Distributed urban drag parameterization for sub-kilometre scale numerical weather prediction

Article

Published Version

Sützl, B. S. ORCID: https://orcid.org/0000-0001-9638-5643, Rooney, G. G. ORCID: https://orcid.org/0000-0002-3787-1198, Finnenkoetter, A. ORCID: https://orcid.org/0000-0003-0210-8288, Bohnenstengel, S. I. ORCID: https://orcid.org/0000-0001-6170-5774, Grimmond, C. S. B. ORCID: https://orcid.org/0000-0002-3166-9415 and Reeuwijk, M. v. ORCID: https://orcid.org/0000-0003-4840-5050 (2021)

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Quarterly Journal of the
Royal Meteorological Society

RMetS

Distributed urban drag parametrization for sub-kilometre scale numerical weather prediction

Birgit S. Sützl[1]* | **Gabriel G. Rooney[2,3]** | **Anke Finnenkoetter[2]** | **Sylvia I. Bohnenstengel[4]** | **Sue Grimmond[5]** | **Maarten van Reeuwijk[1]**

[1]Department of Civil and Environmental Engineering, Imperial College London, London, UK

[2]Met Office, FitzRoy Road, Exeter, UK

[3]School of Mathematics, University of Leeds, Leeds, UK

[4]MetOffice@Reading, Met Office, Meteorology Building, University of Reading, Reading, UK

[5]Department of Meteorology, University of Reading, Reading, UK

Correspondence
B. Sützl, Chair of Building Physics, ETH Zürich, 8092 Zurich, Switzerland.
Email: bsuetzl@ethz.ch

Present Address
*Chair of Building Physics, ETH Zürich, Zurich, Switzerland

Abstract

A recently developed, height-distributed urban drag parametrization is tested with the London Model, a sub-kilometre resolution version of the Met Office Unified Model over Greater London. The distributed-drag parametrization requires vertical morphology profiles in the form of height-distributed frontal-area functions, which capture the full extent and variability of building heights. London's morphology profiles are calculated and parametrized by an exponential distribution with the ratio of maximum to mean building height as the parameter. A case study evaluates the differences between the new distributed-drag scheme and the current London Model setup using the MORUSES urban land-surface model. The new drag parametrization shows increased horizontal spatial variability in total surface stress, identifying densely built-up areas, high-rise building clusters, parks, and the river. Effects on the wind speed in the lower levels include a lesser gradient and more heterogeneous wind profiles, extended wakes downwind of the city centre, and vertically growing perturbations that suggest the formation of internal boundary layers. The surface sensible heat fluxes are underpredicted, which is attributed to difficulties coupling the distributed momentum exchange with the surface-based heat exchange.

KEYWORDS

drag parametrization, MORUSES, regional and mesoscale modelling, urban canopy model, urban meteorology, urban morphology analysis

1 | INTRODUCTION

Urban environments alter aerodynamic, radiative, thermal, and hydrological processes, which can intensify heat waves, flash floods, and air pollution. Accurate urban models are necessary for better warnings of severe weather hazards and to improve weather forecasts and services in the most populated areas in the world. Moreover, urban climate models are crucial for planning how to adapt cities for more extreme weather and how to transform

them to be more sustainable and resilient in a changing climate.

Parametrizations for urban environments in numerical weather prediction (NWP) or climate models aim to represent the effects of the built-up system without resolving it explicitly. Numerous urban schemes at different complexities exist (Grimmond *et al.*, 2010). The schemes are commonly based on Monin–Obukhov similarity theory to calculate exchange fluxes between the atmosphere and the surface, where urban environments are represented by an increased roughness length that modifies surface friction and heat exchange. Single-layer canopy models are the most widely used type of urban scheme (Garuma, 2018). The urban morphology is represented in these schemes by a simple street-canyon geometry with variable height-to-width ratios, which parametrizes most physical processes in the urban canopy such as different flow regimes, shadowing, and radiation trapping (e.g., Porson *et al.*, 2010a). The urban environment is effectively seen as a "bulk surface", which interacts with the atmosphere close to the canopy top, at the level of the displacement height.

However, the use of these models is increasingly challenged by (a) the growing number of high-rise buildings in cities, which protrude deep into the atmospheric boundary layer, (b) increasing vertical and horizontal resolution of models (Lean *et al.*, 2019), so that the resolution scale becomes less than the surface-feature scale instead of vice versa, and (c) the heterogeneity of urban environments across scales, the effects of which become more apparent as grid resolutions increase (Barlow *et al.*, 2017). Individual tall buildings cause wake effects that are not represented by the idealised canyon geometry, and deep urban canopies, such as in central London or Asian megacities, displace and modify the airflow at heights well above the lowest atmospheric level (Hertwig *et al.*, 2021). As regional NWP models are increasingly used with sub-kilometre horizontal resolutions (e.g., Boutle *et al.*, 2016; Lean *et al.*, 2019), heterogeneous urban neighbourhoods, from high-rise building clusters to parks, become resolved by the individual grid boxes. Assumptions behind logarithmic wind profiles are no longer valid, as urban fetches become very short due to the increased heterogeneity and the flow is never adjusted fully with the underlying surface.

The vertical depth of the urban canopy and heterogeneity between the surface grid boxes can be addressed by multilayer canopy models that interact with the atmosphere over several vertical levels (e.g., Martilli *et al.*, 2002). The lowest atmospheric level is situated at ground level (without the displacement height) and wind profiles are modelled by a drag force instead of surface friction. Sützl *et al.* (2021) developed a height-dependent canopy model for urban drag based on large-eddy simulations of idealised heterogeneous urban neighbourhoods. Unlike most models, buildings are not represented by a simple street-canyon geometry but by a detailed vertical morphology profile, which considers the tallest buildings and the height variability between the buildings and therefore also captures the subgrid heterogeneity of the urban environment. This article describes a case study combining the distributed urban drag model from Sützl *et al.* (2021) with the Met Office Unified Model (UM: Davies *et al.*, 2005) in a high-resolution configuration with a limited domain over Greater London and surrounding areas.

This study describes a step towards developing a new multilayer scheme: at this point we only consider momentum exchanges, with a focus on the representation and effects of heterogeneous subgrid morphology. We describe in detail how the additional morphology information is acquired and processed. Section 2 describes how the distributed-drag model is incorporated into the UM and the setup for the case study. Vertical morphology profiles of urban areas are calculated and analysed from 1-m resolution building data, alongside updated urban morphology parameters of plan-area index λ_p (buildings plan area per grid-box plan area), frontal-area index λ_f (buildings frontal area per grid-box plan area), and mean building height z_H per grid box (Section 3). A model comparison between the standard model setup, the model with the updated urban morphology information, and the model including the distributed-drag scheme (Section 4) is conducted. The results are analysed in Section 5. A discussion on the implications of the distributed-drag model (Section 6) is followed by concluding remarks (Section 7).

2 | MODELS, OBSERVATIONS, CASE STUDY

2.1 | The London Model

The London Model (LM: Boutle *et al.*, 2016) is a very high resolution version of the Met Office Unified Model. The model domain covers a region of $125 \times 140 \, \text{km}^2$ and extends over Greater London and surrounding areas (Figure 1a) at a horizontal grid length of 0.003° (approximately 333 m). The initial and boundary conditions are from the operational Met Office forecast model for the United Kingdom (UKV) at 1.5-km horizontal resolution (Tang *et al.*, 2013). The vertical resolutions of the UKV and the London Model are identical, and both model convection explicitly. The numerical setup of the LM follows the RAL1-M science settings as described in Bush *et al.* (2020).

Urban areas in the UM are currently modelled by the Joint UK Land Environment Simulator (JULES: Best
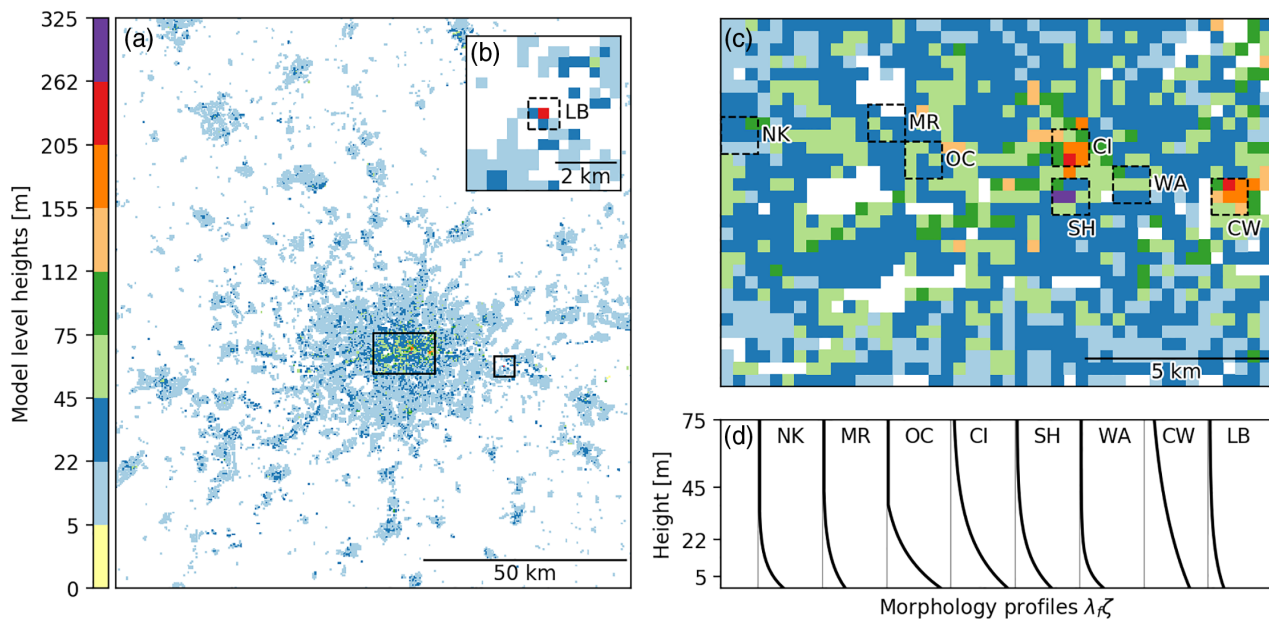
FIGURE 1 Highest vertical model-level heights where the drag parametrization scheme applies (grid boxes with $\lambda_p > 0.1$) for (a) full London Model domain, with rectangles to show the locations of (b) Thames estuary area, and (c) central London study area, with locations of the (d) vertical morphology profiles $\zeta(z)$, scaled by λ_f, from the centre grid box of the neighbourhoods. Notation of reference neighbourhoods (two letter codes) is defined in the text [Colour figure can be viewed at wileyonlinelibrary.com]

