Characterising the vertical structure of buildings in cities for use in atmospheric models

Stretton, M. A. ORCID: https://orcid.org/0000-0002-1444-5735, Hogan, R. J. ORCID: https://orcid.org/0000-0002-3180-5157, Grimmond, S. ORCID: https://orcid.org/0000-0002-3166-9415 and Morrison, W. (2023) Characterising the vertical structure of buildings in cities for use in atmospheric models. Urban Climate, 50. 101560. ISSN 2212-0955 doi: 10.1016/j.uclim.2023.101560 Available at https://centaur.reading.ac.uk/112033/

Urban Climate xxx (xxxx) xxx

Contents lists available at ScienceDirect

Urban Climate

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

Characterising the vertical structure of buildings in cities for use in atmospheric models

Megan A. Stretton [a,*], Robin J. Hogan [a,b], Sue Grimmond [a], William Morrison [a,c]

[a] Department of Meteorology, University of Reading, Reading, UK
[b] European Centre for Medium-Range Weather Forecasts, Reading, UK
[c] Chair of Environmental Meteorology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

ARTICLE INFO

Keywords:
Urban form
SPARTACUS-urban model
Shortwave radiation

ABSTRACT

Urban schemes for numerical weather prediction (NWP) often assume an infinite street canyon with constant height and width, impacting turbulent and radiative fluxes. We develop parameterisations for urban morphology profiles, with five complexity levels, using data from six cities at 2 km × 2 km resolution. Comparisons of parameterised building plan area to these 'true' data show that 90% of building fraction profiles have bias errors (BE) at any height of <0.03. An effective building diameter (D) is used to characterise the proportionality between building plan area and building normalised perimeter length. The six-city mean D is 21 m. Relations for D have normalised BE (nBE) < 16%, increasing to 26% when total wall area is assumed to be unknown. Impacts from using these new morphology relations are tested with SPARTACUS-Urban radiative transfer simulations. The effective shortwave albedo has a nBE 2–10% (cf. 'true'). Within-canyon absorption have larger nBEs, suggesting the bulk albedo hides within-canopy errors. Overall, nBE increase as less morphology data are provided, notably when omitting total wall area. We conclude that urban vertical variability using the proposed relations are acceptable for NWP, requiring only: surface building plan area, mean building height, and effective building diameter.

1. Introduction

High population densities in cities can expose large numbers of people to extreme weather events such as heatwaves, whose effects can be exacerbated by poor air quality and the impacts of climate change. With urban areas being home to 54% of the world's population and rising (> 60% by 2030) (United Nations, 2018), these extreme events threaten public health and may cause severe economic loss. Because of this, there is a need for increasing accuracy and spatial resolution in weather forecasting and climate projections for urban areas (Grimmond et al., 2020).

Cities contain a heterogenous mix of buildings and trees with varying heights and densities. The material properties vary between and within the impervious (e.g., buildings, roads, sidewalks) and pervious (vegetated) areas. These characteristics influence the absorption and loss of shortwave and longwave radiation, while also impacting ventilation within the urban canopy (Guo et al., 2016; Kent et al., 2019). These effects cause trapping of heat (impacting local temperatures), influence the surface energy balance, and can affect precipitation and thunderstorm intensities (Collier, 2006; Liang, 2018; Shepherd, 2005).

Therefore, the 3D structure of cities should be treated with appropriate assumptions in order to understand the impacts of urban

* Corresponding author.
E-mail address: m.a.stretton@reading.ac.uk (M.A. Stretton).

https://doi.org/10.1016/j.uclim.2023.101560
Received 26 September 2022; Received in revised form 21 March 2023; Accepted 16 May 2023

areas on the shortwave and longwave radiation (Arnfield, 1982; Masson, 2000; Martilli et al., 2003). Within numerical weather prediction (NWP) models, surface obstacles (e.g., buildings, vegetation) and the interactions between them need to be parameterised, due to computational and data constraints (among others) that limit surface schemes. NWP urban morphology often assumes an 'urban canyon', consisting of flat roofed buildings of the same height with constant width, that are characterised by the building height to canyon width (H/W) ratio (e.g. Arnfield and Grimmond, 1998; Kusaka et al., 2001; Porson et al., 2010).

Morphometric parameters are key for accurate modelling of urban weather and climate, and include: building heights (Masson, 2020; Zhu et al., 2019) and their probability distribution (e.g., Temperatures of Urban Facets (TUF), Krayenhoff et al. (2014)), building plan area, wall area, and H/W (e.g., Town Energy Balance (TEB), Masson (2000); Lemonsu et al. (2012)) with vertical structure of cities additionally vital (Wentz et al., 2018). One reason why urban NWP schemes remain simple is a lack of global data to describe this vertical structure of cities. If global data for all buildings were available, the required parameters for modelling (e.g., building fraction, wall area) could be derived at any NWP grid resolution without parameterisations.

An additional reason for NWP urban schemes simplicity is a lack of accurate but computationally efficient methods to represent vertically resolved energy exchanges within urban canopies. Model developments, such as the SPARTACUS-Urban radiation scheme (Hogan, 2019a), the Building Effect Parameterization scheme for turbulent fluxes (Martilli et al., 2002; Salamanca et al., 2011; Schubert et al., 2012), Seoul National University Urban Canopy Model (Ryu and Baik, 2012; Ryu et al., 2013), and NJUC-UM-M (Kondo et al., 2005), have potential use in NWP. These consider a multi-layer approach, resolving multi-layer fluxes within the urban canyon, taking into account varying building heights and often vegetation.

Often building footprints, sometimes with height information, are available from municipal sources. Airborne stereophotography, photogrammetry or LiDAR data are used to develop digital elevation models (DEMs) and digital surface models (DSMs) (Gamba and Houshmand, 2002; Xu et al., 2017). These allow high resolution (< 1 m) characterisation of individual cities (Gage and Cooper, 2017; Goodwin et al., 2009; Holland et al., 2008; Lindberg et al., 2011; Lindberg and Grimmond, 2011). However, these data are unavailable for all cities worldwide due to the expense of data collection, storage and processing (Kent et al., 2019). Community volunteer projects (e.g., OpenStreetMap) collect building footprints, which may also have building height attributes provided (e.g., Microsoft providing U.S. cities, Heris et al. (2020)), or derived from other new techniques (e.g., using surrounding building information (Bernard et al., 2022)).

Building heights can be derived from remote sensing observations (Rao, 1972; Champeaux et al., 2005) but few satellite missions have both the coverage and sufficient resolution to provide individual building heights for large areas (Frantz et al., 2021). Datasets have also been derived from combining multiple satellite sources (e.g., Frantz et al. (2021)), and additionally with local building footprint data (e.g., Milojevic-Dupont et al., 2020), but studies such as these may only cover a single continent, or a few cities, and/or rely on the availability of open-source building footprints.

Across the range of current morphology dataset creation studies, methodologies are inconsistent and need to be standardised before building information (e.g., height) can be provided across large areas. Hence, providing parameterisations that use publicly available data to estimate the vertical structure of cities at resolutions suitable for NWP would be beneficial, until better datasets become available globally. Such relations have been developed for urban areas (e.g., plan area fraction of buildings, mean building height, frontal area index) and are used operationally, e.g., UK Met Office use the Bohnenstengel et al. (2011) within the UKV (Hertwig et al., 2020). These parameterisations often require inputs that are not widely available (e.g., urban land cover), are derived from a single city but applied globally, and/or assume very limited variability e.g., one mean building height per built ('urban') fraction with saturation (i.e., tall dense city centres are the same as other areas). Intra-city urban form variability has been included using the Stewart and Oke (2012) local climate zones (LCZ) which provide representative ranges of values of parameters (e.g. plan area fraction of buildings or paved surfaces, roughness element height, and frontal area index) for each class, with datasets available to use as urban model inputs (Demuzere et al., 2022), and integration with models, e.g., within the Weather Research and Forecasting model with a single layer urban canopy model (Brousse et al., 2016; Molnár et al., 2019) with analysis on urban air temperatures.

Given the importance of radiation to surface energy exchange, this work, and multi-layer modelling generally, is motivated by the need to reduce the sources of error in urban radiation calculations, which are from:

(1) the radiation scheme, even if the urban morphology is known exactly
(2) approximating the morphology from a few parameters (e.g. plan area or building cover fraction, mean building height and total building wall area)
(3) incomplete knowledge of the parameters in any given city.

Stretton et al. (2022) address the first of these, evaluating the SPARTACUS-Urban shortwave radiation scheme. They demonstrate that SPARTACUS-Urban accurately predicts profiles of absorption into urban facets (mean absolute error $< 16\%$), and the effective albedo at the top of the canopy (normalised bias error $< 6\%$) for real-world scenes. Overall, errors for all variables are largest when the sun is low in the sky, when the impact of the underlying assumptions on building geometry are largest.

The specific objective of this paper is to address (2) by identifying and parameterising key vertical profiles of the urban form, using methods and coefficients that can be used globally, while retaining some of the realistic intra-city variability. Given the range of urban forms (within and between cities) and the sparsity of data, the relations developed here ideally need to be both simple and universal. We focus on input parameters including building: height, plan area fraction (with height), and wall area (with height), as they define the area of roof and walls exposed for energy exchange with height. Problem (3) is partially addressed here by giving a range of typical values of parameters for six cities. However, the parameterisations developed in this paper will become more applicable once datasets of building cover and mean building height are available globally.

The parameters to describe the urban form are selected and defined (section 2) and used within this study's various methods (section 3). The proposed parameterisations are assessed with 'true' urban morphology (section 4), combined (section 4.4), and used with SPARTACUS-Urban (Hogan, 2019b; Stretton et al., 2022) to simulate absorbed shortwave radiation into the three facets (walls, roof, and ground) (section 5). The conclusions drawn are given in section 6.

2. Definitions of multi-layer urban form parameters

2.1. Plan area fraction (λ_p)

The plan area fraction of buildings (λ_p) is the ratio of the total area covered by buildings to the 'grid-cell' or total horizontal area of interest. Typically, λ_p is assumed to be constant with height, ($\lambda_p(z)$), from the surface ($\lambda_p(z=0)$), unlike real buildings which rarely have equal heights in an area or an individual model grid-cell. Whilst we often model arrays of regular cubes (Kanda et al., 2005a; Kanda et al., 2005b; Morrison et al., 2018; Stretton et al., 2022)). Rather, $\lambda_p(z)$ varies (Fig. 1a) resulting in variations in sunlit and shaded surfaces (e.g., roofs) at any given height, allowing interception of radiation reflected from higher surfaces, increasing the radiation trapping with a canopy.

$\lambda_p(z)$ can be related to $\lambda_p(z=0)$ if the area-weighted mean building height ($\bar{H}$) is known (e.g., determined from building data in real cities), assuming:

$$\lambda_p(z) = \lambda_p(z=0) y(z/\bar{H}), \tag{1}$$

where $y(z/\bar{H})$ is a universal function with the properties: $y(0)=1$, $y(\infty)=0$, and the vertical integral of y must equal 1. A functional form for y is proposed in section 3.2.

2.2. Wall area: building perimeter length (L)

The wall area relates to the normalised building perimeter length, L, at any given height ($L(z)$) (Stretton et al., 2022). L is the total building perimeter (m) normalised by the total area of the grid-cell (m^2). Thus, if the canopy is divided into n layers and layer i has thickness Δz_i, then:

$$\lambda_W = \sum_i^n L_i \Delta z_i, \tag{2}$$

with λ_W the total wall area divided by the grid-cell area (Masson et al., 2020). If all wall orientations are assumed to be equally probable within a grid-cell, then $\lambda_w = \lambda_f \pi$ (Hogan and Shonk, 2013), where λ_f is the frontal area index – the total projected wall area for a particular wind (or azimuth) direction, relative to the grid-cell area. λ_f is often used to parameterise aerodynamic drag of roughness elements (Raupach, 1992; Grimmond and Oke, 1999; Sützl et al., 2020).

2.3. Effective building diameter (D)

Assuming both narrow and wide buildings have an equal probability of extending to any height (Fig. 1a), L is proportional to λ_p. We use this to define an effective building diameter (D) that is independent of height:

$$L(z) = \frac{4\lambda_p(z)}{D}. \tag{3}$$

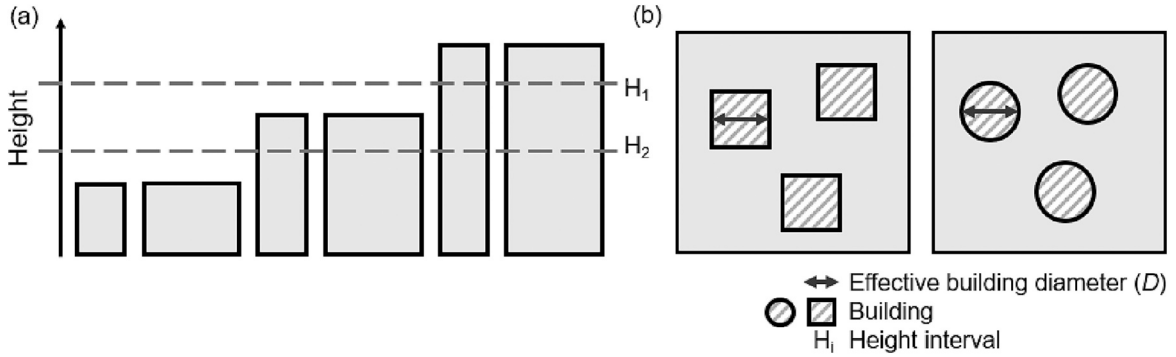

Fig. 1. Effective building diameter (D): (a) assumption it is constant with height (Eq. 3) implies an equal probability of both large and small buildings extending to different heights (cross section); and (b) plan view of two equivalent areas with buildings that are either cuboids of width D or cylinders of diameter D.

This is analogous to the relations used between perimeter length and area of clouds and trees (Jensen et al., 2008; Hogan et al., 2018). The D parameter can be thought of as the diameter (or width) of buildings in an equivalent idealized city, with the same properties (λ_p and L) as a real city (Fig. 1b), where all buildings are identical cylinders (or cubes). Note that we neither assume buildings have a particular shape, nor that they are all the same size in a real city. Rather, D quantifies the assumed proportionality between L and λ_p in Eq. 3, while having a simple physical interpretation.

