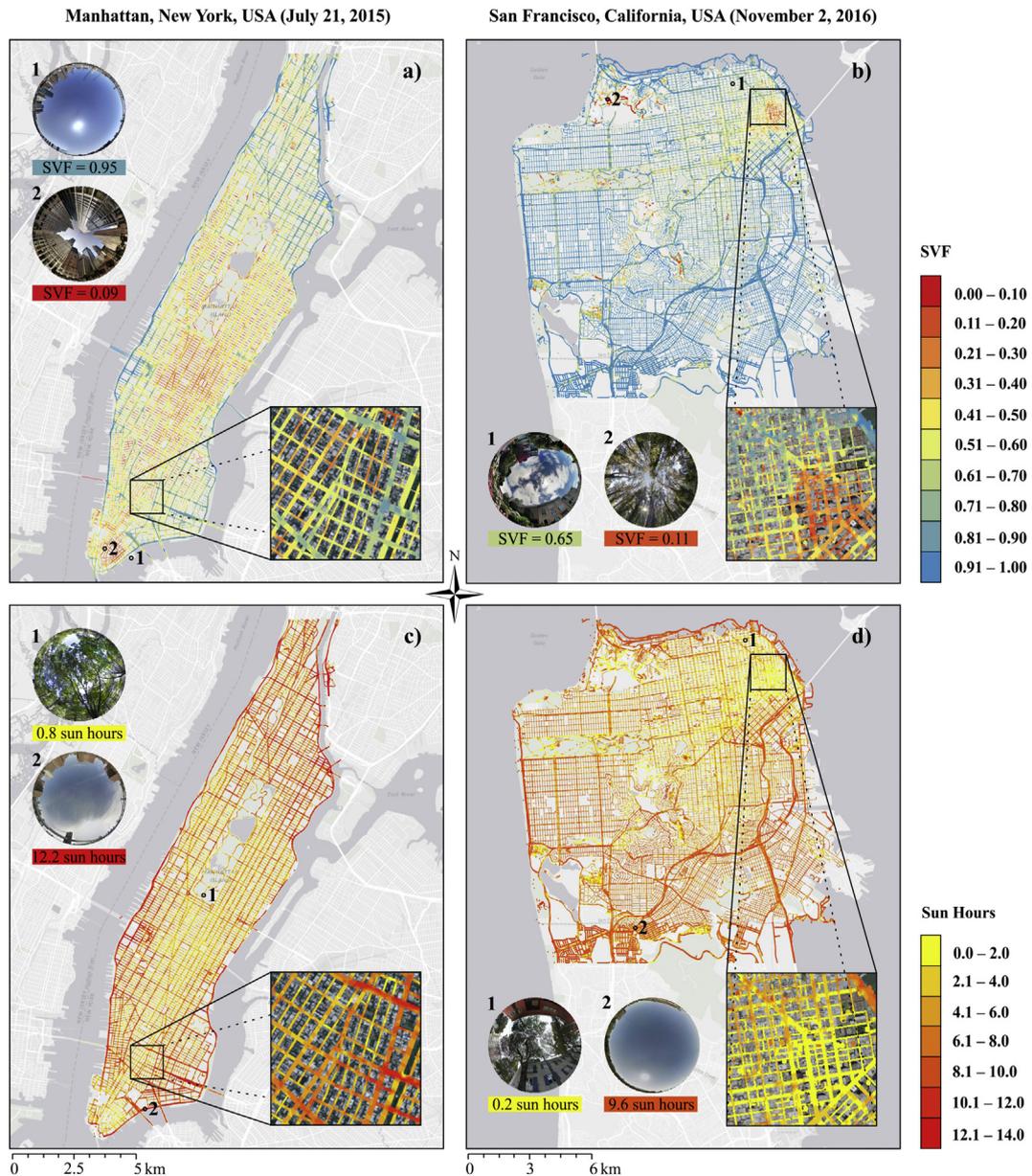


Fig. 9. Sky View Factor (top) and sunshine hours (bottom) for Manhattan (left) and San Francisco (right).

goal.

### Acknowledgements

This research was sponsored by Technische Universität Kaiserslautern, grant “Microclimate Data Collection, Analysis, and Visualization” and supported by National Science Foundation (NSF) Award Number 1635490, “A Simulation Platform to Enhance Infrastructure and Community Resilience to Extreme Heat Events.” and NSF Award Number 1639227, “Flexible Model Compositions and Visual Representations for Planning and Policy Decisions for the Food-Energy-Water Nexus”. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsoring organizations.