et al., 2011; Clark *et al.*, 2011), which parametrizes surface subgrid-scale processes and their exchange with the atmosphere. The model separates different surface types within a grid box into tiles (e.g., different vegetation types, water, soil), where subgrid processes are calculated in parallel for each tile with its own defined properties. The results of these calculations are combined into surface exchange fluxes by weighting the contributions from each tile by their grid-box fraction. The surface fluxes are coupled to the atmosphere implicitly to maintain balance between the land surface and the atmosphere (Best *et al.*, 2004). For the regional models UKV and LM, the calculation of urban properties within JULES is covered by the single-layer Met Office-Reading Urban Surface Exchange Scheme (MORUSES: Porson *et al.*, 2010a; 2010b; Bohnenstengel *et al.*, 2011; Bohnenstengel and Hendry, 2016). MORUSES represents urban areas using two tiles: one for the roof area of the buildings, and one for the street canyon. Material properties of building and road surfaces are fixed, but the street-canyon geometry may vary in each grid box. MORUSES calculates properties such as the heat capacity, albedo, and surface roughness length based on the varying geometry.

2.2 | Distributed urban drag model

A recently developed model for the contributions of urban areas to wind stress is based on building-resolving large-eddy simulations in neutrally stable conditions (Sützl *et al.*, 2021). The canopy model represents building effects by a height-distributed frontal area, which allows for a detailed representation of the vertical structure of the urban morphology. This representation is particularly important for areas containing high-rise buildings or grid boxes with large subgrid heterogeneity, because the tallest buildings impose a disproportionally large amount of drag on the flow (Xie *et al.*, 2008).

Sützl *et al.* (2021) defined a (dimensionless) normalised frontal area as

$$\zeta(z) = \frac{1}{A_F} \int_z^{z_{max}} L(z') \ dz', \tag{1}$$

which represents the fraction of the total frontal area above height z. Here, A_F is the total frontal area of all buildings in a grid box, and z_{max} is the height of the tallest building. The integrand $L(z)$ in Equation 1 represents the height-dependent total width of the buildings, which is the distribution of the surface area of buildings with height z, such that

$$A_F = \int_0^{z_{max}} L(z') \ dz'. \tag{2}$$

To keep the morphology representation relatively simple, $L(z)$ does not depend on wind direction but instead represents the average surface from all wind directions as discussed in Appendix A. The normalised frontal area has $\zeta = 0$ at the maximum building height and $\zeta = 1$ at ground
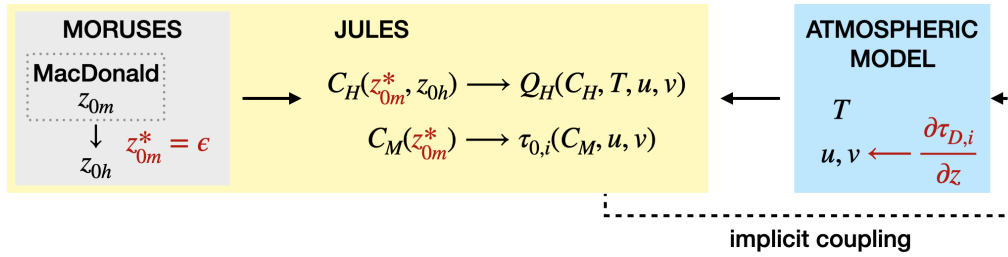
FIGURE 2 Interdependence of model parameters and modifications to the urban surface parametrization for including the distributed-drag scheme. The new model (colour highlight) replaces the momentum roughness length z_{0m} with $\epsilon = 10^{-5}$ m *after* the calculation of the heat roughness length z_{0h}. The notation is as follows: surface exchange coefficient for momentum C_M and heat C_H, surface sensible heat flux Q_H, surface momentum flux $\tau_{0,i}$, temperature T, horizontal velocities u and v, and vertical gradient of distributed urban drag $\partial \tau_{D,i}/\partial z$ with $i \in \{x, y\}$ [Colour figure can be viewed at wileyonlinelibrary.com]

level, and represents the urban morphology as a nonlinear function in between these heights (see Figure 1d for examples of λ_f-scaled $\zeta(z)$ profiles). Only for grid boxes with buildings of uniform heights will ζ depend linearly on z.

The model parametrizes the drag force exerted by buildings on the airflow as an explicit stress term that is added to the momentum equations. The urban drag at time t, height z, for each horizontal component $i \in \{x, y\}$, per grid box, is

$$\tau_{D,i}(t, z) = \tau_{0,i}(t)s(\zeta(z)), \qquad (3)$$

where $\tau_{0,i}(t)$ is the total urban canopy drag (or surface stress) component and s is the drag distribution function. That is, the distributed stress $\tau_{D,i}$, when normalised by the total surface stress $\tau_{0,i}$, collapses on to a single function s that depends on ζ only. Equation 3 was shown to hold by analysing the data from a series of large-eddy simulations of idealised urban morphologies, each with identical surface-cover parameters λ_p and λ_f, but with different building heights and street layouts (Sützl *et al.*, 2021). The drag distribution function s is represented reasonably well by the third-order polynomial

$$s(\zeta) = A\zeta^3 + B\zeta^2 + (1 - A - B)\zeta, \qquad (4)$$

with $A = 1.88$ and $B = -3.89$. The fact that s is not a linear function of ζ is testament to the earlier statement that taller buildings are responsible for a larger proportion of the drag.

In order to predict the total surface stress $\tau_{0,i}(t)$, we adapt the model from Coceal and Belcher (2004) and Belcher (2005) to the following form:

$$\tau_{0,i}(t) = \frac{1}{2}c_D \lambda_f \rho_0 U(t) u_i(t), \qquad (5)$$

with drag coefficient c_D, frontal-area index λ_f, air density ρ_0, horizontal wind velocities $u_i = u, v$, and wind speed $U = \sqrt{u^2 + v^2}$.

2.3 | Modifications to the urban surface scheme

Surface exchanges of momentum and heat between the land surface and the atmosphere are modelled in JULES on the basis of roughness lengths of momentum, z_{0m}, and heat, z_{0h}. The height of the surface exchange is assumed to be above the displacement height, although this is not represented explicitly by the surface scheme. The surface exchange of momentum, which parametrizes friction and aerodynamic effects of the surface elements, is modelled by (Best *et al.*, 2004; Lock *et al.*, 2020)

$$\tau_{0,i}(t) = C_M \rho_0 U(t) u_i(t), \qquad (6)$$

where C_M is the momentum-exchange coefficient, which incorporates the momentum roughness lengths of the surface tiles. The wind components are evaluated at the lowest model level height. The momentum roughness length of the urban tiles is calculated by MORUSES from λ_p, λ_f, and z_H, using the MacDonald *et al.* (1998) parametrization (Figure 2). The surface exchange of sensible heat, Q_H, is a function of the lowest model level winds, temperature differences between the surface and the lowest model level, and the surface exchange parametrization C_H, and has some dependence on atmospheric stability (Porson *et al.*, 2010a).

Figure 2 shows the different model components affected by the momentum-exchange scheme and highlights the modifications made to incorporate the distributed-drag scheme proposed in this study. The distributed canopy drag is included in the horizontal momentum equations via a body force given by the vertical gradient of Equation 3:

$$\frac{\partial \tau_{D,i}}{\partial z}(t, z) = \tau_{0,i}(t)\frac{ds}{d\zeta}[\zeta(z)]\frac{L(z)}{A_F}. \qquad (7)$$

This requires specification of $L(z)$ and z_{max}. The Unified Model has a parametrization for distributed

orographic drag (form drag induced by unresolved small-scale hills: Wood *et al.*, 2001), which was used as the basis for this implementation. Since the urban drag model does not account for the material effects of urban areas, MORUSES parametrizations are used for the radiation and heat exchange.

To prevent applying urban form effects twice, that is, via an effective roughness length and a distributed form drag, the momentum roughness length z_{0m} in MORUSES is reset to the value $\epsilon = 10^{-5}$ m representing a small material roughness length, so the JULES momentum exchange (Equation 6) becomes negligible. The roughness length for heat for urban surfaces, z_{0h}, remains unchanged as parametrized by MORUSES using the MacDonald *et al.* (1998) momentum roughness length and a resistance network (Harman *et al.*, 2004; Porson *et al.*, 2010a). We note that this setup causes an inconsistency for the heat-exchange coefficient C_H in JULES (see equation 19 in Porson *et al.*, 2010a), which depends on both z_{0m} and z_{0h}, and is calculated with the small momentum roughness length and the "normal" (MORUSES) heat roughness length. Unfortunately it is far from trivial to resolve this inconsistency in the code; however, its effects will be described in Section 5.3.

The urban drag parametrization and modifications to the momentum roughness length are applied to grid boxes with plan-area indices of $\lambda_p > 0.1$. This threshold includes suburban grid boxes at the periphery of Greater London and other large towns in the domain, but does not include individual buildings outside these urban patches. Figure 1 shows the highest (model-level) heights of the grid boxes in the London Model domain where the drag parametrization scheme applies, which is derived from the newly calculated λ_p and z_{max} data (see Section 3).