3. Methods

3.1. Deriving $\lambda_p(z)$ and $L(z)$ from high resolution reference data

Profiles of λ_p and L are derived from high resolution reference data from six cities with differing characteristics (e.g., morphology, city layout) (Table 1). The cities chosen are Auckland, New Zealand; Berlin, Germany; Birmingham, UK; London, UK; New York City (NYC), US; and Sao Paulo, Brazil. Multiple global cities are used to capture the variation in building and land-use types seen exhibited across urban areas.

For London, Berlin, and Birmingham, building footprints (Umweltatlas Berlin, 2010; EMU Analytics, 2018)[1] are used to create a 1 m × 1 m raster with mean height assigned to each building. Therefore, buildings have flat roofs and non-tapering vertical walls without building features (e.g., pitched roofs) above the mean height. During rasterization, "false walls" between adjoining (e.g., terraced) and overlapping buildings are removed. Building information for Auckland, NYC, and Sao Paulo are LiDAR datasets provided as 4 m × 4 m raster (Kent et al., 2019). Although some differences in building morphology may arise from the raster resolution chosen (Fig. SM 1, Table SM 1), by using the highest resolution available (i.e., not coarsening to the 4 m × 4 m resolution of some datasets) allows the parameterisations to be closer to 'true' data. Buildings/pixels with height values below 2.5 m (mean storey height in the UK (OPDC, 2018)) or without height information are discarded for consistency across all datasets. For all grid-cells in all cities, the topography is assumed to be flat.

Each city dataset is split into 2 km × 2 km grid-cells, to have a similar resolution to operational limited-area NWP models (e.g. Met Office 1.5 km UKV model (Tang et al., 2013), DWD 2.8 km COSMO-DE model (Baldauf et al., 2011)). At this scale, most grid-cells will have many buildings with different heights, but are less likely to contain multiple neighbourhoods with very different characteristics. Grid-cells with incomplete (i.e., missing) building data (determined from visual inspection of aerial imagery) and/or any areas with extremely small building coverage ($\lambda_p(z = 0) < 0.001$) are removed. Profiles are calculated using a height interval (Δz) of 0.5 m.

3.2. Deriving globally obtainable parameterisations for $\lambda_p(z)$ and $L(z)$

The profiles of λ_p and L are parameterised for the six cities (Table 1). The parameterisations developed for $\lambda_p(z)$ and $L(z)$ are used in five combinations of increasing complexity, with differing input data requirements. These are evaluated against the high-resolution reference datasets (P0). The parameterisations (P#) of $\lambda_p(z)$ and $L(z)$ are given increasing numbers (#, 1 → 5) as the input requirements decrease (i.e., with the intention they are more globally applicable) (Table 2):

Table 1
Cities and high-resolution reference data used to evaluate urban form parameterisations. City area (A_{city}) from Demographia World Urban Areas (2020), US Gazetteer (2021) as basis for the whole city area, an area can be >100% if it contains areas outside of the city, or areas of water. The 2 km × 2 km parameters are derived from either raster (*rDSM*) or vector (*vDSM*) digital surface models with the horizontal resolution (Δx) indicated.

City	City		Whole city mean building			Data			
	Number of grid-cells	Fraction of A_{city}	Edge length $L(z = 0)$ (m^{-1})	Surface fraction $\lambda_p(z = 0)$	Building height $\bar{H}$ (m)	Date	Type	Δx (m)	Source
Auckland	59	0.45	0.051	0.21	5.98	Jul-Nov 2013	rDSM	4	Kent et al. (2019)
Berlin	198	0.58	0.022	0.10	8.86	2009–2010	vDSM	1	Umweltatlas Berlin (2010)
Birmingham	177	1.18	0.020	0.082	5.87	Jan-1998 – Sep-	vDSM	1	EMU Analytics (2018)
London	415	0.96	0.029	0.13	7.32	2014	vDSM	1	EMU Analytics (2018)
New York City	251	1.29	0.033	0.17	12.68	2014	rDSM	4	Kent et al. (2019)
Sao Paulo	329	0.41	0.046	0.24	5.97	2007	rDSM	4	Kent et al. (2019)

[1] EMU Analytics data combine building polygons from Ordnance Survey Open Map and heights from Environment Agency LiDAR data from 2014 and 2015.

Table 2
Inputs required for parameterisations (P1 → P5) to determine building fraction (λ_p) and normalised building edge length (L) from D (Fig. 1b) (assumed constant with height). These are evaluated using the actual 'true' profiles of λ_p and L (P0, section 3.1), mean building height, $\bar{H}$, and building fraction at the surface, $\lambda_p(z = 0)$, are required inputs for all parameterisations.

Code	Inputs required	$\lambda_p(z)$ parameterisation	$L(z)$ parameterisation
P0	$\lambda_p(z)$ and $L(z)$ profiles	Known	Known
P1	$\lambda_p(z)$ profile; λ_w; $\bar{H}$	Known	Wall area conserved D: Eq. 5
P2	$\lambda_p(z = 0)$; λ_w; $\bar{H}$	Variable b: Eq. 1 & 4 with $b(\bar{H})$	Wall area conserved D: Eq. 5
P3	$\lambda_p(z = 0)$; λ_w; $\bar{H}$	Fixed b: Eq. 1 & 4 with $b = 4.7$	Wall area conserved D: Eq. 5
P4	$\lambda_p(z = 0)$; $\bar{H}$	Variable b: Eq. 1 & 4 with $b(\bar{H})$	Linear-fit D: Eq. 6: $D[\lambda_p(z = 0), \bar{H}]$
P5	$\lambda_p(z = 0)$; $\bar{H}$	Fixed b: Eq. 1 & 4 with $b = 4.7$	Fixed D: $D = 20.93$ m

P1 the most data demanding (e.g., from high-resolution raster data, Section 3.1) parameterisation, uses non-globally available (yet) λ_W and $\lambda_p(z)$. An effective building diameter, D, is used to predict $L(z)$ to ensure the latter satisfies Eq. 2 (Table 2). This can be achieved if λ_w is known, by:

$$D = \frac{4}{\lambda_W} \sum_{z=0}^{i} \lambda_{p,i} \Delta z = \frac{4}{\lambda_W} \mathrm{V}, \tag{4}$$

where the normalised building volume, V, is the ratio of building volume and grid-cell area, related to building plan area and mean building height ($\bar{H}$) by $V = \lambda_p(z = 0)\bar{H}$ (parameterisation: *Wall area conserved (CWA) D*, Table 2).

P2 uses λ_W as in P1 but parameterises $\lambda_p(z)$ from $\lambda_p(z = 0)$, and mean building height ($\bar{H}$) using Eq. 1, with:

$$y(x) = \left[1 + (ax)^b\right]^{-1}, \tag{5}$$

where $x = z/\bar{H}$, a is a function of b that ensures the integral of y from zero to infinity is one. Fig. SM 2 demonstrates that this functional form is a good fit to the median building-fraction profile from real cities worldwide. Value of the best-fit parameter, b, derived from the high-resolution reference data have smaller values when there is larger variation in building heights within a grid-cell. To determine b, we take $\lambda_p(z)$, and normalise each axis by $\bar{H}$ and $\lambda_p(z = 0)$, such that the resultant curve represents the $y(x)$ function from Eq. 4 (Fig. SM 2). The normalised curves are interpolated using a common vertical normalised height interval of 0.05, prior to determining the median profile (Fig. SM 2a, b).

For the $\lambda_p(z)$ *variable b* parameterisation (Table 2), grid-cells are categorised in $\bar{H}$ intervals. Values of b are derived using median normalised profiles for each $\bar{H}$ interval calculated from all cities.

P3 is P2 but for $\lambda_p(z)$ the *fixed b* parameterisation (Table 2) is used with $b = 4.7$ (Eq. 4) across all $\bar{H}$ for all grid-cells. This value is derived from the multi-city median $y(x)$. Using one b value allows assessment of cities similarity, and if parameterisations perform better (less error in radiation fluxes) if more data are included.

P4 requires only grid-cell values of $\lambda_p(z = 0)$ and $\bar{H}$. $\lambda_p(z)$ is obtained as in P2 and derives L from calculating an effective building diameter (D) from:

$$D = p\bar{H} + q\lambda_p(\mathrm{z} = 0) + r, \tag{6}$$

where the constants p, q, and r ($p = 0.847$, $q = 5.17$, $r = 11.96$) are derived by fitting D to $\lambda_p(z = 0)$ and $\bar{H}$ across all cities (Table 1, Table 2). This is referred to as *linear-fit D* (Table 2).

P5 the simplest case uses only $\lambda_p(z = 0)$ and $\bar{H}$, so requires the least data. It parameterises $\lambda_p(z)$ as in P3. To obtain L, with D set to 20.93 m for all cities (*fixed D*, Table 2) and λ_w is assumed proportional to V (Eq. 4).

All the 2 km × 2 km grid-cells available are used to derive parameters for the $L(z)$ and *fixed b* parameterisations. For the *variable b* parameterisations (P2, P4), for each height interval a bootstrapped random sample (Padiyedath Gopalan et al., 2019) of 1000 is used to ensure all intervals have the same sample size.

3.3. SPARTACUS-urban radiative transfer model

The original SPARTACUS (Speedy Algorithm for Radiative Transfer through Cloud Sides) 3D radiative exchange model for complex cloud fields simulates lateral radiative exchange between clear and cloudy regions in proportion to the cloud edge length per unit area of a NWP grid-cell, using cloud fraction and cloud edge length (Hogan et al., 2016). A similar approach has been applied to forest vegetation (Hogan et al., 2018) and city buildings (Hogan, 2019b). All three assume any obstacles to radiation are randomly distributed within a horizontal plane, allowing the mean radiation field to be modelled as a function of height. The open-source software SPARTACUS-Surface combines SPARTACUS-Urban and SPARTACUS-Vegetation. As we focus only on buildings, we refer to it as SPARTACUS-Urban.

SPARTACUS-Urban is underpinned by the 1D discrete-ordinate method, solving coupled ordinary differential equations for 2 N

streams of radiation, with N streams per hemisphere. The radiation field is described more accurately as N increases, but with added computational cost. In this work, 16 streams are used (i.e., $N = 8$). A scene (any combination of building geometry, solar zenith angle, and albedo) is split into n layers to calculate radiative interactions per level. Each layer has regions of clear-air and buildings. SPARTACUS-Urban computes the radiative interactions between the three facets (wall, roof, and ground) using the vertical profiles of L and λ_p to characterise the urban form with z within a grid-cell. If vegetation is included, it also needs to be accounted for in each layer.

To compute radiative fluxes, SPARTACUS-Urban uses a prescribed albedo (α). Here, we assign 0.2 to all facets (wall, roof, ground), based on literature for a range of urban form synthesised by Oke et al. (2017): 0.08–0.25 for roofs, 0.2–0.25 for walls and 0.08–0.25 for the ground (Grimmond and Oke, 1995; Feigenwinter et al., 1999; Grimmond et al., 2004; Lemonsu et al., 2004; Offerle et al., 2005; Balogun et al., 2009; Bergeron and Strachan, 2012). This is constant between facets to isolate any radiative errors due to the changes in morphology. SPARTACUS-Urban treats determines both direct and diffuse radiation, but does not separate sunlit and shaded surfaces. For this study, we assume all incoming radiation is direct radiation (i.e., no diffuse) with the normal solar flux at the top of the canopy set to 1000 W m^{-2}. We model for three solar zenith angles (θ_0): overhead (nadir) sun (0°), 45°, and low-sun (75°). Under overhead sun ($\theta_0 = 0°$) we expect fluxes to be sensitive primarily to differences in the λ_p parameterisations (P2, P3), and for low sun simulations ($\theta_0 = 75°$) to respond to differences in L parameterisations (P4, P5).

3.4. Metrics to evaluate parameterisations against high resolution reference datasets

The parameterisations (P1 → P5, Table 2) are evaluated using the bias error ($BE_i = P\#_i - P0_i$) between the 'true' (P0) and parameterised profiles for building fraction and building edge length. The BE of individual λ_p and L vertical profiles across each city are analysed using the median, mean, and 5th - 95th percentiles.

Similarly, the λ_w for L parameterisations is evaluated using the mean BE (MBE). The mean λ_w is also evaluated using a normalised MBE (nMBE), computed by dividing the MBE by the 'true' mean λ_w computed from the P0 data, and multiplying by 100 to give a percentage. The mean absolute error ($MAE = \Sigma(|P\#_i - P0_i|)/n$) for λ_w for each parameterisation is calculated for the number of grid-cells per city (n), to examine the performance across all data.

The impact of the parameterisations to the vertically integrated shortwave absorption into the walls (a_{Wall}), roof (a_{Roof}) and ground (a_{Ground}) per unit area of the entire horizontal grid-cell (W m^{-2}) is assessed using the MBE, nMBE, and MAE as above. The total absorption is also assessed for each grid-cell using the normalised BE (nBE). Similarly, we examine the impact on the shortwave bulk albedo for each grid-cell as above.