### References

Ali-Toudert, F., Mayer, H., 2007. Effects of asymmetry, galleries, overhanging façades and vegetation on thermal comfort in urban street canyons. *Sol. Energy* 81 (6),

- 742–754.
- An, S.M., Kim, B.S., Lee, H.Y., Kim, C.H., Yi, C.Y., Eum, J.H., Woo, J.H., 2014. Three-dimensional point cloud based sky view factor analysis in complex urban settings. *Int. J. Climatol.* 34 (8), 2685–2701.
- 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 (1), 199–219.
- Bechtel, Benjamin, Demuzere, Matthias, Sismanidis, Panagiotis, Fenner, Daniel, Brousse, Oscar, Beck, Christoph, Van Coillie, Fricke, et al., 2017. Quality of crowd-sourced data on urban morphology—the human influence experiment (HUMINEX). *Urban Sci.* 1 (2). <http://dx.doi.org/10.3390/urbansci1020015>.
- Bechtel, B., Alexander, P. J., Beck, C., Brousse, O., Ching, J., Demuzere, M., Gal, T., Hidalgo, J., Hoffman, P., Middel, A. M., Mills, G., Ren, C., See, L., Sismanidis, P., Verdonck, M.-L., Xu, G., Xu, Y., under review, Generating WUDAPT Level 0 data - current status of production and evaluation. *Urban Clim.*
- Blanco-Muriel, M., Alarcón-Padilla, D.C., López-Moratalla, T., Lara-Coira, M., 2001. Computing the solar vector. *Sol. Energy* 70 (5), 431–441.
- Brousse, O., Martilli, A., Foley, M., Mills, G., Bechtel, B., 2016. WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. *Urban Clim.* 17 (Supplement C), 116–134.
- Carter, M., Shepherd, J.M., Burian, S., Jayachandran, I., 2011. Integration of Lidar data into a coupled mesoscale–land surface model: a theoretical assessment of sensitivity of urban–coastal mesoscale circulations to urban canopy parameters. *J. Atmos. Ocean. Technol.* 29 (3), 328–346.
- Chapman, L., Thornes, J., 2004. Real-time sky-view factor calculation and approximation. *J. Atmos. Ocean. Technol.* 21 (5), 730–741.
- Chapman, L., Thornes, J.E., Bradley, A.V., 2001. Rapid determination of canyon geometry parameters for use in surface radiation budgets. *Theor. Appl. Climatol.* 69 (1–2), 81–89.
- Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C.S.B., Grossman-Clarke, S., Loidan, T., Manning, K.W., Martilli, A., Miao, S., Sailor, D., Salamanca, F.P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A.A., Zhang, C., 2011. The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.* 31 (2), 273–288.
- Chen, L., Ng, E., An, X., Ren, C., Lee, M., Wang, U., He, Z., 2012. Sky view factor analysis of street canyons and its implications for daytime intra-urban air temperature differentials in high-rise, high-density urban areas of Hong Kong: a GIS-based simulation approach. *Int. J. Climatol.* 32 (1), 121–136.
- Ching, J., 2013. A perspective on urban canopy layer modeling for weather, climate and air quality applications. *Urban Clim.* 3, 13–39.
- Ching, J., Brown, M., McPherson, T., Burian, S., Chen, F., Cionco, R., Hanna, A., Hultgren, T., Sailor, D., Taha, H., Williams, D., 2009. National urban database and access portal tool. *Bull. Am. Meteorol. Soc.* 90 (8), 1157–1168.
- 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., Brunzell, N., Pesaresi, M., Miao, S., Mu, Q., Chen, F., Theeuwes, N., 2018. World Urban Database and Access Portal Tools (WUDAPT), an urban weather, climate and environmental modeling infrastructure for the Anthropocene. *Bull. Am. Meteorol. Soc.* <http://dx.doi.org/10.1175/BAMS-D-16-0236.1>.
- Compagnon, R., 2004. Solar and daylight availability in the urban fabric. *Energy. Build.* 36 (4), 321–328.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1. 4. *Geosci. Model Dev.* 8 (7), 1991–2007.
- Demuzere, M., Harshan, S., Järvi, L., Roth, M., Grimmond, C.S.B., Masson, V., Oleson, K.W., Velasco, E., Wouters, H., 2017. Impact of urban canopy models and external parameters on the modelled urban energy balance in a tropical city. *Q. J. R. Meteorol. Soc.* 143 (704), 1581–1596. <http://dx.doi.org/10.1002/qj.3028>.
- Demuzere, M., Bechtel, B., Mills, G., under review. Global transferability of local climate zone models. *Urban Clim.*
- Gál, T., Unger, J., 2014. A new software tool for SVF calculations using building and tree-crown databases. *Urban Clim.* 10, 594–606.
- Gál, T., Lindberg, F., Unger, J., 2009. Computing continuous sky view factors using 3D urban raster and vector databases: comparison and application to urban climate. *Theor. Appl. Climatol.* 95 (1–2), 111–123.
- Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E.L., Fei-Fei, L., 2017. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proc. Natl. Acad. Sci.* 114 (50), 13108–13113.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* <http://dx.doi.org/10.1016/j.rse.2017.06.031>.