2.4 | Observations

Automatic Lidar and Ceilometer (ALC) observations from the London Urban Micrometeorological Archive (LUMA network: https://muhd.readthedocs.io) are used to support the model comparison. ALCs measure the light scattered back from aerosols. The distribution of aerosols in the atmosphere is a result of previous atmospheric mixing, and can therefore be used to characterise boundary-layer dynamics. Vertical profiles of aerosol backscatter are measured at the North Kensington (NK) and Marylebone Road (MR) sites in central London using Vaisala CL31 ceilometers. An automatic algorithm detects the mixed-layer height z_{ML} from the attenuated backscatter profiles (Kotthaus and Grimmond, 2018). Kotthaus and Grimmond (2018) report good agreement between z_{ML} inferred from backscatter profiles and the mixed-layer height inferred

from observed temperature inversions for clear-sky summer days. Therefore, the case study date was chosen accordingly.

2.5 | Case study

The case study period is June 26, 2018 with a 36-hr forecast run using the London Model starting from 25 June at 1800 UTC. This cloud-free period had easterly winds and an average wind speed of 5 m·s^{-1} (London City Airport, at 5 m above ground level (agl)). Analysis of the model behaviour will focus on a 10×15 km^2 area of central London (Figure 1c), which is densely built up and includes the historical centre and high-rise area of the City. The cluster of high-rise buildings of Canary Wharf is in the east. In the west there are three large parks: The Regent's Park, Hyde Park, and Battersea Park. Detailed analysis of 1 km^2 (3×3 grid boxes) neighbourhoods includes the following (see Table S1): Oxford Circus (OC), City (CI), Shard (SH), Wapping (WA), Canary Wharf (CW), and Littlebrook Power Station (LB). The first five are within the central London area. While the City and Canary Wharf neighbourhoods contain high-rise buildings above 200 m and the Shard (London's tallest building) is 304 m high, neither the Oxford Circus nor Wapping neighbourhoods has buildings above 80 m height. Wapping, unlike Oxford Circus, is outside the highest density areas of London, situated between the City and Canary Wharf. Littlebrook Power Station (Figure 1b), a former coal-fired power station in the Thames estuary with a 211-m high chimney, provides a reference for an isolated tall building in a low-density urban area.

3 | URBAN MORPHOLOGY OF GREATER LONDON

The London Model default urban morphology, described by the plan-area index λ_p, frontal-area index λ_f, and building mean height z_H, is insufficient to use the distributed-drag parametrization, as it requires vertical morphology profiles $\zeta(z)$ for each grid box. The urban morphology profiles of Greater London and its surroundings are therefore calculated using Ordnance Survey (OS) data of all the individual buildings in this area (Section 3.1). For consistency, λ_p, λ_f, and z_H are also updated (Appendix A). Analysis is conducted to parametrize $\zeta(z)$ based on the ratio of the maximum to mean building height $r = z_{max}/z_H$ (Section 3.2). The great advantage of this approach, apart from the insight it gives about the buildings in London, is that the only parameter needed in addition to those

present in the London Model is the maximum building height z_{max}.

3.1 | Urban morphology calculations

Urban morphology datasets for the London Model domain are derived using the Ordnance Survey MasterMap Topography Layer – Building Height Attribute data at 1 m resolution, updated in April 2019 (Ordnance Survey (GB), 2019). The London Model domain contains approximately 9.4 million buildings. Data for around 5% of the domain area are missing, all of which are outside Greater London. For these regions, the original London Model input data are used and the maximum height is assigned a value of $z_{max} = 2.3 z_H$, which is the average ratio of maximum to mean building height of the OS data in the whole LM domain, for this resolution. (It was found in a limited study that the average ratio in central London increases as horizontal resolution decreases, see Figure S2.)

A building in the OS MasterMap Topography Layer is represented by a polygon defining the footprint of the building and attributed with two heights: the maximum height and the principal height. The maximum height represents the highest point of the building extracted from the Digital Surface Model, relative to the height datum. The principal height captures the height of the main structure of the building excluding the roof. We approximate the three-dimensional building shapes by assuming a building area equal to the building footprint below the principal height, and a pyramid-shaped roof between the principal and maximum building height. The slope of the roof therefore depends on the height difference between these two parameters. To calculate the grid-box quantities, the following steps are taken for each building:

1. assign the building to a grid box(es), allowing it to be in multiple grid boxes;
2. calculate the building's plan area from the building footprint;
3. determine the wind-direction-averaged vertical building width function from the three-dimensional building shape (see Appendix A);
4. add the building's width function to the total width $L(z)$ and the building's plan area to the total plan area A_P in the appropriate grid box(es);
5. compare the building height with the current grid-box maximum and update z_{max} if necessary.

The grid-box profiles $\zeta(z)$ and parameters λ_p, λ_f, and z_H are determined after the calculations for all buildings (Appendix A).

3.2 | Parametrizing London's morphology profiles

To reduce the complexity of the required model input, the vertical profiles $L(z)$ and $\zeta(z)$ are parametrized in terms of the mean building height z_H and maximum building height z_{max} of the grid box. It is possible to compare the vertical functions directly amongst the grid boxes by a change of variables $\hat{z} = z/z_{max}$, such that the rescaled and normalised $\hat{L}(\hat{z})/A_F$ functions and $\zeta(\hat{z})$ are defined on the interval $[0, 1]$. The $\hat{L}(\hat{z})/A_F$ profiles are categorised according to the ratio of the maximum to the mean building height of the corresponding grid box, $r = z_{max}/z_H$ (cf. Figure S3).

The data are split into six bins: $r = 1$, $r \in (1, 2]$, $r \in (2, 3]$, $r \in (3, 4]$, $r \in (4, 5]$, $r > 5$. The first, $r = 1$, classifies grid boxes with uniform buildings, where the maximum height is equal to the mean height. Less than 1% of all London Model domain grid boxes fall into this category, and they have very low building densities ($\lambda_p < 0.02$), usually with only one or two buildings in the area. Approximately half of the grid boxes have a building-height ratio $r \in (1, 3]$; the remaining three categories contain the other 10% of the grid boxes that include buildings; overall, approximately 40% of the grid boxes do not include any buildings. The average building-height ratio is $r = 2.3$. A high r is generally associated with a large heterogeneity in building heights. For example, grid boxes with high-rise building clusters like those in the London City (CI, Figure 1c) have typical ratios of 6–8. The Shard (within SH, Figure 1c), the highest building in the domain, is situated on the border of two grid boxes, which have values of $r = 15$ and $r = 17$. Some sparsely populated grid boxes with a single tall structure, for example, a transmission tower, also fall into the $r > 5$ category.

Averaging the functions $\hat{L}(\hat{z})/A_F$ for each of the bins yields six distinct profiles $\langle \hat{L} \rangle(\hat{z})/A_F$ (Figure 3a). The functions integrate to unity and therefore resemble a distribution function. The data fit well to an exponential distribution that is modified to be non-zero on the interval $[0, 1]$:

$$\frac{L}{A_F}(\hat{z}; \alpha) = \begin{cases} \frac{\alpha e^{-\alpha \hat{z}}}{1 - e^{-\alpha}}, & \hat{z} \in [0, 1], \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

with $\alpha > 0$. Empirical values for α are derived from a curve fit of the six averaged profiles to Equation 8. Afterwards these empirically fitted coefficients $\tilde{\alpha}$ are related to the building-height ratio r. The values $\tilde{r} = 1, 1.5, 2.5, 3.5, 4.5, 6$ are chosen to represent the bins. Linear regression between $\tilde{r}$ and the empirical values $\tilde{\alpha}$

FIGURE 3 Parametrizing total width profiles $L(z)$ as a function of building-height ratio $r = z_{max}/z_H$. (a) Normalised and averaged total width functions $\langle \hat{L} \rangle (\hat{z})/A_F$. (b) Linear regression between bin values $\tilde{r}$ and estimated $\tilde{\alpha}$ from curve fitting $\langle \hat{L} \rangle (\hat{z})/A_F$ to $L(\hat{z}; \alpha)/A_F$. (c) The functions $L(\hat{z}; \alpha(r))/A_F$ parametrizing the total width functions. Note that the normalised height $\hat{z}$ in (a) and (c) is shown on the horizontal axis [Colour figure can be viewed at wileyonlinelibrary.com]
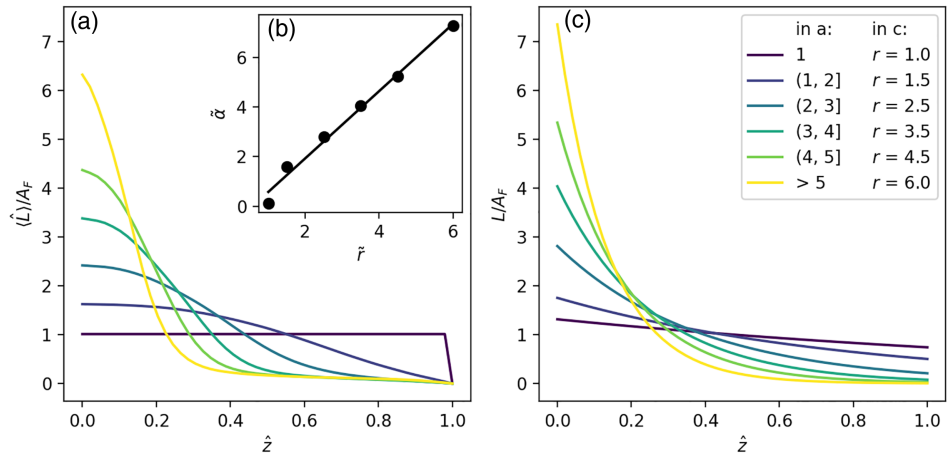
gives the following approximation:

$$\alpha(r) = 1.355r - 0.7807. \tag{9}$$

The fit of the empirical coefficients to the building-height ratio works remarkably well (Figure 3b). The resulting profiles $L/A_F(\hat{z}; \alpha(r))$ are estimated entirely based on r (Figure 3c). The functions capture well the gradual change from a uniform profile ($r = 1$) to one that has its weight at lower $\hat{z}$ values (i.e., relatively closer to the ground) with increasing r values. Repeated analysis with a larger number of bins yields very similar results, which suggests a robust relation between the building-height ratio r and the shape of the vertical functions $L(z)$. The averaged morphology profiles also remain invariant over different grid resolutions in a limited analysis of central London, suggesting the parametrization may be suitable across resolutions (Figure S2). The parametric function in Equation 8 is therefore an appropriate parametrization for the morphology profiles.