4. Results and discussion

4.1. Spatial variation of morphology in high resolution reference datasets

For all six cities, the largest values of both plan are fraction at the surface ($\lambda_p(z = 0)$) and mean height ($\bar{H}$) are found near the city centre or central business district (CBD) (Fig. 2). The largest $\bar{H}$ values occur in NYC (e.g., Manhattan) (~ 30 m for 2 km × 2 km grid-

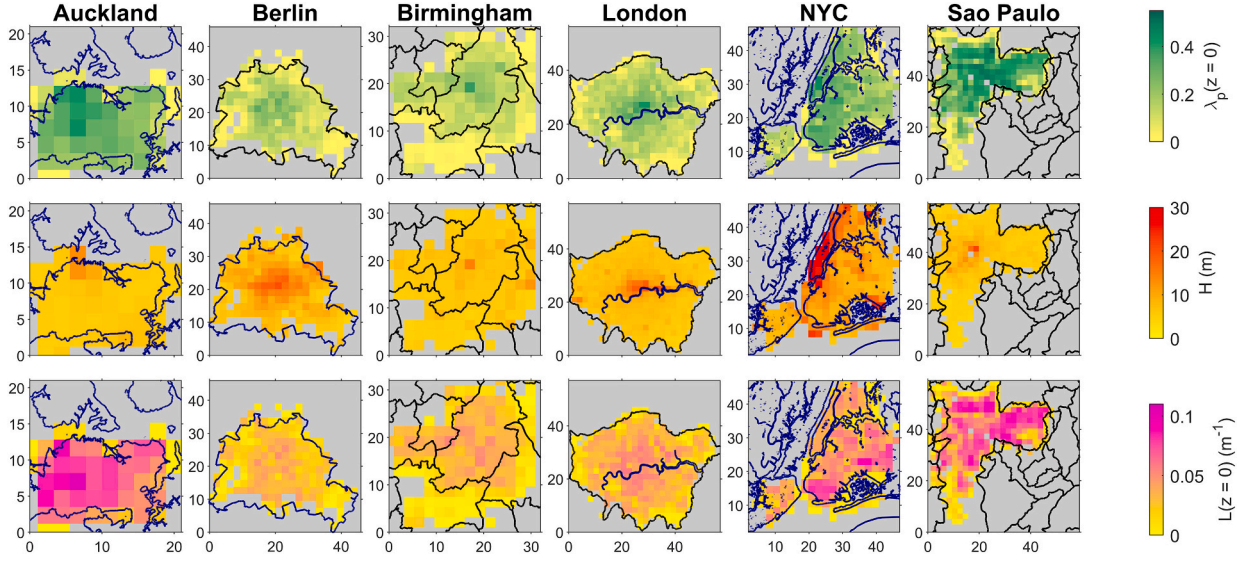

Fig. 2. Parameters derived from building data (Table 1) at 2 km × 2 km resolution for (columns) six cities: (row 1) building fraction at the surface ($\lambda_p(z = 0)$), (row 2) mean building height ($\bar{H}$), and (row 3) normalised building perimeter length at the surface ($L(z = 0)$) (Eq. 3); with city boundaries (black) and water bodies (blue) shown. Data sources are given in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cells) and Berlin. The lower values in other cities may be due to large variations in building types within a grid-cell, e.g., in Sao Paulo very tall buildings are often surrounded by areas of smaller buildings within the same grid-cells. Values of $\lambda_p(z = 0)$ are higher in NYC, Sao Paulo, and Auckland. Similar values of $L(z = 0)$ are found across each individual city, lowest in Berlin, Birmingham, and London ($<$ 0.05 m^{-1}) and largest in Auckland, NYC, and Sao Paulo (0.05–0.1 m^{-1}). These could be impacted by fraction of the whole city analysed and the definitions used (Table 1), as this varies between 45% and $>$ 100%. For example, in Auckland, a smaller fraction (0.45) of the city is analysed, so likely less variability in land-use and building type are characterised. However, each city has grid-cells that contain parks, residential areas, and parts of the CBD. Although only six cities are analysed, they cover a wide range of λ_p and $\bar{H}$ combinations, city layouts, locations, employment types/industries, and demographic characteristics. However, as other cities are analysed the result details will vary.

Across all cities the 2 km $\times$ 2 km grid-cells generally have a $\bar{H}$ of $<$30 m (colour bar, Fig. 3). The few grid-cells with $\bar{H} > 40$ m often have larger $\lambda_p(z = 0)$ values. This suggests that grid-cells are not dominated by a large number of tall buildings, although this depends on the exact grid-cell location. Effective building diameter (D) values are always $>$8 m but mostly $<$40 m (Fig. 3). Values of D tends to increase with $\lambda_p(z = 0)$ (between 10 and 20 m, when surface building fractions are 0.0 to 0.5) in grid-cells with low values of $\bar{H}$ ($<$ 10 m) (Fig. 3). Grid-cells with large values of D (up to 100 m) tend to have smaller $\lambda_p(z = 0)$.

4.2. Evaluation of parameterisations for normalised perimeter length (L)

We assess if the assumption that the effective building diameter (D) is constant with height (z) (Section 3.2) is appropriate. Within and between cities, D varies with z (Fig. 4). The mean and median D values across each city appear to follow similar vertical relations, with the mean D at each height larger than the median, especially with increasing height. In Birmingham, London, and NYC, D is approximately constant with height between 7 and 10 m, and at ~25 m (Birmingham and London), and ~ 35 m (NYC). Similar behaviour is seen in Berlin, although the mean and median D values are much closer in magnitude. In Sao Paulo, D is roughly constant with height above 20 m. In Auckland, D decreases with height, with $D > 20$ m until $z > 50$ m; but Auckland has both the fewest grid-cells (only 59) and smallest fraction of the city analysed (Table 1).

As the vertical variation of D is much smaller than the vertical variation of plan area fraction of buildings (λ_p, Fig. 4), we test the consequence of assuming D is constant with height. In four of the six cities, wider buildings tend to extend higher into the urban canopy (not Auckland or Sao Paulo). Auckland appears to be an outlier (cf. other cities), possibly because of the smaller fraction of the city assessed (0.45, Table 1). In grid-cells with taller buildings (larger mean $\bar{H}$), the assumption of a constant D is appropriate, but could be replaced with values obtained from two (or more) height intervals to better characterise the vertical profile; however, this additional complexity is not introduced here.

Examining the normalised perimeter length at each height level ($L(z)$, Eq. 3, Table 2) parameterisations derived using the high-resolution reference dataset λ_p profile for each grid-cell, an increase in $L(z)$ bias error (BE) between the *linear-fit D* and *fixed D* parameterisations is evident (Fig. 5). In most cities there is a tendency to overestimate $L(z)$ higher in the canopy (Fig. 4, Fig. 5),

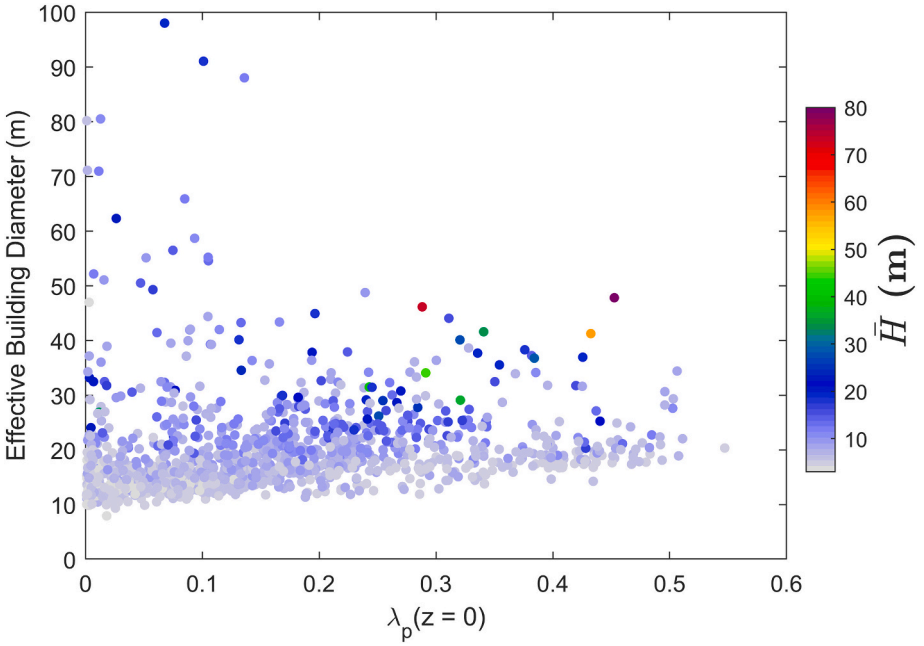

Fig. 3. Relation between the three parameters used in the 'linear-fit D' parameterisation (Table 2) for all grid-cells ($N = 1429$): building fraction at the surface ($\lambda_p(z = 0)$), effective building diameter (D) calculated using normalised wall area (λ_w) and mean building height ($\bar{H}$).

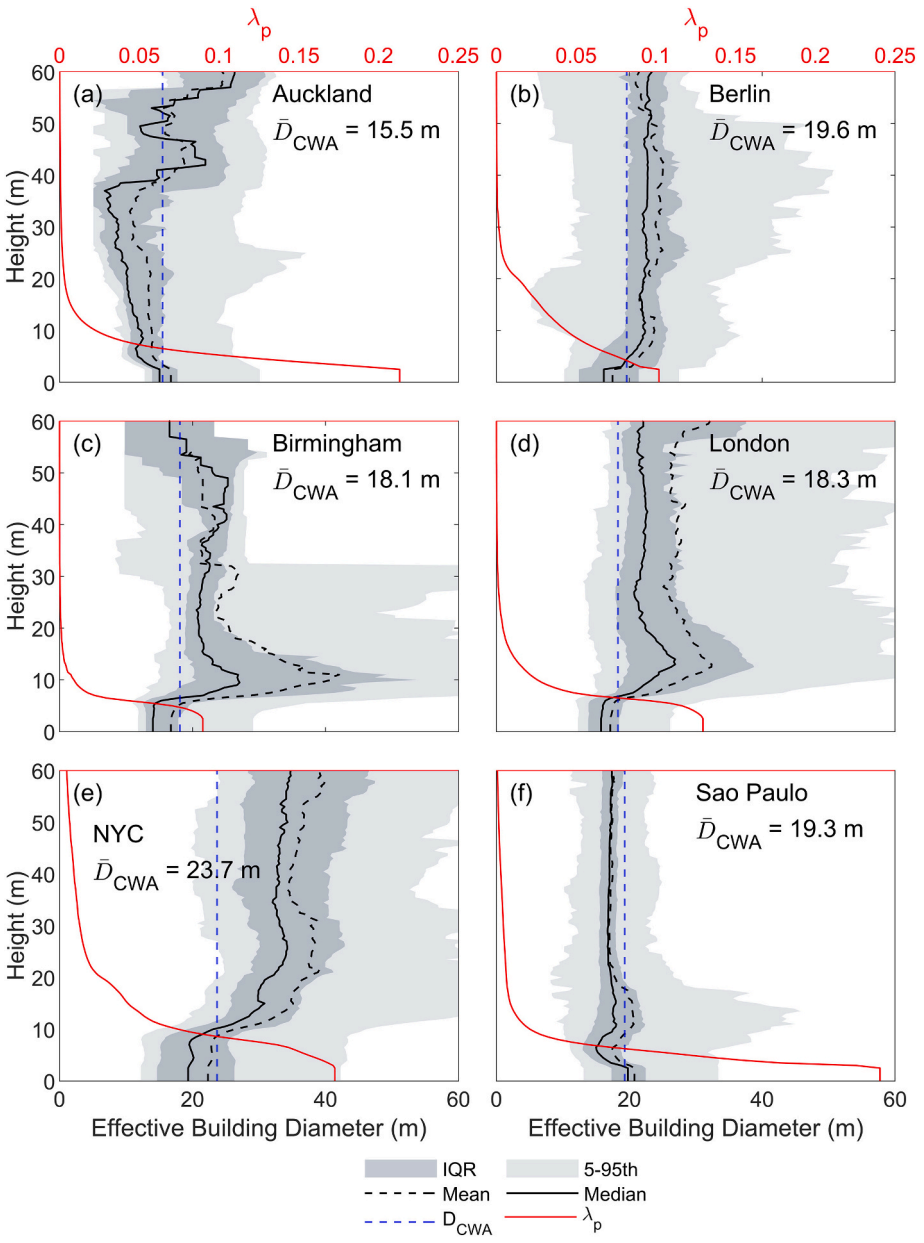

Fig. 4. Vertical profiles of effective building diameter (D) calculated from the true building data (P0, Table 2) [mean (dashed), median (solid), inter-quartile range (IQR, shading), 5th and 95th percentile (shading)], building fraction (λ_p) [red, mean weighted by number of grid-cells ($\lambda_p > 0.001$) per height level] and mean 'wall area conserved D' (CWA) parameterisation (blue dashed) for: (a) Auckland, (b) Berlin, (c) Birmingham, (d) London, (e) NYC, and (f) Sao Paulo. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

indicating an under-estimation in effective building diameter (i.e., from Eq. 3). Hence, buildings with a larger horizontal size are more likely to be taller.

The λ_w nMBE for the *linear-fit D* parameterisation are lowest in NYC (0.01%), and largest in Auckland (−15%) (Table 3). The *fixed D* parameterisation of λ_w have larger nMBE than the *linear-fit D* parameterisation, with values between −6.4% (Berlin) and − 26% (Auckland). Generally, all nMBE indicate λ_w is underestimated when using both *linear-fit D* and the *fixed D* parameterisations for $L(z)$ (Table 3). For Auckland, London, and Birmingham, the *linear-fit D* parameterisation performance is always better than the *fixed D* parameterisation. The distribution of BE in $L(z)$ (Fig. 5) for both the *linear-fit D* and *fixed D* parameterisations, indicate the median is <0.005 m^{-1} (mean $L(z = 0)$ spans 0.02–0.05 m^{-1}). Exceptions to this include the *fixed D* parameterisation in Auckland. The 5th – 95th percentiles of profiles (shading, Fig. 5) for Berlin, London, and Birmingham do not exceed 0.04 m^{-1} at any height but are slightly larger for the remaining cities.