- Grimmond, C.S.B., Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Meteorol.* 38 (9), 1262–1292.
- Grimmond, C., Potter, S., Zutter, H., Souch, C., 2001. Rapid methods to estimate sky-view factors applied to urban areas. *Int. J. Climatol.* 21 (7), 903–913.
- Guo, G., Zhou, X., Wu, Z., Xiao, R., Chen, Y., 2016. Characterizing the impact of urban morphology heterogeneity on land surface temperature in Guangzhou, China. *Environ. Model. Softw.* 84 (Supplement C), 427–439.
- Hammerberg, K., Brousse, O., Martilli, A., Mahdavi, A., 2018. Implications of employing detailed urban canopy parameters for mesoscale climate modelling: a comparison between WUDAPT and GIS databases over Vienna, Austria. *Int. J. Climatol.* <http://dx.doi.org/10.1002/joc.5447>.
- Hofierka, J., Kaňuk, J., 2009. Assessment of photovoltaic potential in urban areas using open-source solar radiation tools. *Renew. Energy* 34 (10), 2206–2214.
- Iannelli, G.C., Dell'acqua, F., 2017. Extensive exposure mapping in urban areas through deep analysis of street-level pictures for floor count determination. *Urban Sci.* 1 (2), 16.
- Johansson, E., 2006. Influence of urban geometry on outdoor thermal comfort in a hot dry climate: a study in Fez, Morocco. *Build. Environ.* 41 (10), 1326–1338.
- Johnson, G.T., Watson, I.D., 1984. The determination of view-factors in urban canyons. *J. Clim. Appl. Meteorol.* 23 (2), 329–335.
- Laungrunthip, N., Mckinnon, A.E., Churcher, C.D., Unsworth, K., 2008. Edge-based detection of sky regions in images for solar exposure prediction. In: Paper Presented at the 2008 23rd International Conference Image and Vision Computing New Zealand.
- Li, X., Li, W., Middel, A., Harlan, S., Brazel, A., Turner, B., 2016. Remote sensing of the surface urban heat island and land architecture in Phoenix, Arizona: combined effects of land composition and configuration and cadastral–demographic–economic factors. *Remote Sens. Environ.* 174, 233–243.
- Li, Y., Miao, S., Chen, F., Liu, Y., 2017. Introducing and evaluating a new building-height categorization based on the fractal dimension into the coupled WRF/urban model. *Int. J. Climatol.* 37 (7), 3111–3122.
- Lindberg, F., Grimmond, C., 2010. Continuous sky view factor maps from high resolution urban digital elevation models. *Clim. Res.* 42 (3), 177–183.
- Lindberg, F., Holmer, B., Thorsson, S., 2008. SOLWEIG 1.0—modelling spatial variations of 3D radiant fluxes and mean radiant temperature in complex urban settings. *Int. J. Biometeorol.* 52 (7), 697–713.
- Lukaszczuk, J., Middel, A., Zakrzewski, S., Arnold, M., Maciejewski, R., under review, Urban form and composition of street canyons: a human-centric big data and deep learning approach. *Landsch. Urban Plan.*
- Marquez, L.O., Smith, N.C., 1999. A framework for linking urban form and air quality. *Environ. Model. Softw.* 14 (6), 541–548.
- Martilli, A., Clappier, A., Rotach, M.W., 2002. An urban surface exchange parameterisation for mesoscale models. *Bound.-Layer Meteorol.* 104 (2), 261–304.
- Matzarakis, A., Rutz, F., Mayer, H., 2010. Modelling radiation fluxes in simple and complex environments: basics of the RayMan model. *Int. J. Biometeorol.* 54 (2), 131–139.
- Middel, A., Häb, K., Brazel, A.J., Martin, C., Guhathakurta, S., 2014. Impact of urban form and design on microclimate in Phoenix, AZ. *Landsch. Urban Plan.* 122, 16–28.
- Middel, A., Selover, N., Hagen, B., Chhetri, N., 2016. Impact of shade on outdoor thermal comfort—a seasonal field study in Tempe, Arizona. *Int. J. Biometeorol.* 1–13.
- Middel, A., Lukaszczuk, J., Maciejewski, R., 2017. Sky view factors from synthetic fisheye photos for thermal comfort routing—a case study in Phoenix, Arizona. *Urban Plan.* 2 (1), 19–30.
- Mills, G., Ching, J., See, L., Bechtel, B., Foley, M., 2015. An introduction to the WUDAPT project. In: 9th International Conference on Urban Climate, Toulouse.
- Ng, E., Yuan, C., Chen, L., Ren, C., Fung, J.C., 2011. Improving the wind environment in high-density cities by understanding urban morphology and surface roughness: a study in Hong Kong. *Landsch. Urban Plan.* 101 (1), 59–74.
- Oke, T.R., 1981. Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. *J. Climatol.* 1 (3), 237–254.
- Ratti, C., Baker, N., Steemers, K., 2005. Energy consumption and urban texture. *Energy. Build.* 37 (7), 762–776.
- Salamanca, F., Martilli, A., Tewari, M., Chen, F., 2010. A study of the urban boundary layer using different urban parameterizations and high-resolution urban canopy parameters with WRF. *J. Appl. Meteorol. Climatol.* 50 (5), 1107–1128.