Substitution of Equation 8 into Equation 1 results in an analytical expression for the normalised frontal area:

$$\zeta(z) = \begin{cases} \frac{1 - e^{\alpha(1 - z/z_{max})}}{1 - e^{\alpha}}, & z \le z_{max}, \\ 0, & z > z_{max}, \end{cases} \tag{10}$$

with $\alpha = 1.355z_{max}/z_H - 0.7807$. Figure 1d illustrates $\lambda_f \zeta(z)$ profiles calculated with Equation 10 for selected neighbourhoods (Sections 2.4 and 2.5). The mean and maximum building height are obtained from the central grid boxes of the neighbourhoods. This scaled quantity illustrates the frontal-area index above height z (with λ_f at ground level), which gives an indication of the magnitude of the stress applied at each height.

4 | MODEL COMPARISON

4.1 | Description of model configurations

Three different model runs are compared: (a) the *original* London Model with the current experimental configuration, (b) the *control* run with updated morphology input fields λ_p, λ_f, and z_H derived from the 1-m resolution Ordnance Survey data, and (c) the *distributed drag* model run with the vertical parametrization of urban drag. Comparisons between (a) and (b) assess the effects of updated urban morphology inputs in MORUSES. Note that the land-cover fractions of the grid boxes have not been assessed (since the analysis did not include vegetation and other surface types) and the control run therefore does not represent a fully updated model configuration. This implies that changes to the plan area of buildings and aspect ratio of street canyons are not reflected in the tile weighting in JULES.

Comparisons between (b) and (c) assess the effects of distributing urban drag over several vertical levels. The simulations are also used to assess spatial variability, since urban drag is modelled differently between the two runs. 14% of all grid boxes in the London Model domain are above the threshold $\lambda_p > 0.1$, where the effective roughness formulation is replaced with the distributed-drag scheme (Figure 1a). Most commonly, the distributed-drag scheme affects the lowest two or three vertical model levels (up to 22 and 45 m agl, respectively), whereas in central London (Figure 1c) the scheme frequently exerts drag on the fourth model level (up to 75 m agl) and occasionally higher. It reaches the ninth vertical level around the tallest building in the domain.

Note that, in the comparison between (a), (b), and (c), the output variables are compared at equal vertical model levels. Although the bottom of the atmospheric model in

MORUSES is conceptually situated above the displacement height (not above the ground, Hertwig *et al.*, 2020), the model does not actually adjust the surface heights of the individual grid boxes, and the displacement heights therefore have no influence on the surface-layer dynamics. This inconsistency in the typical modelling approach makes it difficult to interpret the model output close to the surface. In the distributed-drag scheme, the bottom of the atmospheric model is at ground level, the building parametrization is immersed in the atmosphere, and the lower levels therefore represent profiles within the urban canopy.

The output variables compared are the following: the magnitude of the surface stress, $\tau_0(t, x, y) = \sqrt{\tau_{0,x}^2(t, x, y) + \tau_{0,y}^2(t, x, y)}$; the surface sensible heat flux $Q_H(t, x, y)$; the wind speed $U(t, x, y, z)$ at all vertical model-level heights; and the mixed-layer height $z_{ML}(t, x, y)$. The outputs of τ_0, Q_H, and U at 10 m are averaged over 15-min time intervals. Wind speed at other heights and z_{ML} are instantaneous outputs. A 24-hr time series is analysed for each model run, which discards the first 6 hr as spin-up. Outputs for the measurement sites (Section 2.4) and 1-km^2 neighbourhoods (Section 2.5) are analysed as 3×3 grid-box averages. The central grid box contains the point of interest (i.e., measurement site or landmark). Note that the data location varies slightly by output variable, as the Unified Model uses a staggered Arakawa-C grid (Davies *et al.*, 2005). Space averages over central London (Figure 1c) are denoted by $\langle \cdot \rangle$, 24-hr time averages by $\overline{\cdot}$.

4.2 | Matching the drag coefficient

The aim of this study is to compare the effects of distributing drag at similar levels of total urban drag, despite the different ways of obtaining $\tau_{0,i}$ in MORUSES and the distributed-drag scheme. For this purpose, the drag coefficient c_D of the distributed-drag model is set so the surface stress magnitude τ_0 matches between the control and distributed-drag run in the central London study area. A suitable c_D is estimated using a least-squares fit for

$$c_D = \frac{2\langle \tau_0 \rangle}{\rho_0 \langle \lambda_f U^2 \rangle}, \tag{11}$$

with the central London space-averages $\langle \tau_0 \rangle(t)$ and $\langle \lambda_f U^2 \rangle(t)$ from the control run for the 24-hr period. The wind speed for the least-squares fit and in the distributed-drag scheme (Equation 5) is evaluated at the fifth model level ($z = 93$ m). Since the distributed drag is expected primarily to affect the low-level winds, this

height is sufficiently above most buildings to avoid interference of the drag scheme with the reference wind speed.

When a test run with the resulting drag coefficient $c_D = 0.067$ found a space- and time-averaged surface stress over the central London area $\langle \overline{\tau_0} \rangle$ slightly exceeding that of the control run, it was changed to $c_D = 0.06$. This yielded an almost exact match, where $\langle \overline{\tau_0} \rangle$ of the control and distributed-drag run differ by less than 0.01% (for comparison, $\langle \overline{\tau_0} \rangle$ of the original run is about 13% lower). The space-averaged surface stresses are also similar over time, as shown by the $\langle \tau_0 \rangle(t)$ time series for the three model runs (Figure 6a). Note that using Equation 11 with the wind speed at the lowest level height ($z = 2.5$ m) yields $c_D \approx 1$. This indicates the close link between surface winds and boundary-layer stresses through the implicit coupling of JULES and the atmospheric model component.

4.3 | Changes in input morphology

The urban morphology input data for the Unified Model are from the Institute of Terrestrial Ecology (ITE: Bunce *et al.*, 1990). The plan-area index, frontal-area index and mean building height are derived from the "urban" (i.e., impervious) land-cover fraction F_U of a grid box using empirical relations (Bohnenstengel *et al.*, 2011; Bohnenstengel and Hendry, 2016). The relations are derived from high-resolution building data of London (Evans *et al.*, 2006) at 1 km grid resolution. With λ_p, λ_f, and z_H all functions of F_U, the three parameters are linked and reduced to one degree of freedom. The original run uses these morphology datasets. The new control-run datasets are derived directly from the 1-m resolution Ordnance Survey building data.

Figure 4 shows the frequency distributions of the original and new morphology grid-box values λ_p, λ_f, z_H, and z_{max} in central London. The frequency distributions of the plan-area index λ_p (Figure 4a) are largely similar between the original and new data, with differences only evident in grid boxes with $\lambda_p < 0.1$ and around the mean value (original = 0.3; new = 0.28). The maximum value increases slightly from 0.61 to 0.66. The frequency of low λ_p grid-box values is resolution-dependent (higher resolutions increase the number of grid boxes that contain very few buildings, cf. Figure S4). Using the 1-km based empirical relations on the London Model grid therefore underestimates grid boxes with very low λ_p in the original data. Spatially, the original and new plan-area indices of central London appear relatively similar overall (Figure S5). The new data have a larger cluster of high-density grid boxes north of the river Thames; the original data have more individual grid boxes
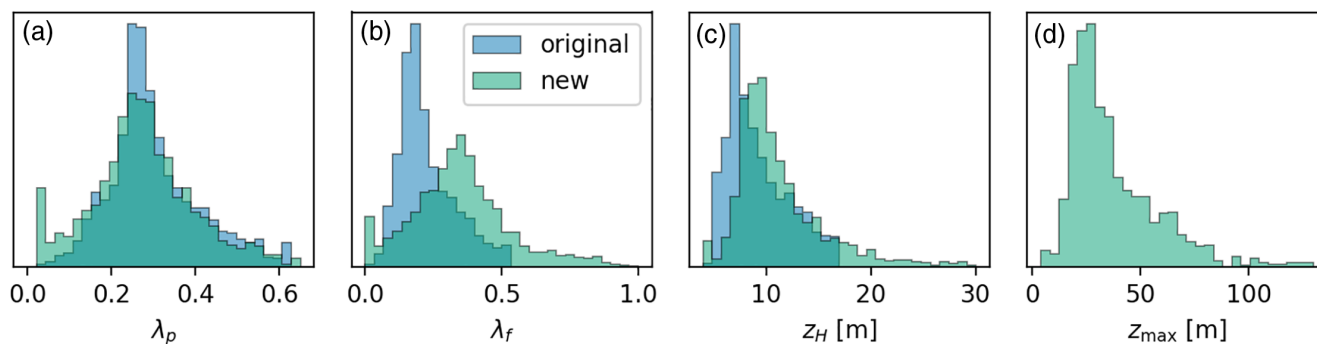
FIGURE 4   Frequency distributions of original and new urban morphology grid-box values in central London for (a) plan-area index λ_p, (b) frontal-area index λ_f, (c) mean building height z_H, and (d) maximum building height z_{max} (new data only) [Colour figure can be viewed at wileyonlinelibrary.com]

with high λ_p values outside the most densely built-up area.