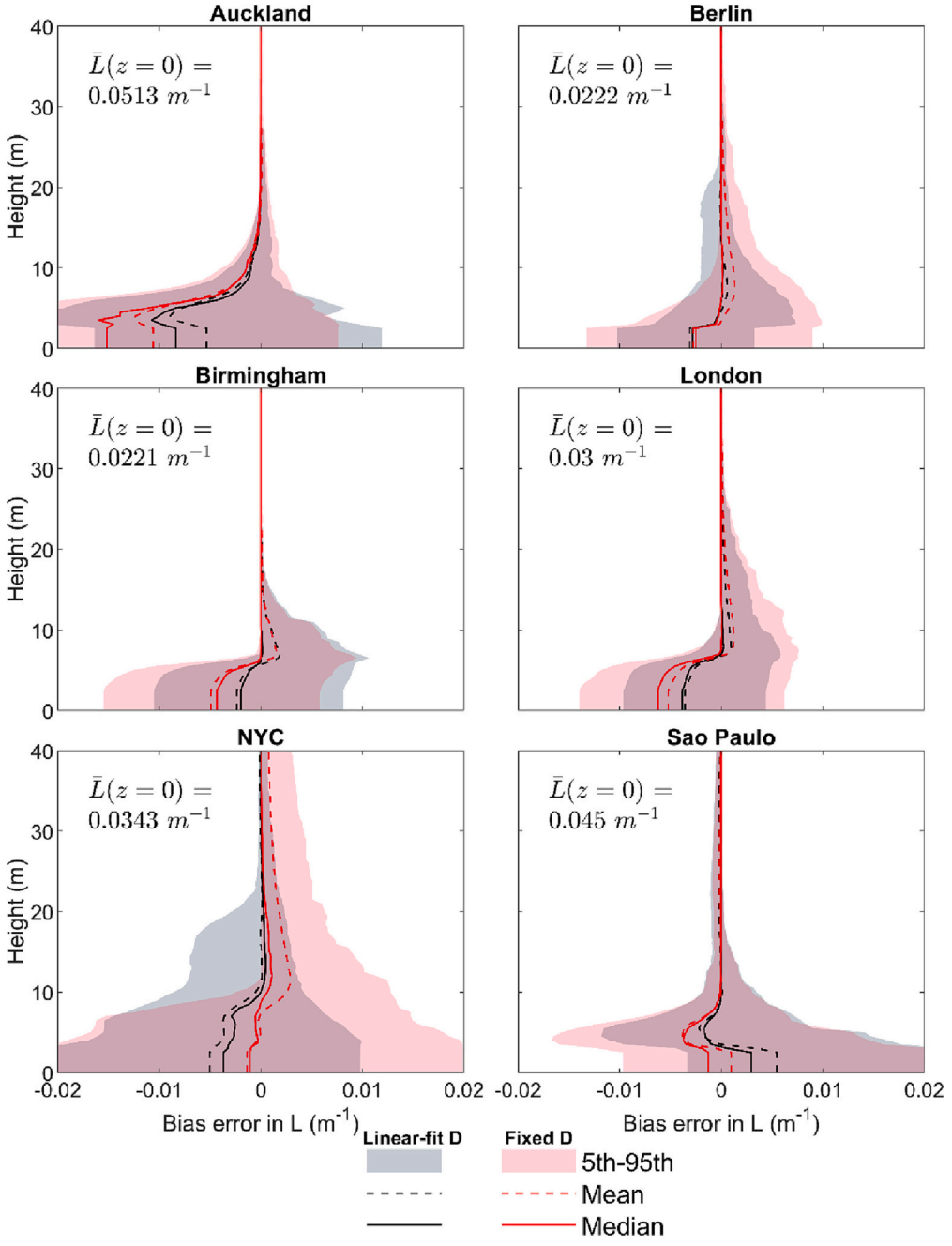

Fig. 5. Mean building edge length at the surface ($\bar{L(0)}$) and vertical profiles of bias error (BE, section 3.4) of normalised building edge length (L) determined using the 'linear-fit D' parameterisation (grey) and 'fixed D' parameterisation (red) (Table 2) for (a) Auckland, (b) Berlin, (c) Birmingham, (d) London, (e) NYC, and (f) Sao Paulo. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Both UK cities (London and Birmingham) have similar shaped BE profiles, suggesting a similar urban structure. In London, the L skill changes with height from an underestimation ($z < \sim 7$ m) to overestimation ($z > \sim 9$ m) (Fig. 4). This could result from the larger $\bar{H}$ variation across the city (Fig. 2), cf. Birmingham which has lower $\bar{H}$ with less variation, as the CBD covers fewer grid-cells. NYC has the largest $L(z)$ differences of the six cities, with the *fixed D* parameterisation median BE and MBE generally smaller than the *linear-fit D* parameterisation for $z \leq 10$ m. For all cities the largest $L(z)$ errors occur for $z < 10$ m, with most overestimating $L(z)$ for both parameterisations below 5 m.

Table 3
Evaluation of mean normalised wall area (λ_w) using the linear-fit D and fixed D parameterisations (Table 2) using the conserved wall area (CWA) D parameterisation for each city assessed with metrics (section 4.4): mean bias error (MBE), normalised mean bias error (nMBE, %) and mean absolute error (MAE). Table 1 gives number of grid-cells analysed per city.

City	*Conserved wall area D* $\bar{\lambda_w}$ (m^2/m^2)	Linear-fit D			Fixed D		
		MBE (m^{-1})	nMBE (%)	MAE (m^{-1})	MBE (m^{-1})	nMBE (%)	MAE (m^{-1})
Auckland	0.482	−0.054	−15	−0.054	−0.082	−26	−0.082
Birmingham	0.165	−0.0025	1.5	−0.0025	−0.018	−14	−0.018
Berlin	0.187	−0.0073	−3.3	−0.0073	0.0071	−6.4	0.0071
London	0.250	−0.011	−4.8	−0.011	−0.016	−13	−0.016
NYC	0.302	−0.043	0.013	−0.043	0.074	13	0.074
Sao Paulo	0.287	0.0061	5.8	0.0061	−0.015	−7.6	−0.015

4.3. *Evaluation of parameterisations for plan area fraction of buildings (λ_p)*

Parameterisation methods for λ_p (*variable b* and *fixed b*, Table 2), are evaluated using the $\lambda_p(z)$ vertical profiles from the high-resolution reference datasets (P0, Table 2) for each grid-cell (Fig. 2). For the *variable b* parameterisation, b varies with $\bar{H}$ intervals ($b = 2.1$–6.5, Fig. 6), whereas *fixed b* uses one multi-city value ($b = 4.7$) independent of city. As both parameterisations assume $\lambda_p(z = 0)$ is known, they agree with the 'truth' at the surface.

Generally, b decreases when mean building heights exceed 7 m (Fig. 6). The multi-city *fixed b* is approximately equivalent to the *variable b* parameterisation value of b at ~7 m. Grid-cells with the lowest $\bar{H}$ have the smallest variation in building heights and are associated with the largest b values. Areas with low $\bar{H}$ are often found in suburbs, dominated by two-storey buildings. This contrasts with the CBD, where a wider variation of building heights occurs, with lower values of b.

The variability around the median multi-city curve (black, Fig. 6) shows the variability in b is higher when $\bar{H}$ is larger (whiskers,

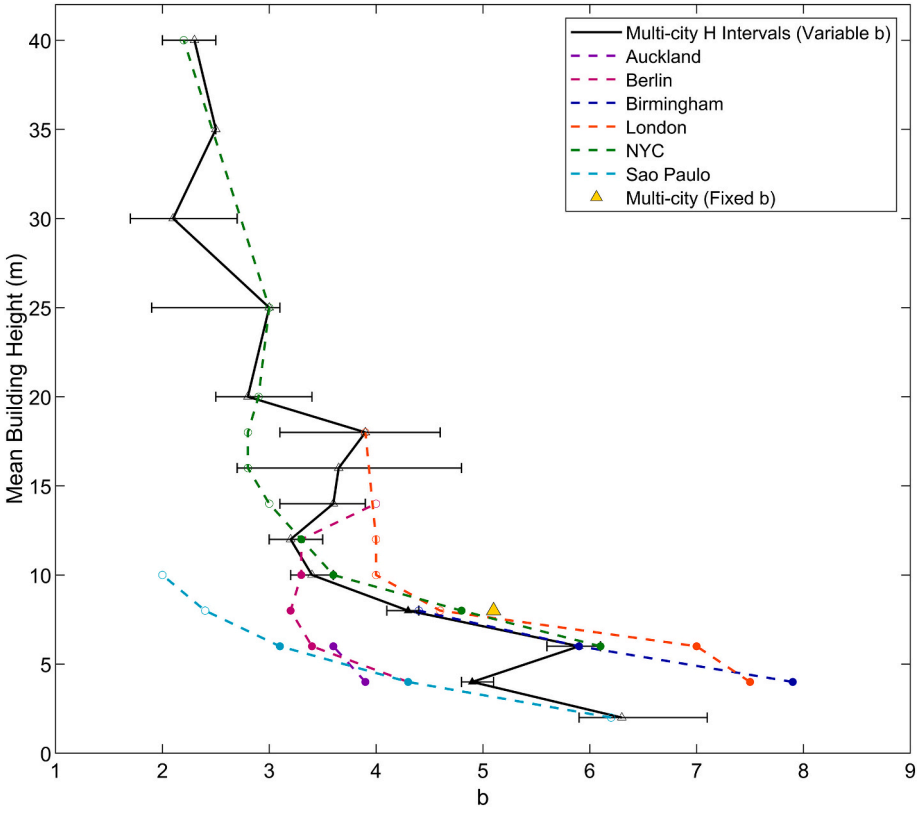

Fig. 6. Relation between b (Eq. 4) and mean building height ($\bar{H}$) for each city (Table 1) and across all cities, for $\bar{H}$ intervals (2 m: 2–20 m, 5 m: 20–40 m). $\bar{H}$ intervals with sample size ≥ 5, containing $> 10\%$ of individual city grid-cells or across all (multi-city) (filled, otherwise open) the 'variable b' spread (whiskers) in median b with height are calculated from 1000 bootstrap samples, random with repetition to the same sample size of each interval, for the full dataset (P0, Table 2).

Fig. 6), which may be associated with smaller sample sizes. The multi-city b with $\bar{H}$ relation follows NYC closely (green, Fig. 6), as NYC has the most grid-cells sampled. Differences between the multi-city median and the NYC b at $\bar{H} = 40$ m arise as extra grid-cells from other cities (i.e., with <5 grid-cells) are included. Auckland and Sao Paulo have b values are most different from both other cities and the multi-city relation (Fig. 6). The UK cities, London and Birmingham, are similar to ~10 m but differ above this as London has higher $\bar{H}$. Other cities with different variations in morphology may lie either side, or elsewhere in relation to this multi-city curve, e.g., taller

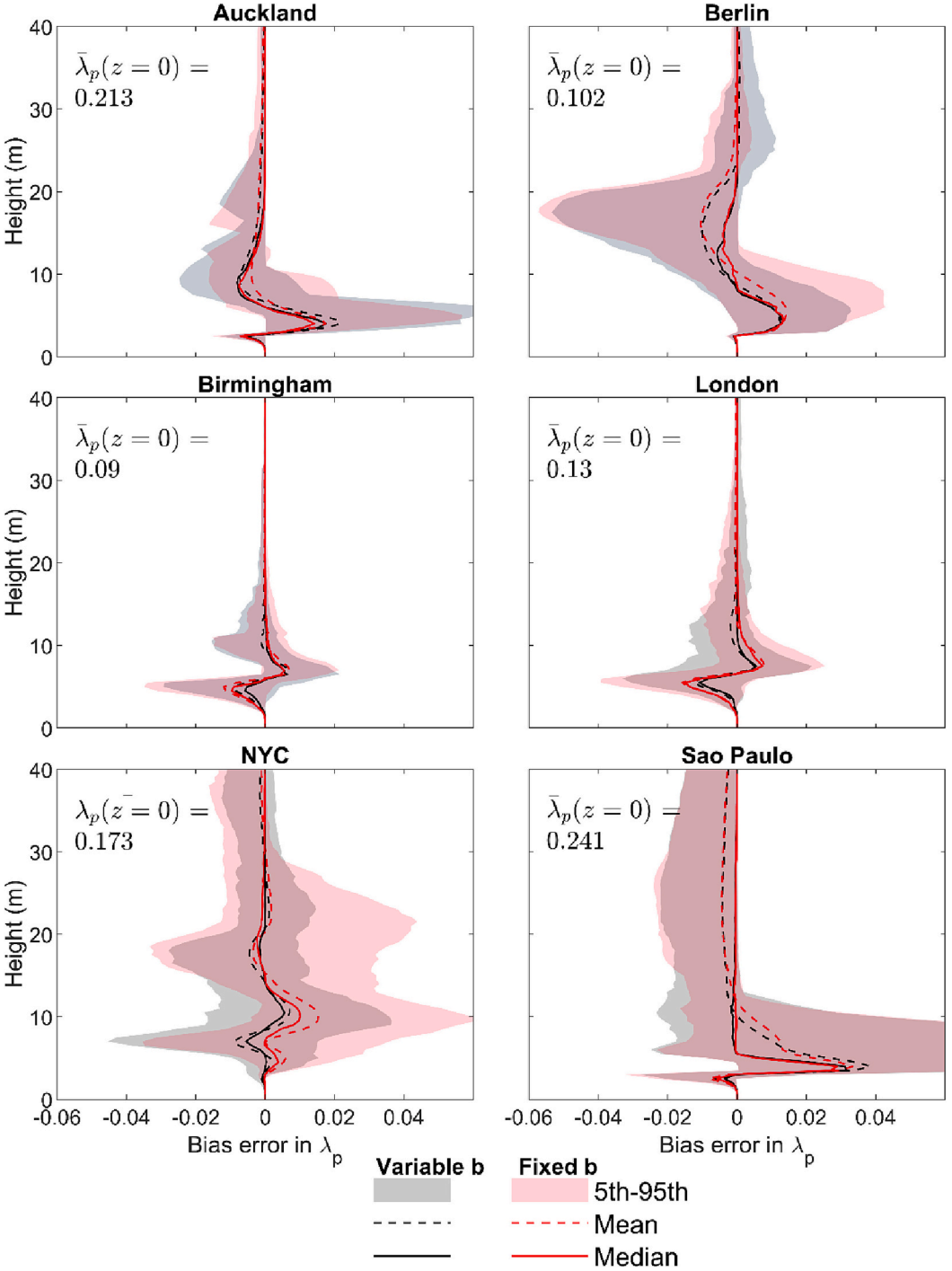

Fig. 7. As Fig. 5, but mean city plan area building fraction at the surface ($\bar{\lambda}_{p,0}$) and bias error (BE) vertical profiles (λ_p) determined 'variable b' (grey) and 'fixed b' (red) (Table 2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
As Table 3, wall area (λ_w) for six cities calculated using P2-P5 (Table 2) assessed using P0 with bias and normalised bias error (MBE and nMBE, section 3.4).