- Sobel, I., Feldman, G., 1968. A  $3 \times 3$  Isotropic Gradient Operator for Image Processing. presented at the Stanford Artificial Intelligence Project (SAIL) in .
- Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. *Bull. Am. Meteorol. Soc.* 93 (12), 1879–1900.
- Steyn, D.G., 1980. The calculation of view factors from fisheye-lens photographs: research note. *Atmosphere-Ocean* 18, 254–258.
- Steyn, D.G., Hay, J., Watson, I.D., Johnson, G.T., 1986. The determination of sky view-factors in urban environments using video imagery. *J. Atmos. Ocean. Technol.* 3 (4), 759–764.
- Stone, B., 2008. Urban sprawl and air quality in large US cities. *J. Environ. Manag.* 86 (4), 688–698.
- Svensson, M.K., 2004. Sky view factor analysis—implications for urban air temperature differences. *Meteorol. Appl.* 11 (3), 201–211.
- Unger, János, 2004. Intra-urban relationship between surface geometry and urban heat island: review and new approach. *Clim. Res.* 27, 253–264.
- Unger, J., 2009. Connection between urban heat island and sky view factor approximated by a software tool on a 3D urban database. *Int. J. Environ. Pollut.* 36 (1–3), 59–80.
- Verdonck, M.-L., Okujeni, A., van der Linden, S., Demuzere, M., De Wulf, R., Van Coillie, F., 2017. Influence of neighbourhood information on 'local climate zone' mapping in heterogeneous cities. *Int. J. Appl. Earth Obs. Geoinf.* 62. <http://dx.doi.org/10.1016/j.jag.2017.05.017>.
- Wang, C., Middel, A., Myint, S.W., Kaplan, S., Brazel, A.J., Lukaszczuk, J., 2018. Assessing local climate zones in arid cities: the case of Phoenix, Arizona and Las Vegas, Nevada. *ISPRS J. Photogramm. Remote. Sens.* 141, 59–71. <http://dx.doi.org/10.1016/j.isprsjprs.2018.04.009>. <http://www.sciencedirect.com/science/article/pii/S0924271618301217> (ISSN 0924-2716, Keywords: Local climate zone; Land use land cover; Sky view factor; Land surface temperature; Phoenix; Las Vegas).
- Xu, Y., Ren, C., Ma, P., Ho, J., Wang, W., Lau, K.K.-L., Lin, H., Ng, E., 2017. Urban morphology detection and computation for urban climate research. *Landsc. Urban Plan.* 167, 212–224.
- Zakšek, Klemen, Oštir, Kristof, Kokalj, Žiga, 2011. Sky-view factor as a relief visualization technique. *Remote Sens.* 3 (2), 398–415.
- Zhang, Q., Li, B., Thau, D., Moore, R., 2015. Building a better urban picture: combining day and night remote sensing imagery. *Remote Sens.* 7 (9), 11887–11913.