The new frontal-area index distribution (Figure 4b) differs from the original data distribution, with a larger mean (original = 0.22; new = 0.35), the maximum reaching 1.2 (0.5 previously), and a much broader distribution of values. The new frontal area indices are higher overall, and much higher in the most densely build-up areas north of the river and at the high-rise buildings of Canary Wharf (Figure 5). Both datasets show low frontal-area indices along the river and at the three large parks in the west. The higher overall values of the new data may partly be attributed to the definition we applied for the frontal-area index, which estimates higher frontal areas than other methods (see Appendix A); however, the data also clearly reflect the vast increase in high-rise buildings in London over the last decades.

The average value for the mean building height z_H increases from 9 to 11 m with the new data (Figure 4c). The distribution has a longer tail, with more grid boxes with higher z_H in the new dataset. The maximum z_H is 59 m (up from 17 m!). Decoupling the mean-height calculations from the urban land-cover fraction results in considerable spatial differences between the original and new z_H data (Figure S6). The new data indicate three areas with high average building heights (Figure 1c shows locations): the City (area around CI), Canary Wharf (around CW), and a narrow band southwest of the centre. This area contains the Battersea Power Station development site, which had few completed buildings in 2019. Unlike the original data (linked to impervious fraction), the river is not evident in the new z_H, as these grid boxes have both river and buildings, which the z_H map now reflects. The distribution of the new input parameter, maximum building height z_{max}, is of comparable shape to the z_H distribution of the new data (Figure 4d). The mean is 39 m and the maximum is 304 m.
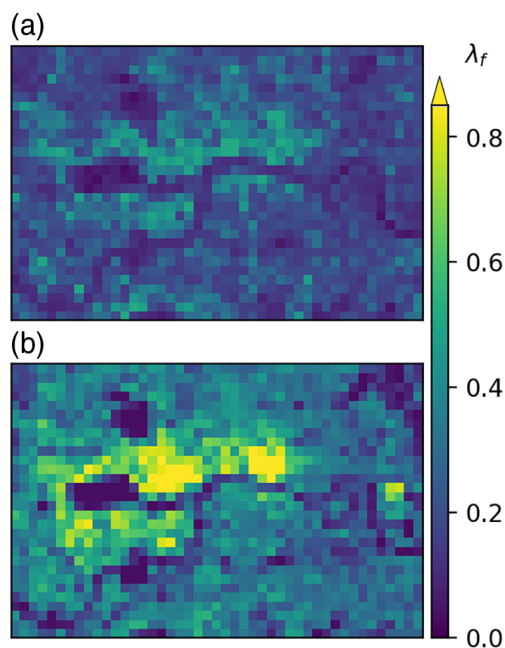

FIGURE 5   Frontal-area index λ_f in central London: (a) original and (b) new datasets. For spatial scale see Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]

5 | CASE STUDY RESULTS

5.1 | Surface stresses

Figure 6 shows the magnitude of the surface stress τ_0 for the central London area. The space-averaged $\langle\tau_0\rangle(t)$ are largely similar over time for the three model runs, but the surface stress of the original run is consistently less than the other two model runs (Figure 6a). Spatially, the surface stresses have large variations over central London for the 24-hr time-averaged $\overline{\tau_0}(x, y)$ fields (Figure 6b–d). In the original run (Figure 6b), surface stresses across urban areas are relatively uniform and the spatial structure

of London is only visible through the lack of built-up areas (i.e., large parks and the river). High surface stresses in the control run (Figure 6c) occur in Canary Wharf. Higher stresses than in the original run also occur in areas with comparatively low building density in southwest London. Open areas such as the river are not evident. The distributed-drag run (Figure 6d) has high surface stresses for Canary Wharf, the City, and over the most densely built-up area in the centre of London. Distinctively low stresses are found at the parks and river. Overall, the spatial variability of surface stresses is largest with the distributed-drag model, highlighting various local features, such that the spatial patterns of τ_0 reflect the heterogeneity of the urban area.

The spatial variability between the different model runs stems from changes in morphology inputs and the different parametrizations for the surface stress equations (Equations 5 and 6). The surface stress of the distributed-drag run clearly correlates with the frontal-area index λ_f (cf. Figure 5b), which results from incorporating λ_f in the distributed-drag model (Equation 5). The surface stresses of the original and control run are closely linked to the momentum roughness length z_{0m}, which is the key element of the standard momentum-exchange parametrization (Equation 6). The roughness length, calculated in MORUSES with the MacDonald *et al.* (1998) parametrization, is most strongly

correlated with the mean building height z_H. The MacDonald *et al.* parametrization derives z_{0m} as a fraction of z_H, where λ_p and λ_f affect the proportionality. The ratio of z_{0m} to z_H is higher for low values of λ_p and λ_f, or low values of λ_p and high values of λ_f.

The original dataset values of z_{0m} have little spatial variability (Figure S7). Non-urban features such as the parks and the river are most distinct. Roughness lengths from the new morphology inputs are higher and values have a wider range. In Canary Wharf, where z_H and λ_f are high and λ_p relatively low, the highest z_{0m} is 12 m. Roughness-length values that are ten times higher than typical for urban areas seem to be a gross overestimation, although there are almost no measured values in high-rise areas to say this with any certainty (Grimmond and Oke, 1999). High z_{0m} values also occur at the Battersea development site, despite low building densities (derived from high z_H, low λ_p, λ_f). The correlation with z_H causes higher roughness lengths along the Thames river, such that the river is no longer distinct from other neighbourhoods.

The profound differences in surface stress between the runs demonstrate the importance of high-quality and up-to-date morphology data. However, the control-run τ_0 fields show that more realistic mean building heights do not necessarily improve the parametrization. The data suggest a limit on the grid length in MacDonald *et al.* (1998), such that the roughness-length parametrization
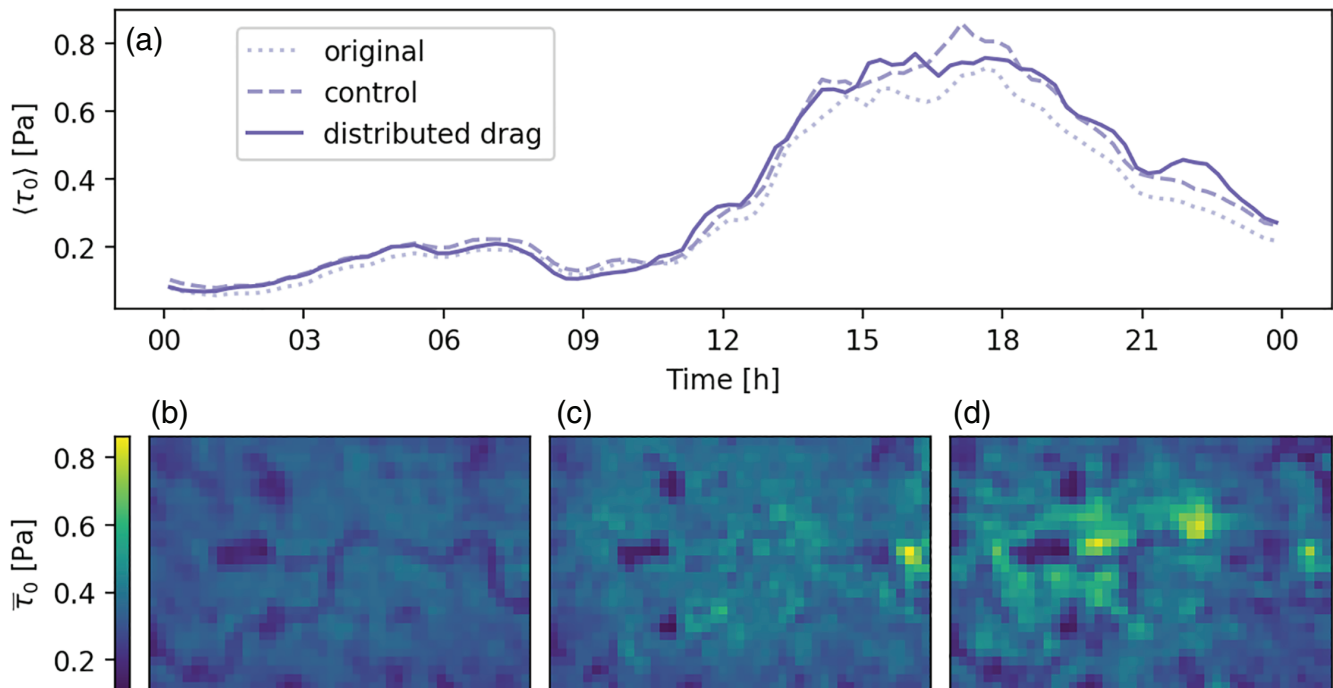

FIGURE 6 Central London's surface stress τ_0 with (a) space-averaged $\langle \tau_0 \rangle(t)$ time series from original (dotted line), control (dashed line), and distributed-drag run (solid line), together with 24-hr time-averaged $\overline{\tau}_0$ fields with (b) original, (c) control, and (d) distributed-drag run. For spatial scale see Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]

is less appropriate for high-resolution models like the LM.

5.2 | Wind speed

Vertical profiles of wind speed $U(z)$ at 1900 UTC for the reference neighbourhoods illustrate representative features of the different model runs (Figure 7), because vertical mixing through convection is lower than during times with stronger solar radiation. The wind profiles of the original and control run are relatively similar in all neighbourhoods and have the typical boundary-layer profile of air flowing over a flat surface. The wind-speed profiles with the distributed-drag scheme are much more diverse. The City and Oxford Circus differ the most in the lower model levels compared with the other neighbourhoods. The distributed-drag wind profiles change gradually in height, as the wind speeds in the lower vertical levels are lower. Wind speeds become similar to the other two model runs at several hundred metres above the ground. Significant differences between the original and control run are only found for Canary Wharf, where the much higher τ_0 in the control run is reflected in smaller wind speeds in the lower model levels. The effects of height distribution are particularly evident for this neighbourhood, because the distributed-drag wind speed is higher than in the control run near ground level, but below the control run between approximately 50 and 150 m. All model runs are similar for the isolated tall chimney of the Littlebrook Power Station.