City	P0	P2			P3			P4			P5		
	λ_w (m^2/m^2)	MBE (m^2/m^2)	nMBE (%)	MAE (m^{-1})	MBE (m^2/m^2)	nMBE (%)	MAE (m^{-1})	MBE (m^{-1})	nMBE (%)	MAE (m^{-1})	MBE (m^{-1})	MBE (%)	MAE (m^{-1})
Auckland	0.482	−0.0078	−2.3	−0.0078	−0.0077	−2.3	−0.0077	−0.060	−17	−0.0603	−0.088	−27	−0.088
Birmingham	0.165	−0.0030	−2.2	−0.0029	−0.0029	−2.2	−0.0029	−0.0054	−0.75	−0.0054	−0.020	−15	−0.020
Berlin	0.187	−0.0044	−2.1	−0.0044	−0.0045	−2.1	−0.0045	−0.012	−5.3	−0.012	0.0025	−8.4	0.0025
London	0.250	−0.0051	−2.2	−0.0050	−0.0050	−2.2	−0.0050	−0.016	−6.9	−0.016	−0.021	−15	−0.021
NYC	0.302	−0.011	−2.4	−0.011	−0.0067	−2.1	−0.0067	−0.053	−2.3	−0.053	0.069	11	0.069
Sao Paulo	0.287	−0.0069	−2.1	−0.0069	−0.0068	−2.0	−0.0068	−0.00088	3.7	−0.00088	−0.021	−9.5	−0.021

cities (Asian megacities) may sit with Sao Paulo, or more with NYC.

Using the *variable b* and *fixed b* parameterisations (Table 2) gives absolute BE for λ_p of <0.025 for all heights in all cities (black lines, Fig. 7), except Sao Paulo. There is little difference between the two parameterisations. Generally, 90% of the data (5th - 95th percentile, shading Fig. 7) in each city are within 0.03 of the 'true' values at each height. These results suggest Eq. 1 is applicable to real-world cities.

Focusing on individual cities, the UK cities have similar vertical profiles (Fig. 7), as expected given their similar *b* values (Fig. 6). For Auckland and Sao Paulo, λ_p is underestimated below 10 m and overestimated above, the opposite to London and Birmingham. For NYC, large errors extend to higher building heights, especially for the *fixed b* parameterisation. In Sao Paulo the largest BE (> 0.05) occur when $z < 10$ m. Berlin, up to 20 m, has the second largest maximum magnitude of BE (~0.06), but low mean and median absolute BE (< 0.02).

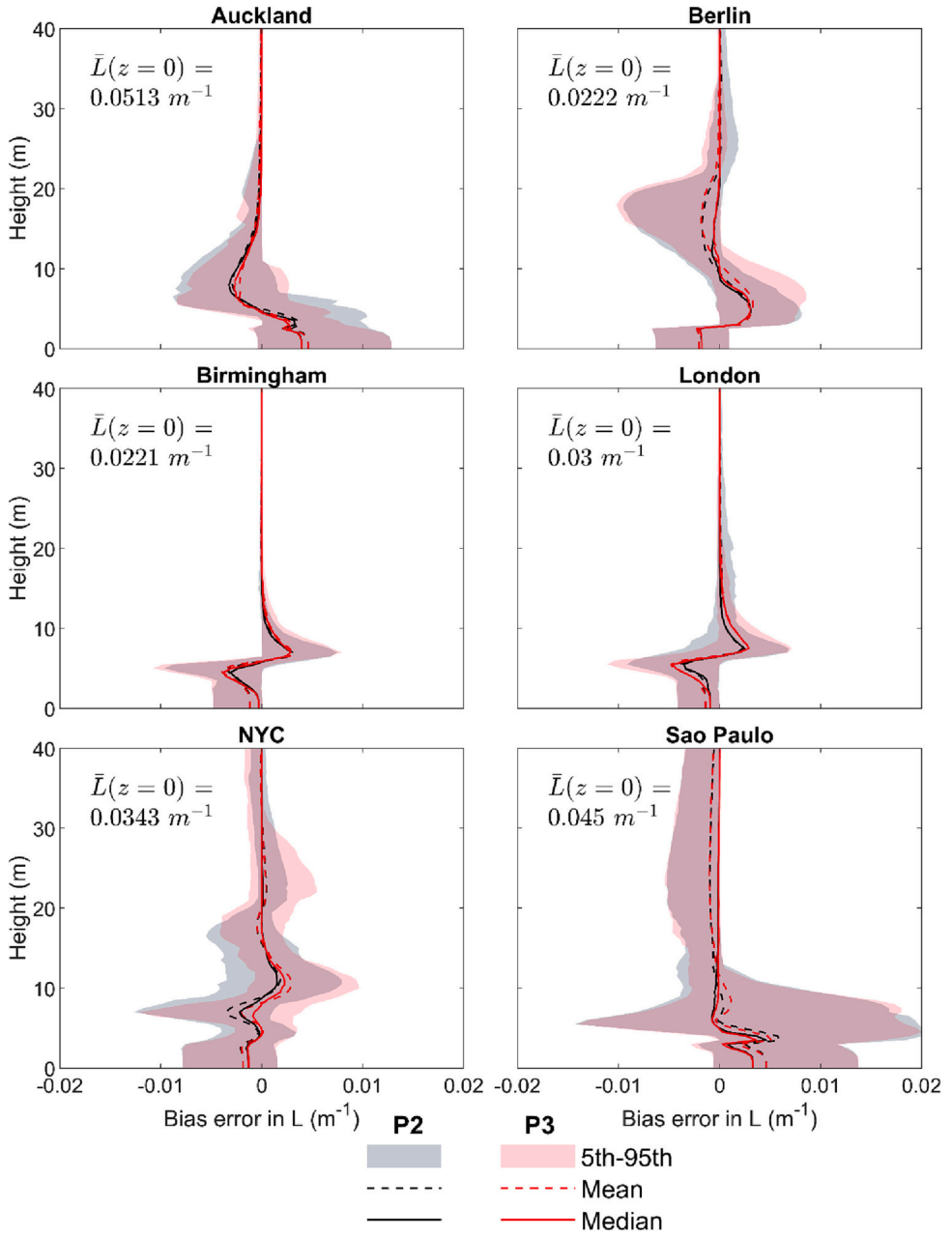

Fig. 8. As Fig. 5, but for P2 and P3.

Overall, the *fixed b* parameterisation b (Eq. 4) gives a reasonable fit in all cities. Comparing both parameterisations (*variable b*, *fixed b*) shows that only small improvements when b values are dependent on mean building height.

4.4. Evaluation of combined parameterisations

The combined $L(z)$ and λ_p parameterisations (P1 → P5, Table 2) are evaluated using the P0 high resolution reference profiles for each city. This evaluation also assesses λ_w for parameterisations of L.

MBE for λ_w is smallest if more city-specific information is used (Table 2). For mean λ_w, the nMBE are <2.5% for P2 and P3, but up to −17% for P4 (Table 4). The largest nMBE occur in Auckland and Berlin, and smallest in Birmingham and London. Using *fixed*

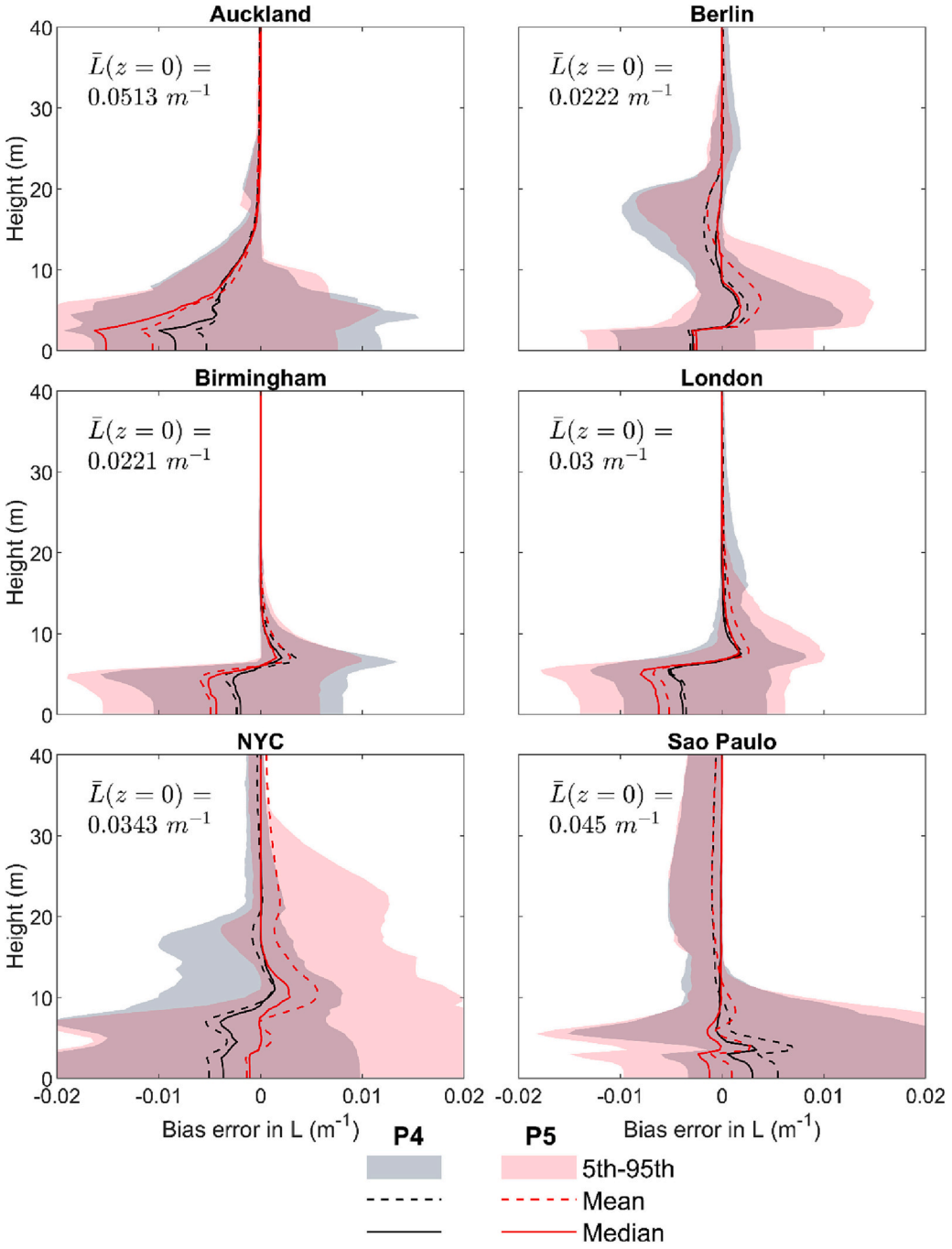

Fig. 9. As Fig. 8, but for P4 and P5.

Table 5
As Table 4, but for SPARTACUS-Urban simulated total absorption when the solar zenith angle is 75° for (a) walls, (b) roof, and (c) ground facets.

City	P0	P2			P3			P4			P5		
(a) Wall	Absorption ($W\ m^{-2}$)	MBE ($W\ m^{-2}$)	nMBE (%)	MAE ($W\ m^{-2}$)	MBE ($W\ m^{-2}$)	nMBE (%)	MAE ($W\ m^{-2}$)	MBE ($W\ m^{-2}$)	nMBE (%)	MAE ($W\ m^{-2}$)	MBE ($W\ m^{-2}$)	nMBE (%)	MAE ($W\ m^{-2}$)
Auckland	68.3	−2.1	−2.7	−2.1	−2.02	−2.6	−2.0	−10	−15	−10	−16	−25	−16
Birmingham	31.9	−0.55	−1.9	−0.55	−0.53	−1.8	−0.53	−1.5	−1.1	−1.5	−4.9	−15	−4.9
Berlin	44.6	−0.74	−1.9	−0.74	−0.9	−2.0	−0.90	−2.1	−5.1	−2.1	−1.2	−9.4	−1.2
London	47.8	−0.68	−1.7	−0.68	−0.72	−1.7	−0.72	−3.2	−6.3	−3.2	−5.6	−14	−5.6
NYC	73.2	−0.89	−1.5	−0.89	−1.59	−2.1	−1.6	−4.8	−0.78	−4.8	0.22	5.4	0.22
Sao Paulo	62.5	−2.7	−3.3	−2.7	−2.66	−3.2	−2.7	−1.4	1.3	−1.4	−5.5	−10	−5.5
(b) Roof													
Auckland	37.9	1.8	3.5	1.8	1.7	3.2	1.7	2.4	4.7	2.4	2.6	5.2	2.6
Birmingham	18.0	−0.24	−0.95	−0.24	−0.31	−1.2	−0.31	−0.28	−0.92	−0.28	−0.28	−0.85	−0.28
Berlin	19.0	0.18	0.83	0.18	0.5	1.7	0.50	0.21	0.98	0.21	0.37	1.5	0.37
London	25.3	−0.47	−1.3	−0.47	−0.46	−1.5	−0.46	−0.46	−1.14	−0.46	−0.45	−1.1	−0.45
NYC	30.3	−0.31	−0.74	−0.31	0.8	1.9	0.80	0.12	0.39	0.12	0.21	0.51	0.21
Sao Paulo	41.4	3.0	5.0	3.0	2.8	4.9	2.8	2.9	5.0	2.9	2.9	5.2	2.9
(c) Ground													
Auckland	111	−0.31	−0.46	−0.097	−0.10	−0.40	0.0070	0.0070	−0.32	6.2	6.2	7.5	10
Birmingham	162	0.11	0.086	0.70	0.70	0.51	0.75	0.75	0.55	1.5	1.5	0.86	4.3
Berlin	150	0.066	0.043	0.44	0.44	0.40	0.25	0.25	0.17	1.5	1.5	1.4	0.58
London	140	0.15	0.16	1.0	1.0	1.0	1.1	1.1	0.95	3.1	3.1	2.4	5.1
NYC	115	0.26	0.36	1.1	1.1	1.7	0.53	0.53	0.12	4.0	4.0	9.0	−0.47
Sao Paulo	113	−0.40	−0.83	−0.72	−0.72	−1.8	−0.58	−0.58	−1.6	−1.7	−1.7	−2.4	1.6

coefficients (P5) has the largest nMBE in all cities, with Auckland the poorest (−27%).

For P2 and P3, profiles of $L(z)$ have similar BE in all cities (Fig. 8). In NYC, P3 tends to underestimate $L(z)$, and P2 overestimate $L(z)$, least. Across all cities, 90% (5-95th percentile, shading Fig. 8) of both the P2 and P3 profiles are <0.01 m^{-1} different from P0, except for Sao Paulo between 5 and 10 m. In London and Birmingham, both P2 and P3 overestimate $L(z)$ below $z = 5$ m but underestimate $L(z)$ above, while in Sao Paulo and Auckland the opposite occurs. Overall, the largest errors occur in Sao Paulo.

Consistent with the errors in total λ_w (Table 4), P4 and P5 perform the poorest for $L(z)$ (Fig. 9), Generally, the $L(z)$ absolute value of BE are <0.01 for P4 (black) and P5 (red, Fig. 9) with close to 90% of the profiles being within magnitude of 0.02 of P0. For all cities when P5 is used, the range of the 5 - 95th percentile of profiles is wider.

5. Impacts of applying parameterisations for λ_p and L in the SPARTACUS-urban radiative transfer model

5.1. Impact on vertically integrated absorbed shortwave radiation

To assess the impact of approximating the urban morphology using parameterisations (P2-P5, Table 2) on simulated (section 3.3) total integrated shortwave absorption into urban facets, results are compared to the P0 in each grid-cell. Although three solar zenith

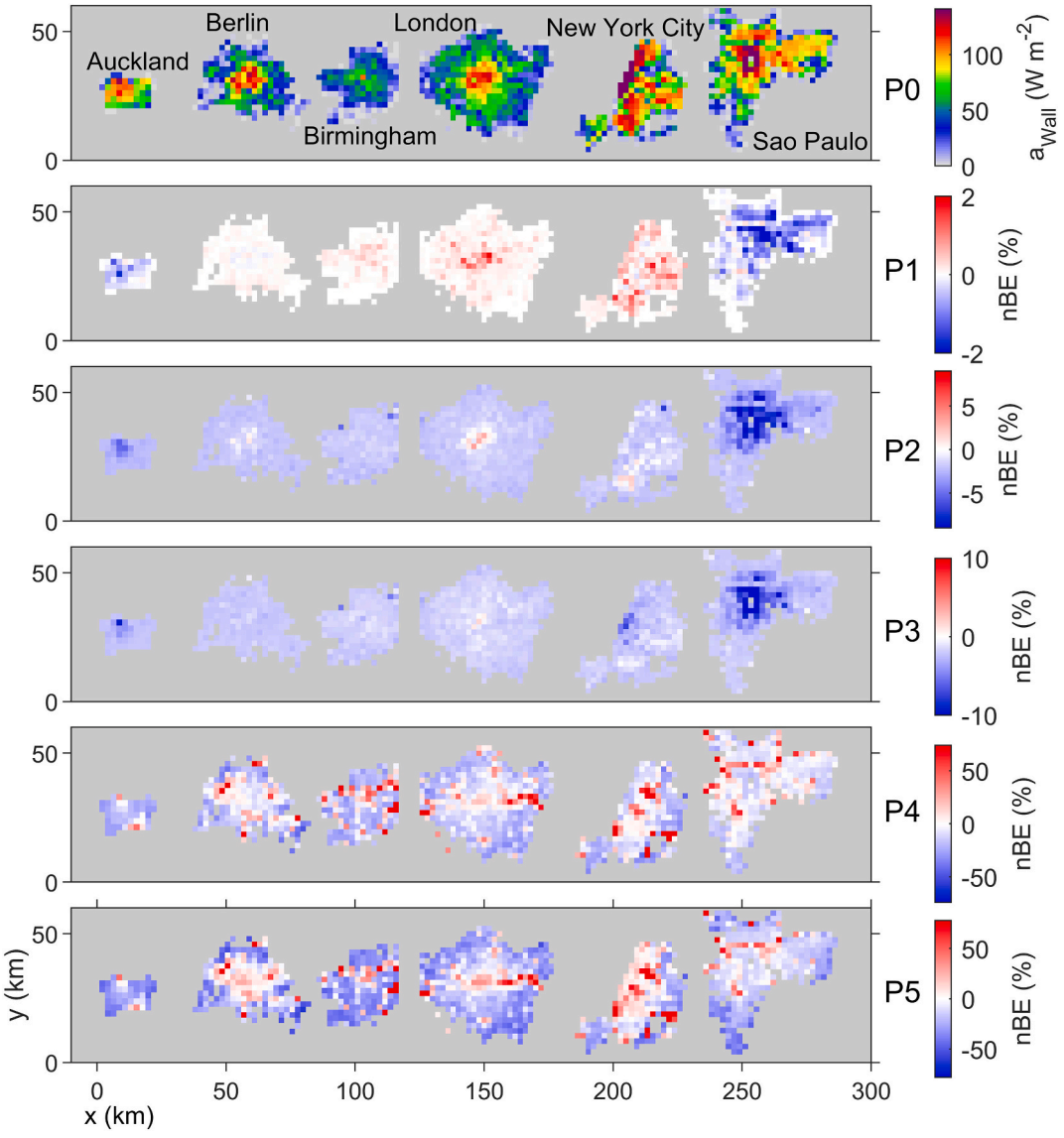

Fig. 10. Total absorbed shortwave radiation into walls (a_{Wall}) modelled with SPARTACUS-Urban for a solar zenith angle (θ_0) of 75° for six cities using 'true' building fraction and building edge length vertical profiles (P0, Table 2), and normalised bias error (nBE, section 3.4) when using parameterised profiles (P1-P5). Note scales differ between rows. See Fig. SM 3 for $\theta_0 = 0°$ and Fig. SM 4 for $\theta_0 = 45°$.

angles (θ_0) are simulated, as $\theta_0 = 75^\circ$ has the largest normalised bias errors (nBE) we focus on these, with the $\theta_0 = 0^\circ$ and 45° results given in Fig. SM 3 – Fig. SM 6.

The SPARTACUS-Urban simulated wall absorption (a_{Wall}) with a $\theta_0 = 75^\circ$, generally have nMBE magnitudes >3% for P2 (cf. > 15% as complexity increases to P5, Table 5) across all cities. Given poorer skill of the urban form parameterisations in Auckland (section 4), the poorest a_{Wall} performance is consistent (P4 15%; P5 25%) (Table 5). The largest nBE in a_{Wall} occur when using P1-P3 in Auckland, NYC, and Sao Paulo, in areas with the largest $\lambda_p(z = 0)$ and $L(z = 0)$ (Fig. 2, cf. Fig. 10). However, for P4 and P5 no clear relation is evident between grid-cell morphology and a_{Wall} nBE. Generally, P1 - P3 underestimate the absorption of radiation by walls (Table 5), which is consistent with the morphology results (Table 4).

Simulated a_{Roof}, using P2-P5 morphology, have all cities have a mean nBE < 5% (Table 5), and across all grid-cells the largest nBE magnitudes <25% (Fig. 11). Given the skill in morphology parameterisations (section 4), Sao Paulo and Auckland overestimate a_{Roof}, while London and Birmingham underestimate a_{Roof}. The highest nBE occur in areas with the largest $\lambda_p(z = 0)$ (e.g., CBD, industrial), and therefore P0 a_{Roof} is highest.

Across all parameterisations, many cities have an absolute value of nMBE in mean a_{Ground} of <2% (Table 5), apart from P5 in Auckland (7.5%) and NYC (9%). The spatial patterns of a_{Ground} show underestimates in Auckland and Sao Paulo, but overestimates in the other cities when using P1-P3 (Fig. 12).

Overall, the lowest nMBE occur when $\theta_0 = 0^\circ$ with a_{Roof} and a_{Ground} nBE < 1% for all cities (Table SM **3**); but with larger nMBE for

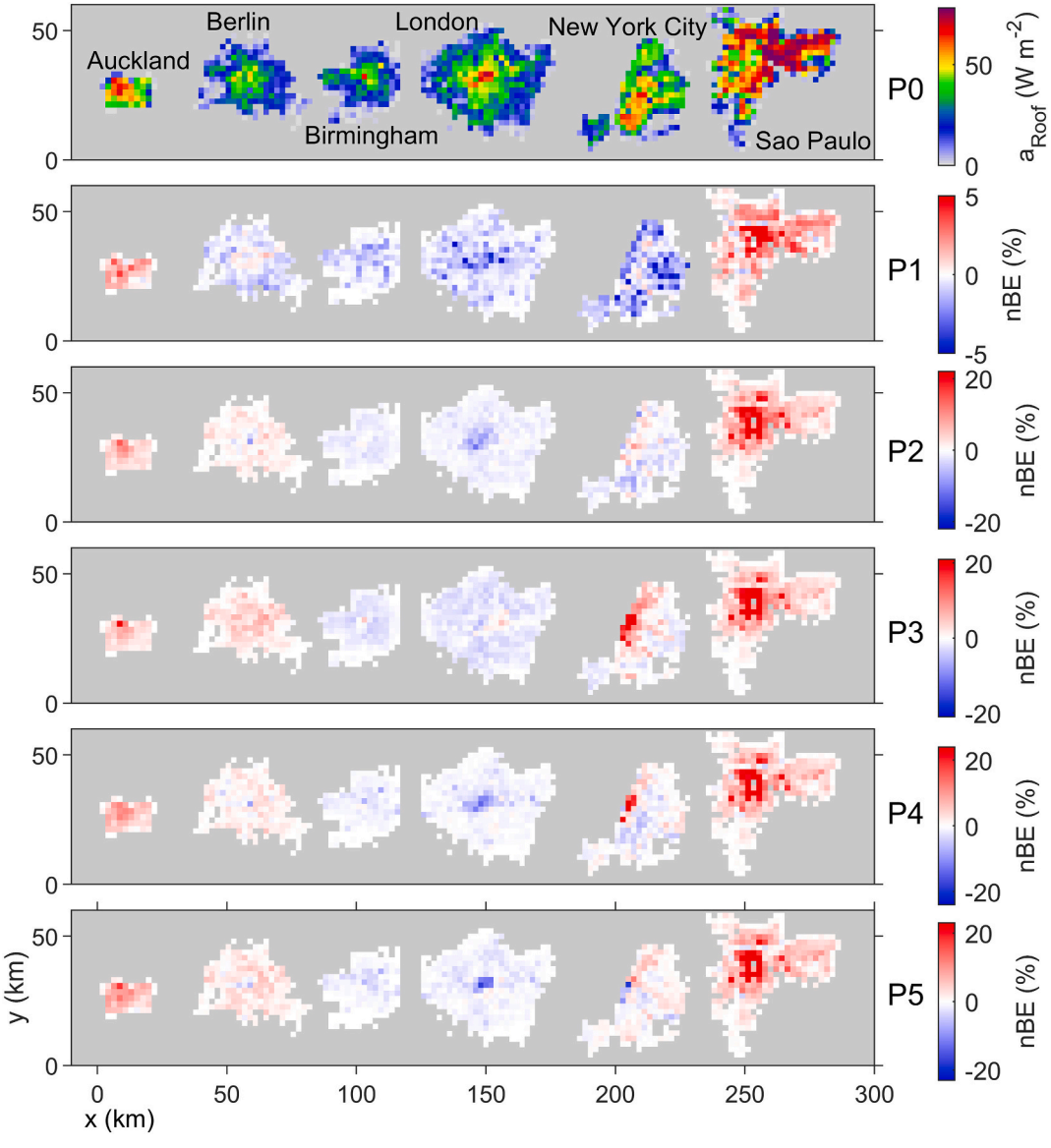

Fig. 11. As Fig. 10, but for total roof absorption (a_{Roof}). See Fig. SM 5 for $\theta_0 = 45^\circ$.

a_{Wall} (2–5% for P1-P3, up to 50% for P4 and P5, Fig. SM 3). When $\theta_0 = 45°$ the nMBE for a_{Wall} are similar to $\theta_0 = 0°$ (P1-P3 < 5%, and P4 and P5 > 50%; Table SM 2, Fig. SM 4) and a_{Roof} (2–5%) with the largest nBE associated with larger $\lambda_p(z = 0)$. The $\theta_0 = 45°$ a_{Ground} values have nBE that are <10% (Table SM 2, Fig. SM 6).

5.2. Impact on shortwave bulk albedo

We additionally assess the error in approximating the urban morphology using P2-P5 on the simulated bulk albedo for each grid-cell (i.e., ratio of upwelling to downwelling shortwave radiation at the top of the urban canopy). For all θ_0 values tested, the nBE have similar magnitudes across the six cities, so we focus on $\theta_0 = 45°$. The nBE results for $\theta_0 = 75°$ are slightly larger (Fig. SM 7).

Simulated bulk albedo generally increases with parameterisation complexity (i.e., P1-P5), but all have an absolute nBE < 10% (Fig. 13). This is smaller than the total facets absorption nBE (Fig. 10 - Fig. 12) that have a maximum of ±20–50% for P5 (largest for a_{Wall}). This suggests bulk effects result from within canopy compensation (i.e., some absorption overestimates and underestimates). The net result is a good approximation of the shortwave bulk albedo for all parameterisations. Unlike absorption, albedo has a similar trend for all six cities, being overestimated for P1-P3 in all grid-cells and for P4-P5 in most grid-cells. For the latter, notably P5, the albedo is underestimated in areas with the higher surface building fractions (Fig. 2).

Hence, the crudest assumptions about urban form have low impact on the bulk albedo. However, this contrasts with the error in shortwave absorption (section 5.1), suggesting that the within-canopy impacts work to cancel one another out.