The frontal-area index function $\lambda_f \zeta(z)$ for the centre grid box of each neighbourhood illustrates the different underlying morphology profiles and stress magnitudes (via the λ_f-scaling) used in the drag parametrization (Figure 7). The effects of height-distributed drag can go beyond the height of the parametrization ζ, as the Oxford Circus case illustrates. The maximum building height in the neighbourhood is 62 m (37 m in the central grid box), but the distributed-drag wind speed does not approach the speed of the other model runs until 250 m model height. The high-density urban surroundings of Oxford Circus are likely contributing to the large effects on the wind.

Indeed, the effect of large-scale urban areas under the distributed-drag scheme is evident in Figure 8, which shows 24-hr time-averaged wind-speed fields $\overline{U}(x, y)$ at 10-m model height for central London. The distributed-drag run shows relatively high average wind speeds to the east side of the study area (Figure 8c), where easterly winds with high wind speeds along the Thames estuary approach central London. A large wake with low average wind speeds is found downwind of the city centre to the west. This spatial pattern suggests that
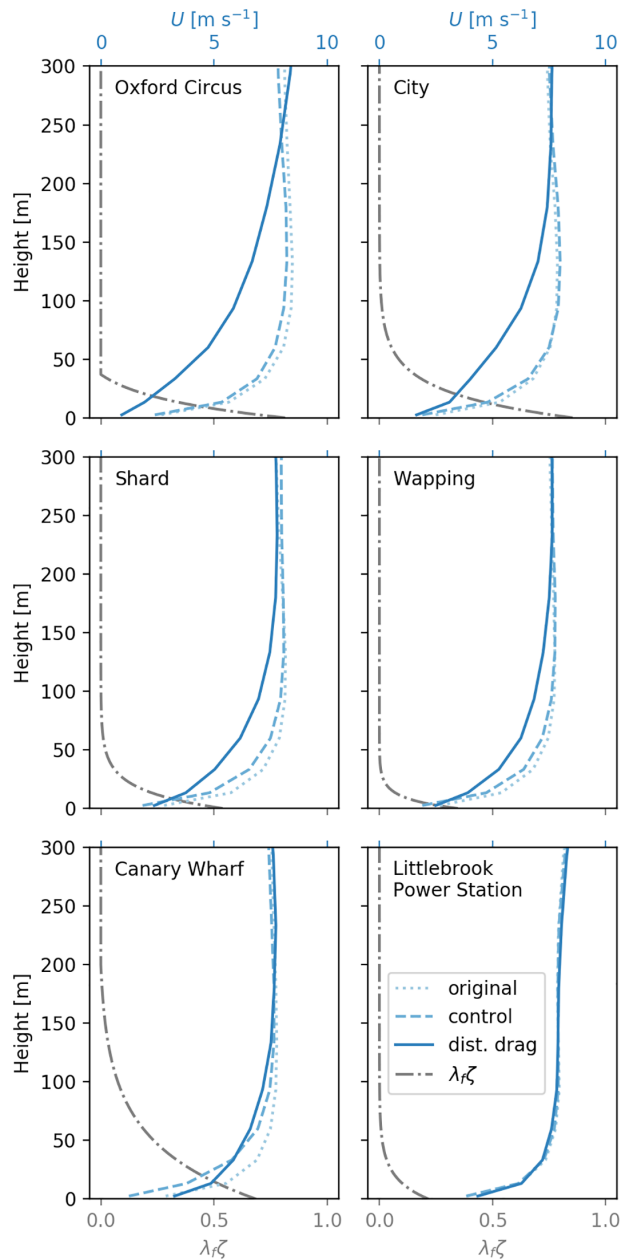

FIGURE 7 Vertical profiles of wind speed $U(z)$ at 1900 UTC for selected neighbourhoods (Figure 1) with original (dotted line), control (dashed line), and distributed-drag run (solid line). The profiles are 3×3 grid-box averages. Frontal-area index functions $\lambda_f \zeta(z)$ (dash–dotted line) are shown for the centre grid box of the neighbourhoods. Height indicates atmospheric model height, which is agl for the distributed-drag run and $\lambda_f \zeta(z)$, and above displacement height for the original and control run. [Colour figure can be viewed at wileyonlinelibrary.com]

local disturbances introduced by the distributed drag are advected and grow vertically in the downstream direction, forming an internal boundary layer. Therefore, with an easterly wind, a neighbourhood downwind of the city centre such as Oxford Circus experiences large impacts of the distributed-drag model.
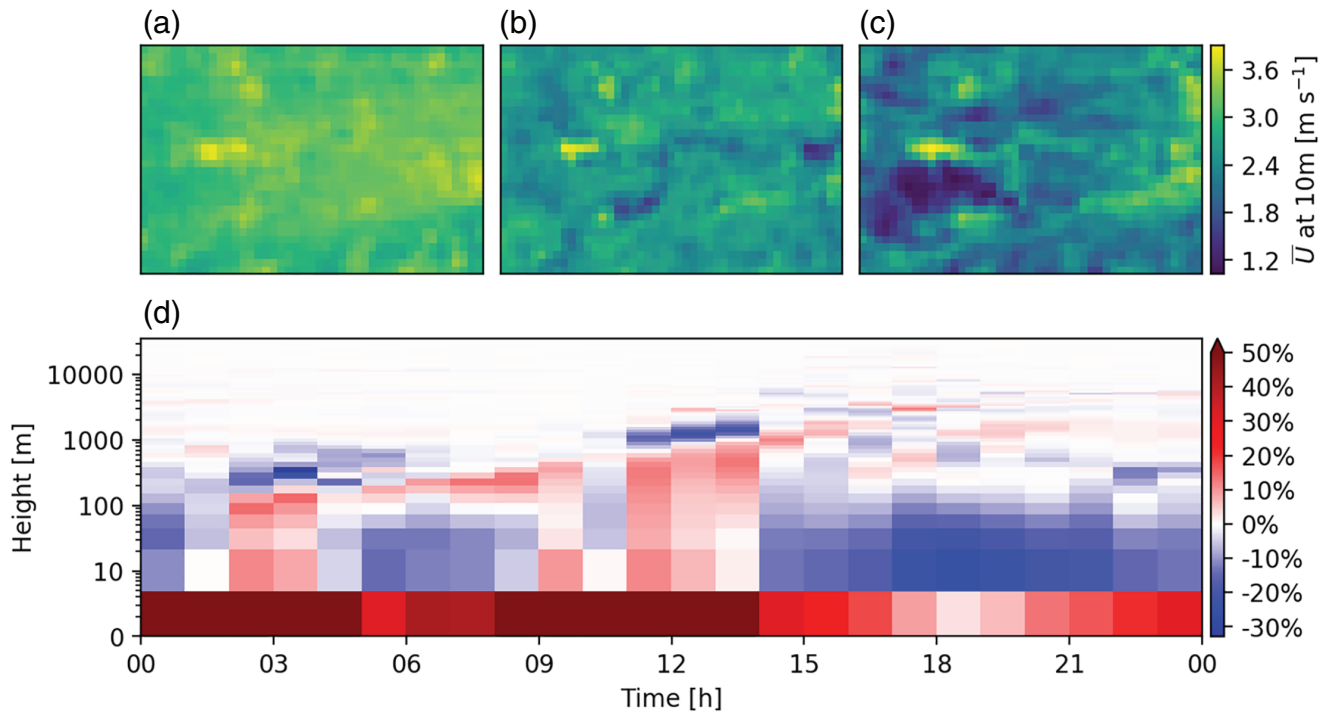
FIGURE 8 Central London 24-hr time-averaged wind speed $\overline{U}(x,y)$ at 10 m model height for (a) original, (b) control, and (c) distributed-drag run (for spatial scale see Figure 1), and (d) change in central London space-averaged wind speed $\langle U \rangle(t,z)$ between control and distributed-drag run. Note that the height axis is logarithmic (except near the surface) and the colour scale is capped at 50%, which is exceeded in the lowest model height level [Colour figure can be viewed at wileyonlinelibrary.com]

In contrast, the spatial patterns of the original and control run largely resemble the patterns of the surface stresses (cf. Figure 6). The original run (Figure 8a) has higher average wind speeds over the parks and river, where friction is smaller, and the control run (Figure 8b) has particularly low wind speeds around Canary Wharf and the southwest, where friction is high (cf. Section 5.1). However, there is little evidence in the 10-m wind field of any extended wake regions, or the growth of internal boundary layers.

Figure 8d shows the change between the control and distributed-drag run for space-averaged wind speed $\langle U \rangle(t,z)$ in central London, giving a direct comparison between the standard parametrization and the distributed-drag scheme. The standard approach produces the desired effects on wind speeds by slowing down the wind speed at the lowest level and relying on diffusion to adjust the wind speed in the adjacent levels. With the building drag distributed over the full depth of the canopy, wind speeds are considerably higher at the lowest model level (often by more than 50%), but tend to be lower for model heights just above the lowest level and up to several hundred metres. A band of higher wind speeds on top of lower winds in the morning hours suggests a lower boundary-layer height for the distributed-drag run at these times. This is probably related to a reduction in

the sensible heat flux (Section 5.3). At other times, there is no clear trend above the lowest few hundred metres and the relative difference becomes smaller at higher levels. This suggests that the distributed-drag forcing can produce the same effects in the upper levels, and at the same time produce wind profiles within the canopy that may be more realistic (or comparable with building-resolving simulations).