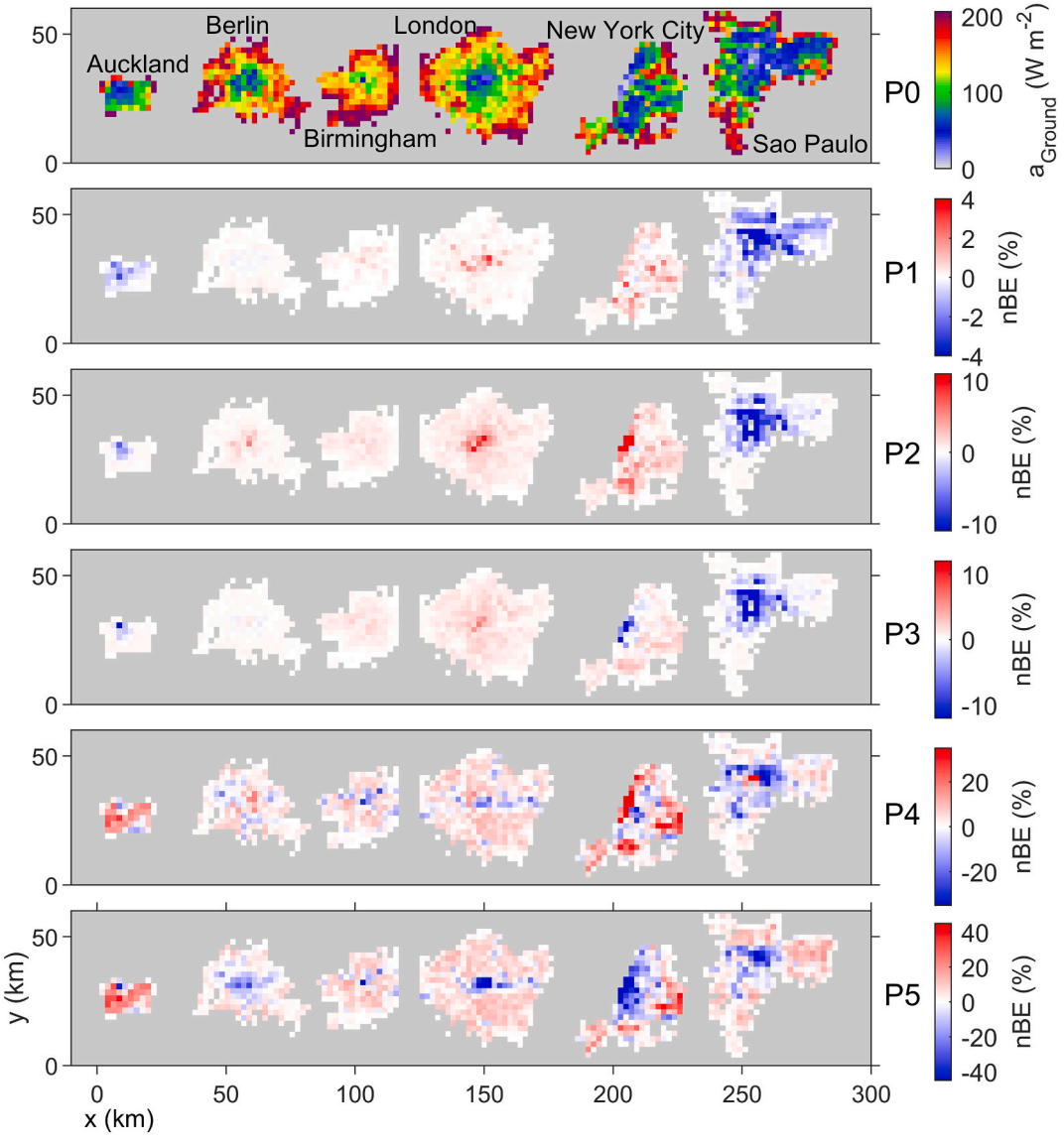

Fig. 12. As Fig. 10, but for total ground absorption (a_{Ground}). See Fig. SM 6 for $\theta_0 = 45°$.

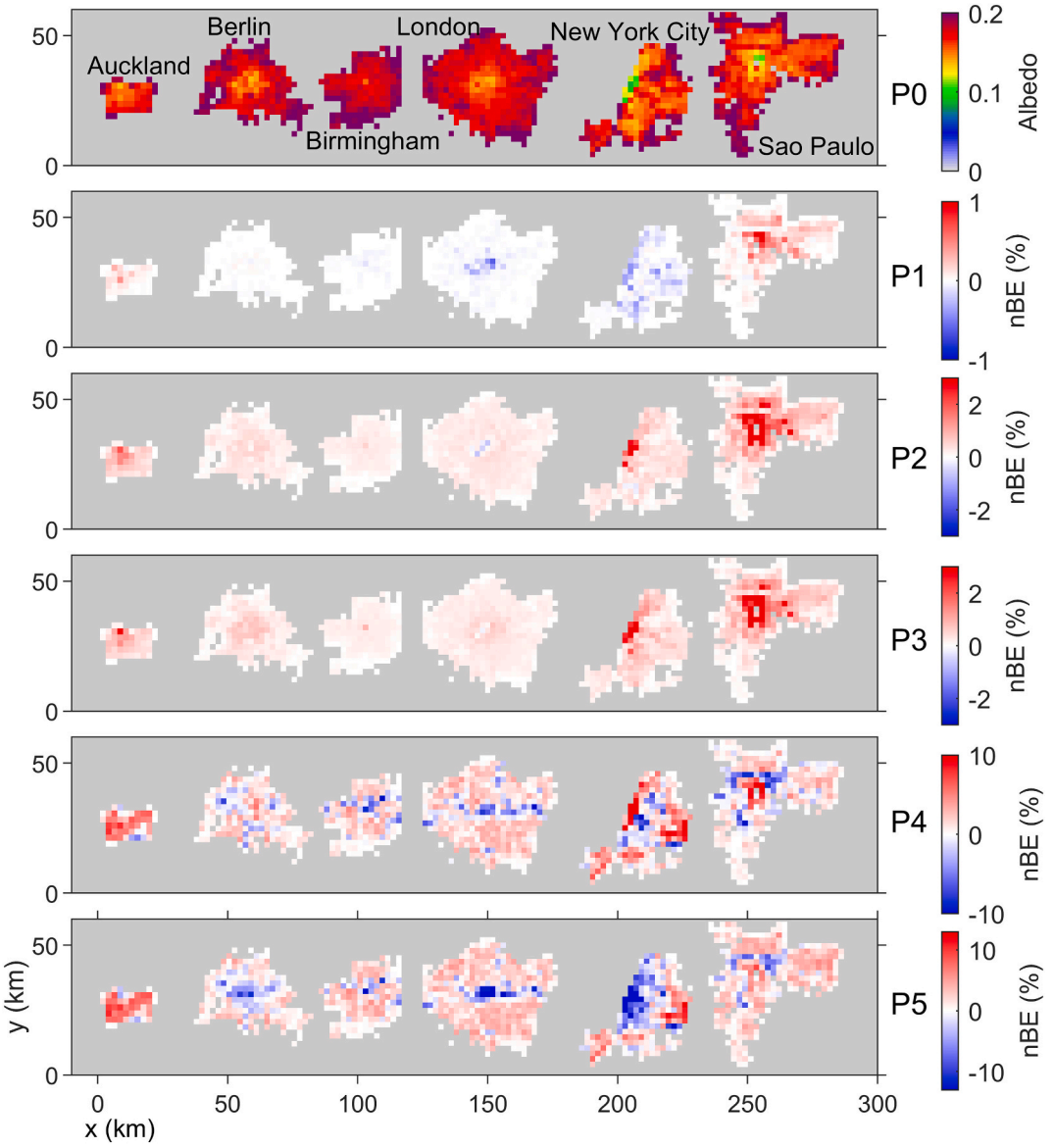

Fig. 13. As Fig. 10, but for bulk albedo for $\theta_o = 45°$. See Fig. SM 7 for $\theta_o = 75°$.

6. Conclusions

Given the challenges in collecting urban morphology data (Frantz et al., 2021; Masson, 2020), that are critical to simulating meteorological processes in cities (e.g. Hogan, 2019b; Krayenhoff et al., 2020), we propose methods to parameterise vertical profiles of building plan area fraction (λ_p) and normalised building perimeter length (L). The latter is simply the vertically resolved normalised wall area.

Although profiles can be derived from high resolution building height data, few datasets with large spatial extent exist. Here, data are analysed for six cities (in Europe, North America, Oceania, and South America) to derive both city specific and multi-city parameters, and to evaluate model skill.

The impact of parameter uncertainty is assessed using SPARTACUS-Surface (Hogan, 2019b), shortwave radiation simulations of albedo, and shortwave radiation absorption by three urban facets (roof, walls, ground). The grid-cell bulk albedo has an absolute normalised bias error of $\leq$10% when using the urban form parameterisations. Albedo is more likely to be underestimated in high-density central-city areas when the most assumptions are made about the urban morphology. Reducing the assumptions made, and therefore increasing the knowledge of the urban morphology used within the parameterisation, decreases the absolute value of normalised bias errors to $\leq$2% for all areas of cities.

The low error in bulk albedo hides errors within the urban canopy introduced by the parameterisations, as errors in shortwave

absorption appear to cancel out. This may make the parameterisations acceptable for use with more complex models. Generally, low ($< 20\%$) absolute values of normalised errors in absorption occur at when the sun is overhead, when building fraction parameterisation is more influential (cf. wall area parameterisation). The largest normalised bias errors (up to 50%) occur in wall absorption for larger solar zenith angles, when wall area is assumed unknown (Section 4.4, 5.1). Thus, trying to overcome a lack of urban form data through parameterisations introduces larger errors. Hence, if vertical profiles of radiation absorption are the focus, appropriate wall area data are critical.

We conclude that building height variations can be approximated using a function of mean building height and surface λ_p. Parameters are derived for both all-cities and each city. Overall, 90% of the profiles have building fraction error $< \pm 0.03$ at any height (cf. 'true' data). As there is little difference between the two methods in four out of six cities analysed, it suggests the multi-city approach may be acceptable.

Additionally, we find intra-city building horizontal extent can be approximated using an effective building diameter D derived either as function of mean building height and λ_p, or from the 'true' wall area. The latter is most constrained by data availability. Although we treat D as constant with height, we find evidence suggesting two values may improve representation of vertical variability. The near ground (for areas with $\lambda_p > 0.05$) mean D across all six cities is $\sim$21 m. For all methods used to determine D, the absolute value of median bias error is $<$6%, with the magnitude of bias error of 90% of all profiles $<0.02\ \mathrm{m}^{-1}$ (cf. $0.02–0.05\ \mathrm{m}^{-1}$ at the surface).

The proposed parameterisations can be combined to give a full description of the vertical morphology, with five combinations used here, with decreasing data requirements. Little skill is lost from using a more general (cf. detailed) parameterisation for λ_p. Parameterisation of L has a larger impact on wall area. If wall area is known but λ_p is parameterised, the mean total wall area is calculated within 3% for all cities, whereas if wall area is unknown this error can increase by up to an order of magnitude. The parameterisations for the six cities perform best in areas with both low λ_p and L (e.g., suburban areas). In denser areas, errors are generally larger and increase without knowledge of wall area. Additionally, in denser areas the errors in absorbed shortwave and bulk albedo may have opposite signs of those in the outer city.

Overall, as our parameterisations can characterise vertical profiles of λ_p and L they have utility in multi-layer urban canopy models (e.g., SPARTACUS-Urban, BEP, TEB). However, the uncertainty in radiative fluxes shown here will impact other modelled fluxes (e.g., storage and sensible heat fluxes) so future work should assess these cascading effects. Alongside this, as the morphology within a given grid cell influences the surface roughness, future work could include assessing the impact of the parameterisations developed here on the momentum transfer. Here, only λ_p and mean building height data are used to approximate urban form. Including building wall area information would improve this, but it is the most challenging of the three variables to obtain. Additionally, we note that the coefficients and best-fit parameters derived here, which impact absorption and albedo, may change as the range and mean of cities properties are expanded and analysed at different scales.

If only mean building height and plan area fraction are available for a city at grid resolution, building height and λ_p can be derived from existing empirical relations, such as Bohnenstengel et al. (2011), that are currently applied globally, despite being developed from data from a single city. But, these relations require input data (i.e., built or 'urban' [*sic*] fraction). Thus, there is a clear and increasing need to improve global datasets of urban morphological parameters for weather and climate models, but also for a wide range of other applications that use the outputs from these models for decision making and managing cities day-to-day, in emergencies and for the long-term changes (both climate and urban).

Author contributions

MS performed all modelling, all data analysis, and wrote the initial manuscript. MS and RH developed the parameterisation method. RH is the main author of the SPARTACUS-Urban code. MS, RH, and SG designed the manuscript structure, and all co-authors read and provided feedback on the manuscript. RH and SG formulated the initial idea. SG applied for and obtained the funding to support all (including earlier Kent obtained data) except for RH.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data and code are archived at https://zenodo.org/10.5281/zenodo.5851021. SPARTACUS-Surface is freely available at https://github.com/ecmwf/spartacus-surface.

Acknowledgements

Funding from the NERC Scenario Doctoral Training Partnership Grant, EPSRC 2130186, EPRSC DARE EP/P002331/1, ERC Synergy *urbisphere (855005)*, and Newton Fund/Met Office CSSP China NGC (SG). Processed data for Berlin: Karthik Reddy, Andreas Christen, Environmental Atlas Berlin; Auckland, New York City, and Sao Paulo: Christoph Kent for rasterised data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.uclim.2023.101560.

References

Arnfield, A.J., 1982. An approach to the estimation of the surface radiative properties and radiation budgets of cities. Phys. Geogr. 3, 97–122. https://doi.org/10.1080/02723646.1982.10642221.

Arnfield, A.J., Grimmond, C.S.B., 1998. An urban canyon energy budget model and its application to urban storage heat flux modeling. Energy Build. 27, 61–68. https://doi.org/10.1016/S0378-7788(97)00026-1.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T., 2011. Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Mon. Weather Rev. 139, 3887–3905. https://doi.org/10.1175/MWR-D-10-05013.1.

Balogun, A.A., Adegoke, J.O., Vezhapparambu, S., Mauder, M., McFadden, J.P., Gallo, K., 2009. Surface energy balance measurements above an exurban residential neighbourhood of Kansas City, Missouri. Boundary-Layer Meteorol. https://doi.org/10.1007/s10546-009-9421-3.

Bergeron, O., Strachan, I.B., 2012. Wintertime radiation and energy budget along an urbanization gradient in Montreal, Canada. Int. J. Climatol. https://doi.org/10.1002/joc.2246.

Bernard, J., Bocher, E., Le, E., Wiederhold, S., Leconte, F., Masson, V., 2022. Estimation of missing building height in OpenStreetMap data: a French case study using GeoClimate 0.0.1. Geosci. Model Dev. 15, 7505–7532. https://doi.org/10.5194/gmd-15-7505-2022.

Bohnenstengel, S.I., Evans, S., Clark, P.A., Belcher, S.E., 2011. Simulations of the London urban heat island. Q. J. R. Meteorol. Soc. 137, 1625–1640. https://doi.org/10.1002/QJ.855.

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, 116–134. https://doi.org/10.1016/J.UCLIM.2016.04.001.

Champeaux, J.L., Masson, V., Chauvin, F., 2005. ECOCLIMAP: a global database of land surface parameters at 1 km resolution. Meteorol. Appl. 12, 29–32. https://doi.org/10.1017/S1350482705001519/FORMAT/PDF.

Collier, C.G., 2006. The impact of urban areas on weather. Q. J. R. Meteorol. Soc. https://doi.org/10.1256/qj.05.199.

Demographia World Urban Areas, 2020. Demographia World Urban Areas, 16th Annual Edition 2020.06. *Demographia*.

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I.D., Van Vliet, J., Bechtel, B., 2022. A global map of local climate zones to support earth system modelling and urban-scale environmental science. Earth Syst. Sci. Data 14, 3835–3873. https://doi.org/10.5194/essd-14-3835-2022.

EMU Analytics, 2018: Building heights – top 25 urban areas.

Feigenwinter, C., Vogt, R., Parlow, E., 1999. Vertical structure of selected turbulence characteristics above an urban canopy. Theor. Appl. Climatol. https://doi.org/10.1007/s007040050074.

Frantz, D., Schug, F., Okujeni, A., Navacchi, C., Wagner, W., van der Linden, S., Hostert, P., 2021. National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series. Remote Sens. Environ. 252, 112128 https://doi.org/10.1016/J.RSE.2020.112128.

Gage, E.A., Cooper, D.J., 2017. Relationships between landscape pattern metrics, vertical structure and surface urban Heat Island formation in a Colorado suburb. Urban Ecosyst. https://doi.org/10.1007/s11252-017-0675-0.

Gamba, P., Houshmand, B., 2002. Joint analysis of SAR, LIDAR and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings. Int. J. Remote Sens. https://doi.org/10.1080/01431160110114952.

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. https://doi.org/10.5589/m09-015.

Grimmond, S., et al., 2020. Integrated urban hydrometeorological, climate and environmental services: concept, methodology and key messages. Urban Clim. 33, 100623 https://doi.org/10.1016/J.UCLIM.2020.100623.

Grimmond, C.S.B., Oke, T.R., 1995. Comparison of heat fluxes from summertime observations in the suburbs of four North American cities. J. Appl. Meteorol. https://doi.org/10.1175/1520-0450(1995)034<0873:COHFFS>2.0.CO;2.

Grimmond, C.S.B., Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. J. Appl. Meteorol. https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2.

Grimmond, C.S.B.B., Salmond, J.A., Oke, T.R., Offerle, B., Lemonsu, A., 2004. Flux and turbulence measurements at a densely built-up site in Marseille: heat, mass (water and carbon dioxide), and momentum. J. Geophys. Res. D Atmos. 109, 1–19. https://doi.org/10.1029/2004JD004936.

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, 84, 427–439.

Heris, M.P., Foks, N.L., Bagstad, K.J., Troy, A., Ancona, Z.H., 2020. A rasterized building footprint dataset for the United States. Sci. Data. https://doi.org/10.1038/s41597-020-0542-3.

Hertwig, D., et al., 2020. Urban signals in high-resolution weather and climate simulations: role of urban land-surface characterisation. Theor. Appl. Climatol. 142, 701–728. https://doi.org/10.1007/s00704-020-03294-1.

Hogan, R.J., 2019a. An exponential model of urban geometry for use in radiative transfer applications. Boundary-Layer Meteorol. https://doi.org/10.1007/s10546-018-0409-8.

Hogan, R.J., 2019b. Flexible treatment of radiative transfer in complex urban canopies for use in weather and climate models. Boundary-Layer Meteorol. https://doi.org/10.1007/s10546-019-00457-0.

Hogan, R.J., Shonk, J.K.P., 2013. Incorporating the effects of 3D radiative transfer in the presence of clouds intol two-stream multilayer radiation schemes. J. Atmos. Sci. https://doi.org/10.1175/JAS-D-12-041.1.

Hogan, R.J., Schäfer, S.A.K., Klinger, C., Chiu, J.C., Mayer, B., 2016. Representing 3-D cloud radiation effects in two-stream schemes: 2. Matrix formulation and broadband evaluation. J. Geophys. Res. 121, 8583–8599. https://doi.org/10.1002/2016JD024875.

Hogan, R.J., Quaife, T., Braghiere, R., 2018. Fast matrix treatment of 3-D radiative transfer in vegetation canopies: SPARTACUS-vegetation 1.1. Geosci. Model Dev. 11, 339–350. https://doi.org/10.5194/gmd-11-339-2018.

Holland, D.E., Berglund, J.A., Spruce, J.P., McKellip, R.D., 2008. Derivation of effective aerodynamic surface roughness in urban areas from airborne Lidar terrain data. J. Appl. Meteorol. Climatol. 47, 2614–2626. https://doi.org/10.1175/2008JAMC1751.1.

Jensen, M.P., Vogelmann, A.M., Collins, W.D., Zhang, G.J., Luke, E.P., 2008. Investigation of regional and seasonal variations in marine boundary layer cloud properties from MODIS observations. J. Clim. 21, 4955–4973. https://doi.org/10.1175/2008JCLI1974.1.

Kanda, M., Kawai, T., Kanega, M., Moriwaki, R., Narita, K., Hagishima, A., 2005a. A simple energy balance model for regular building arrays. Boundary-Layer Meteorol. 116, 423–443. https://doi.org/10.1007/s10546-004-7956-x.

Kanda, M., Kawai, T., Nakagawa, K., 2005b. A simple theoretical radiation scheme for regular building arrays. Boundary-Layer Meteorol. 114, 71–90. https://doi.org/10.1007/s10546-004-8662-4.

Kent, C.W., Grimmond, S., Gatey, D., Hirano, K., 2019. Urban morphology parameters from global digital elevation models: implications for aerodynamic roughness and for wind-speed estimation. Remote Sens. Environ. 221, 316–339. https://doi.org/10.1016/j.rse.2018.09.024.

Kondo, H., Genchi, Y., Kikegawa, Y., Ohashi, Y., Yoshikado, H., Komiyama, H., 2005. Development of a multi-layer urban canopy model for the analysis of energy consumption in a big city: structure of the urban canopy model and its basic performance. Boundary-Layer Meteorol. 116, 395–421. https://doi.org/10.1007/s10546-005-0905-5.

Krayenhoff, E.S., et al., 2020. A multi-layer urban canopy meteorological model with trees (BEP-tree): street tree impacts on pedestrian-level climate. Urban Clim. 32 https://doi.org/10.1016/j.uclim.2020.100590.

Krayenhoff, E.S., Christen, A., Martilli, A., Oke, T.R., 2014. A multi-layer radiation model for urban neighbourhoods with trees. Boundary-Layer Meteorol. 151, 139–178. https://doi.org/10.1007/s10546-013-9883-1.

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. Boundary-Layer Meteorol. 101, 329–358. https://doi.org/10.1023/A:1019207923078.

Lemonsu, A., Grimmond, C.S.B., Masson, V., 2004. Modeling the surface energy balance of the core of an Old Mediterranean City: Marseille. J. Appl. Meteorol. https://doi.org/10.1175/1520-0450(2004)043<0312:MTSEBO>2.0.CO;2.

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 reas. Geosci. Model Dev. https://doi.org/10.5194/gmd-5-1377-2012.

Liang, X., et al., 2018. Surf: understanding and predicting urban convection and haze. Bull. Am. Meteorol. Soc. https://doi.org/10.1175/BAMS-D-16-0178.1.

Lindberg, F., Grimmond, C.S.B., 2011. The influence of vegetation and building morphology on shadow patterns and mean radiant temperatures in urban areas: model development and evaluation. Theor. Appl. Climatol. 105, 311–323. https://doi.org/10.1007/s00704-010-0382-8.

Lindberg, F., Grimmond, C.S.B., Lindberg, F., 2011. Nature of vegetation and building morphology characteristics across a city: Influence on shadow patterns and mean radiant temperatures in London, 14, pp. 617–634. https://doi.org/10.1007/s11252-011-0184-5.

Martilli, A., Clappier, A., Rotach, M.W., 2002. An urban surface exchange parameterisation for mesoscale models. Boundary-Layer Meteorol. 104, 261–304. https://doi.org/10.1023/A:1016099921195.

Martilli, A., Roulet, Y.-A., Junier, M., Kirchner, F., Rotach, M.W., Clappier, A., 2003. On the impact of urban surface exchange parameterisations on air quality simulations: the Athens case. Atmos. Environ. 37, 4217–4231. https://doi.org/10.1016/S1352-2310(03)00564-8.

Masson, V., 2000. A physically-based scheme for the urban energy budget in atmospheric models. Boundary-Layer Meteorol. 94, 357–397. https://doi.org/10.1023/A:1002463829265.

Masson, V., et al., 2020. City-descriptive input data for urban climate models: model requirements, data sources and challenges. Urban Clim. https://doi.org/10.1016/j.uclim.2019.100536.

Milojevic-Dupont, N., et al., 2020. Learning from urban form to predict building heights. PLoS One 15, e0242010. https://doi.org/10.1371/JOURNAL.PONE.0242010.

Molnár, G., Gyöngyösi, A.Z., Gál, T., 2019. Integration of an LCZ-based classification into WRF to assess the intra-urban temperature pattern under a heatwave period in Szeged, Hungary. Theor. Appl. Climatol. 138, 1139–1158. https://doi.org/10.1007/S00704-019-02881-1/FIGURES/11.

Morrison, W., Kotthaus, S., Grimmond, C.S.B., Inagaki, A., Yin, T., Gastellu-Etchegorry, J.P., Kanda, M., Merchant, C.J., 2018. A novel method to obtain three-dimensional urban surface temperature from ground-based thermography. Remote Sens. Environ. 215, 268–283. https://doi.org/10.1016/j.rse.2018.05.004.

Offerle, B., Jonsson, P., Eliasson, I., Grimmond, C.S.B., 2005. Urban modification of the surface energy balance in the West African Sahel: Ouagadougou, Burkina Faso. J. Clim. https://doi.org/10.1175/JCLI3520.1.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban climates.

OPDC, 2018. Tall Buildings Statement LOCAL PLAN SUPPORTING STUDY.

Padiyedath Gopalan, S., Kawamura, A., Amaguchi, H., Takasaki, T., Azhikodan, G., 2019. A bootstrap approach for the parameter uncertainty of an urban-specific rainfall-runoff model. J. Hydrol. 579, 124195 https://doi.org/10.1016/J.JHYDROL.2019.124195.

Porson, A., Clark, P.A., Harman, I.N., Best, M.J., Belcher, S.E., 2010. Implementation of a new urban energy budget scheme into MetUM. Part II: validation against observations and model intercomparison. Q. J. R. Meteorol. Soc. Q. J. R. Meteorol. Soc 136, 1530–1542. https://doi.org/10.1002/qj.572.

Rao, P., 1972. Remote sensing of urban "heat islands" from an environmental satellite. Bull. Am. Meteorol. Soc. 53, 647–648.

Raupach, M.R., 1992. Drag and drag partition on rough surfaces. Boundary-Layer Meteorol. https://doi.org/10.1007/BF00155203.

Ryu, Y.H., Baik, J.J., 2012. Quantitative analysis of factors contributing to urban heat island intensity. J. Appl. Meteorol. Climatol. 51 https://doi.org/10.1175/JAMC-D-11-098.1.

Ryu, Y.H., Baik, J.J., Kwak, K.H., Kim, S., Moon, N., 2013. Impacts of urban land-surface forcing on ozone air quality in the Seoul metropolitan area. Atmos. Chem. Phys. 13, 2177–2194. https://doi.org/10.5194/acp-13-2177-2013.

Salamanca, F., Martilli, A., Tewari, M., Chen, F., 2011. A study of the urban boundary layer using different urban parameterizations and high-resolution urban canopy parameters with WRF. J. Appl. Meteorol. Climatol. https://doi.org/10.1175/2010JAMC2538.1.

Schubert, S., Grossman-Clarke, S., Martilli, A., 2012. A double-canyon radiation scheme for multi-layer urban canopy models. Boundary-Layer Meteorol. 145, 439–468. https://doi.org/10.1007/s10546-012-9728-3.

Shepherd, J.M., 2005. A review of current investigations of urban-induced rainfall and recommendations for the future. Earth Interact. https://doi.org/10.1175/EI156.1.

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.

Stretton, M.A., Morrison, W., Hogan, R.J., Grimmond, S., 2022. Evaluation of the SPARTACUS-urban radiation model for vertically resolved shortwave radiation in urban areas. Boundary-Layer Meteorol. 184, 301–331. https://doi.org/10.1007/s10546-022-00706-9.

Sützl, B.S., Rooney, G.G., van Reeuwijk, M., 2020. Drag distribution in idealized heterogeneous urban environments. Boundary-Layer Meteorol. https://doi.org/10.1007/s10546-020-00567-0.

Tang, Y., Lean, H.W., Bornemann, J., 2013. The benefits of the met Office variable resolution NWP model for forecasting convection. Meteorol. Appl. 20, 417–426. https://doi.org/10.1002/MET.1300.

Umweltatlas Berlin, 2010. Building and vegetation heights, p. 2009/2010.

United Nations, 2018. World Urbanization Prospects, p. 2018.

US Gazetteer, 2021. Urban Areas.

Wentz, E.A., et al., 2018. Six fundamental aspects for conceptualizing multidimensional urban form: a spatial mapping perspective. Landsc. Urban Plan. 179, 55–62. https://doi.org/10.1016/j.landurbplan.2018.07.007.

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. https://doi.org/10.1016/j.landurbplan.2017.06.018.

Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T.A., Taubenböck, H., 2019. Understanding an urbanizing planet: strategic directions for remote sensing. Remote Sens. Environ. 228, 164–182. https://doi.org/10.1016/j.rse.2019.04.020.