5.3 | Mixed-layer height

Roughness and thermal effects of urban areas generate turbulence, which induces vertical mixing in the atmosphere. The height of the mixed layer z_{ML} above cities is therefore an indication of how much turbulence is generated by surface processes. The model mixed-layer height is determined as the maximum of the surface-based mixed-layer height, which is the depth through which a positively buoyant parcel released at the surface would ascend, and the boundary-layer depth diagnosed by a threshold for the Richardson number, which marks the transition to the stable conditions of the capping inversion (Lock *et al.*, 2020). Model mixed-layer heights are evaluated against mixed-layer heights estimated from ALC observations of aerosol-backscatter data. The observed
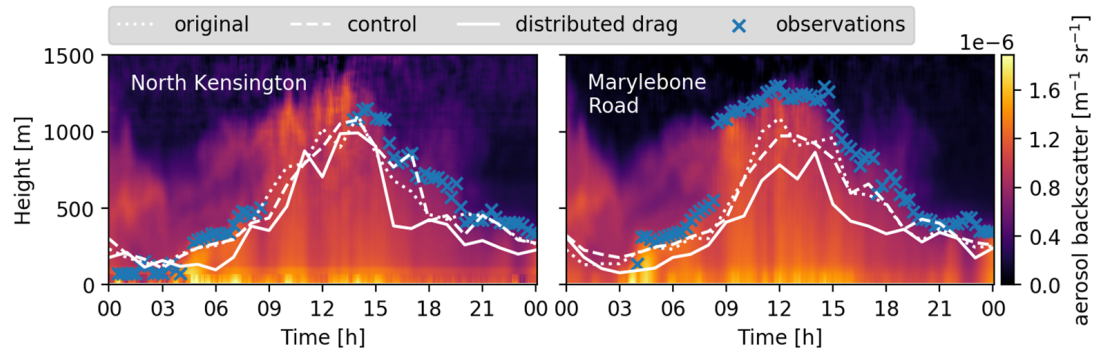
FIGURE 9   Aerosol backscatter profiles and mixed-layer height observed at the measurement sites (markers) and hourly model output by the original (dotted line), control (dashed line), and distributed-drag run (solid line). The model data are 3×3 grid-box neighbourhood averages around the measurement site (Figure 1c) [Colour figure can be viewed at wileyonlinelibrary.com]
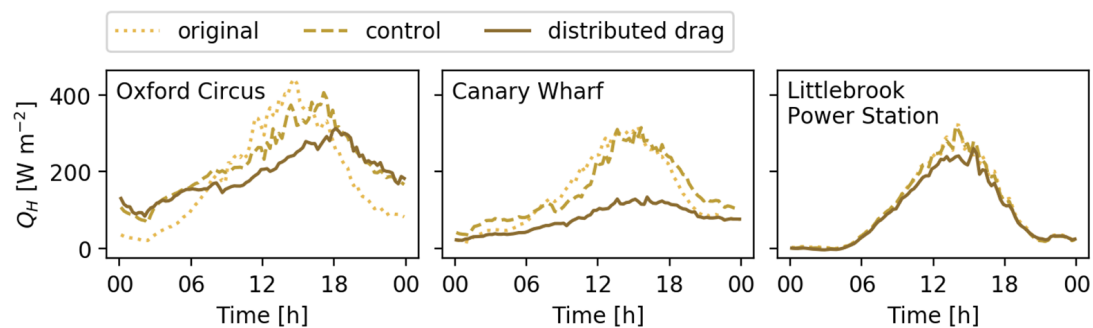

FIGURE 10   Surface sensible heat fluxes $Q_H(t)$ of the neighbourhoods Canary Wharf, Oxford Circus, and Littlebrook Power Station for original (dotted line), control (dashed line), and distributed-drag run (solid line). The time series are 3×3 grid-box neighbourhood averages [Colour figure can be viewed at wileyonlinelibrary.com]

mixed-layer heights at the two measurement sites in central London are largely similar in the morning and afternoon, and differ somewhat at their maximum height at midday. Reasonably good agreement between the observed and modelled mixed-layer heights is found, but the observations suggest that all simulations underestimate z_{ML} (Figure 9). However, z_{ML} values of the distributed-drag run are below those of the original and control run for the day investigated in this study. A typical error estimate for a measurement is around ± 10 m, thus much smaller than the difference between observations and modelled z_{ML}.

The lower mixed-layer height in the distributed-drag run is likely caused by an underprediction of the surface sensible heat flux Q_H. Recall that Q_H is modelled as a function of the lowest model level winds, temperature differences, and the surface heat-exchange coefficient C_H. The distributed-drag scheme yields higher velocities than the standard MORUSES parametrization at the lowest level; however, the exchange coefficient C_H is based on the heat roughness length and the altered momentum roughness length $z_{0m} = \epsilon$ (cf. Figure 2), and C_H is therefore much lower than in the control run. As a result,

surface sensible heat fluxes throughout the day are greatly reduced compared with the standard scheme (Figure S8). For example, Q_H at Canary Wharf (similarly Shard and Wapping neighbourhoods, not shown) is reduced to half at midday (Figure 10). A smaller but significant reduction is also found at Oxford Circus (similarly the City, not shown). Night-time surface sensible heat fluxes are more alike and can be even slightly higher in the distributed-drag run (e.g., at Oxford Circus).

6 | DISCUSSION

The results indicate that incorporating distributed drag has a profound effect on the wind in the lower atmospheric levels. The distributed-drag scheme causes much more variation in the wind speed over central London than the classical parametrizations. The pressure drag generated by clusters of tall buildings causes extensive wakes that reduce wind speeds over extended areas downstream.

However, the mixed-layer height was underestimated with the distributed-drag model, which was the result of

the interaction of the distributed-drag scheme with the sensible heat flux prediction from JULES. This is a fundamental challenge for the next generation of land-surface schemes, since momentum and scalar (heat, moisture) exchange are intrinsically coupled; modelling drag in a distributed manner and heat solely at ground level is bound to lead to inconsistencies. Note that a similar conclusion was drawn for the distributed orographic scheme in a study on inland water (Rooney and Bornemann, 2013): while distributed orographic drag produces more realistic low-level winds than the orographic roughness scheme, it also produces unrealistic lower temperatures with the current model physics settings in JULES.

In order to obtain accurate predictions of sensible and latent heat fluxes with the current model, a recalibration of the scalar exchange coefficients would be required. However, the fundamental mismatch of having part of the physics represented in a distributed manner inside the model and part of the physics represented at ground level as a boundary condition to the model would remain. A "complete" multilayer model for urban areas will need to represent both the drag and scalar exchange within the atmospheric model in a distributed manner.

7 | CONCLUSIONS

We present results from a trial implementation of a height-distributed urban drag scheme in the London Model high-resolution numerical weather prediction model. The distributed-drag scheme (Sützl *et al.*, 2021) relies on a generalised frontal area $\zeta(z)$ to represent the urban morphology as a function of height. Analysis of the $\zeta(z)$ profiles for the Greater London area showed that ζ could be represented by an exponential distribution with the ratio between the maximum and mean building height (z_{max}/z_H) as the independent variable, and the model therefore only requires z_{max} as additional input. This is remarkable, as it is not evident why London's morphology would be described by such a distribution. It would be intriguing to find out whether this distribution is specific to London or whether this is a universal relation amongst cities. The analysis also revealed that there are very few grid boxes that contain entirely uniform buildings, despite many studies of urban airflow and other processes being based on these idealised building forms. The lack of real morphology data hampers weather and climate predictions globally, as well as the delivery of integrated urban services to city residents.

A comparison between the standard model configuration (with original and updated morphology inputs) and the model with the distributed-drag parametrization shows considerable differences in the lower levels of the atmospheric boundary layer. Spatial variations in the surface stresses highlight the importance of the urban morphology inputs and modelling decisions. Original plan-area index, frontal-area index, and mean building height all depend on the impervious land-cover fraction in each grid box. After decoupling these parameters, regions with high-rise buildings experience higher surface stresses, but other local distinctions like the Thames river disappear, putting the applicability of roughness-length-based parametrizations like MacDonald *et al.* (1998) at such high resolutions into question. The distributed-drag parametrization uses a bluff-body modelling approach (i.e., considering frontal areas) and the resulting surface stresses show the greatest spatial variability: local features such as densely built-up areas, high-rise building clusters, parks, and the river are all clearly distinct.

Vertical effects of the drag distribution include a different distribution of wind speed in the first several hundred metres, with a lesser gradient and higher wind speeds at the lowest level. Wind profiles are more heterogeneous, and the perturbations introduced grow vertically in the downwind direction, suggesting the formation of internal boundary layers. The extended wakes downstream of agglomerations of densely built-up neighbourhoods and high-rise buildings are evidence for a heterogeneous urban boundary-layer flow that is not just the aggregated flow over each neighbourhood, but depends crucially on interactions between the neighbourhoods. These findings show that height-distributed drag and the inclusion of subgrid heterogeneity bear potential for improved urban wind modelling in regions close to the surface.

The model in its current form has several limitations. The morphology profiles are relatively simple: for example, they do not depend on wind direction, and sheltering between buildings is not represented. The drag distribution function is derived from a limited number of building-resolving simulations; using data from a wider range of morphologies could improve the parametrization further. Further research is needed to determine a generally suitable drag coefficient. The current implementation requires adaptation of the scalar exchange, which highlights that the distribution of momentum exchange can only be the first step towards a comprehensive multilayer model that is not based on a simple street-canyon geometry.

AUTHOR CONTRIBUTIONS

Birgit S. Sützl: Conceptualization; Formal analysis; Investigation; Software; Visualization; Writing – original draft. **Gabriel G. Rooney:** Conceptualization; Investigation; Software; Writing – review & editing. **Anke Finnenkoetter:** Conceptualization; Software; Writing – review & editing. **Sylvia I. Bohnenstengel:** Conceptualization; Writing – review & editing. **Sue Grimmond:** Conceptualization; Writing – review & editing. **Maarten van Reeuwijk:** Conceptualization; Supervision; Writing – review & editing.

ORCID

*Birgit S. Sützl** https://orcid.org/0000-0001-9638-5643
Gabriel G. Rooney https://orcid.org/0000-0002-3787-1198
Anke Finnenkoetter https://orcid.org/0000-0003-0210-8288
Sylvia I. Bohnenstengel https://orcid.org/0000-0001-6170-5774
Sue Grimmond https://orcid.org/0000-0002-3166-9415
Maarten van Reeuwijk https://orcid.org/0000-0003-4840-5050

REFERENCES

Barlow, J.F., Best, M.J., Bohnenstengel, S.I., Clark, P.A., Grimmond, S., Lean, H.W., Christen, A., Emeis, S., Haeffelin, M., Harman, I.N., Lemonsu, A., Martilli, A., Pardyjak, E., Rotach, M.W., Ballard, S., Boutle, I.A., Brown, A.R., Cai, X.-M., Carpentieri, M., Coceal, O., Crawford, B., Di Sabatino, S., Dou, J., Drew, D.R., Edwards, J.M., Fallmann, J., Fortuniak, K., Gornall, J., Gronemeier, T., Halios, C.H., Hertwig, D., Hirano, K., Holtslag, A.A.M., Luo, Z., Mills, G., Nakayoshi, M., Pain, K., Schlünzen, K.H., Smith, S., Soulhac, L., Steeneveld, G.-J., Sun, T., Theeuwes, N.E., Thomson, D., Voogt, J.A., Ward, H.C., Xie, Z.-T. and Zhong, J. (2017) Developing a research strategy to better understand, observe and simulate urban atmospheric processes at kilometer to subkilometer scales. *Bulletin of the American Meteorological Society* 98(10), ES261–ES264. https://doi.org/10.1175/bams-d-17-0106.1.

Belcher, S.E. (2005) Mixing and transport in urban areas. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 363, 2947–2968. https://doi.org/10.1098/rsta.2005.1673.

Best, M.J., Beljaars, A., Polcher, J. and Viterbo, P. (2004) A proposed structure for coupling tiled surfaces with the planetary boundary layer. *Journal of Hydrometeorology*, 5(6), 1271–1278. https://doi.org/10.1175/JHM-382.1.

Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A., Gedney, N., Mercado, L.M., Sitch, S., Blyth, E., Boucher, O., Cox, P.M., Grimmond, S. and Harding, R.J. (2011) The Joint UK Land Environment Simulator (JULES), model description – part 1: energy and water fluxes. *Geoscientific Model Development*, 4(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011.

Bohnenstengel, S.I., Evans, S., Clark, P.A. and Belcher, S.E. (2011) Simulations of the London urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 137(659), 1625–1640. https://doi.org/10.1002/qj.855.

Bohnenstengel, S.I. and Hendry, M.A. (2016) *Report on Implementation and Evaluation of MORUSES in the UKV (ps37)*. Exeter, UK: Met Office. Technical report.

Boutle, I.A., Finnenkoetter, A., Lock, A.P. and Wells, H. (2016) The London model: forecasting fog at 333 m resolution. *Quarterly Journal of the Royal Meteorological Society*, 142(694), 360–371. https://doi.org/10.1002/qj.2656.

Bunce, R.G.H., Barr, C.J., Clarke, R.T., Howard, D.C. and Lane, A.M.J. (1990) *ITE land classification of Great Britain 1990*. Lancester, UK: NERC Environmental Information Data Centre, https://doi.org/10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f

Bush, M.R., Allen, T., Bain, C., Boutle, I.A., Edwards, J.M., Finnenkoetter, A., Franklin, C., Hanley, K., Lean, H.W., Lock, A.P., Manners, J., Mittermaier, M., Morcrette, C., North, R., Petch, J., Short, C., Vosper, S.B., Walters, D., Webster, S., Weeks, M., Wilkinson, J., Wood, N. and Zerroukat, M. (2020) The first Met Office Unified Model–JULES regional atmosphere and land configuration, RAL1. *Geoscientific Model Development*, 13(4), 1999–2029. https://doi.org/10.5194/gmd-13-1999-2020.

Clark, D.B., Mercado, L.M., Sitch, S., Jones, C.D., Gedney, N., Best, M.J., Pryor, M., Rooney, G.G., Essery, R.L.H., Blyth, E., Boucher, O., Harding, R.J., Huntingford, C. and Cox, P.M. (2011) The Joint UK Land Environment Simulator (JULES), model description - part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3), 701–722. https://doi.org/10.5194/gmd-4-701-2011.

Coceal, O. and Belcher, S.E. (2004) A canopy model of mean winds through urban areas. *Quarterly Journal of the Royal Meteorological Society*, 130(599), 1349–1372. https://doi.org/10.1256/qj.03.40.

Davies, T., Cullen, M.J.P., Malcolm, A.J., Mawson, M.H., Staniforth, A., White, A.A. and Wood, N. (2005) A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131(608), 1759–1782. https://doi.org/10.1256/qj.04.101.

Evans, S., Hudson-Smith, A. and Batty, M. (2006) 3-D GIS: virtual London and beyond. *Cybergeo: European Journal of Geography*. https://doi.org/10.4000/cybergeo.2871.

Garuma, G.F. (2018) Review of urban surface parametrizations for numerical climate models. *Urban Climate*, 24, 830–851. https://doi.org/10.1016/j.uclim.2017.10.006.

Grimmond, S., Blackett, M., Best, M.J., Barlow, J.F., Baik, J.-J., Belcher, S.E., Bohnenstengel, S.I., Calmet, I., Chen, F., Dandou, A., Fortuniak, K., Gouvea, M.L., Hamdi, R., Hendry, M.A., Kawai, T., Kawamoto, Y., Kondo, H., Krayenhoff, E.S., Lee, S.-H., Loridan, T., Martilli, A., Masson, V., Miao, S., Oleson, K., Pigeon, G., Porson, A., Ryu, Y.-H., Salamanca, F., Shashua-Bar, L., Steeneveld, G.-J., Tombrou, M., Voogt, J.A., Young, D. and Zhang, N. (2010) The international urban energy balance models comparison project: first results from phase 1. *Journal of Applied Meteorology and Climatology*, 49(6), 1268–1292. https://doi.org/10.1175/2010jamc2354.1.

Grimmond, S. and Oke, T.R. (1999) Aerodynamic properties of urban areas derived from analysis of surface form. *Journal of*

Applied Meteorology, 38(9), 1262–1292. https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2.

Harman, I.N., Barlow, J.F. and Belcher, S.E. (2004) Scalar fluxes from urban street canyons part ii: model. *Boundary-Layer Meteorology*, 113(3), 387–410. https://doi.org/10.1007/s10546-004-6205-7.

Hertwig, D., Grimmond, S., Hendry, M.A., Saunders, B., Wang, Z., Jeoffrion, M., Vidale, P.L., McGuire, P.C., Bohnenstengel, S.I., Ward, H.C. and Kotthaus, S. (2020) Urban signals in high-resolution weather and climate simulations: role of urban land-surface characterisation. *Theoretical and Applied Climatology*. 142(1), 701–728. https://doi.org/10.1007/s00704-020-03294-1.

Hertwig, D., Grimmond, S., Kotthaus, S., Vanderwel, C., Gough, H., Haeffelin, M. and Robins, A. (2021) Variability of physical meteorology in urban areas at different scales: implications for air quality. *Faraday Discussion*, 226, 149–172. https://doi.org/10.1039/D0FD00098A.

Kotthaus, S. and Grimmond, S. (2018) Atmospheric boundary-layer characteristics from ceilometer measurements. Part 1: a new method to track mixed layer height and classify clouds. *Quarterly Journal of the Royal Meteorological Society*, 144(714), 1525–1538. https://doi.org/10.1002/qj.3299.

Lean, H.W., Barlow, J.F. and Halios, C.H. (2019) The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length versions of the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*. 145(721), 1674–1689. https://doi.org/10.1002/qj.3519.

Lock, A.P., Edwards, J.M. and Boutle, I.A. (2020) *The Parametrization of Boundary Layer Processes*. Exeter, UK: Met Office. Unified Model Documentation Paper 024.

MacDonald, R.W., Griffiths, R.F. and Hall, D.J. (1998) An improved method for the estimation of surface roughness of obstacle arrays. *Atmospheric Environment*, 32(11), 1857–1864. https://doi.org/10.1016/S1352-2310(97)00403-2.

Martilli, A., Clappier, A. and Rotach, M.W. (2002) An urban surface exchange parametrization for mesoscale models. *Boundary-Layer Meteorology*, 104(2), 261–304. https://doi.org/10.1023/A:1016099921195.

Moszynska, M. (2006) *Selected Topics in Convex Geometry* (1st edn). Boston, MA: Birkhäuser.

Ordnance Survey (GB) (2019) *OS MasterMap building height attribute*. Edinburgh, UK: EDINA Digimap Ordnance Survey Service, https://digimap.edina.ac.uk.

Porson, A., Clark, P.A., Harman, I.N., Best, M.J. and Belcher, S.E. (2010a) Implementation of a new urban energy budget scheme in the MetUM. Part I: description and idealized simulations. *Quarterly Journal of the Royal Meteorological Society*, 136(651), 1514–1529. https://doi.org/10.1002/qj.668.

Porson, A., Clark, P.A., Harman, I.N., Best, M.J. and Belcher, S.E. (2010b) Implementation of a new urban energy budget scheme into MetUM. Part II: validation against observations and model intercomparison. *Quarterly Journal of the Royal Meteorological Society*, 136(651), 1530–1542. https://doi.org/10.1002/qj.572.

Rooney, G.G. and Bornemann, F.J. (2013) The performance of FLake in the Met office unified model. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1), 21363. https://doi.org/10.3402/tellusa.v65i0.21363.

Sützl, B.S., Rooney, G.G. and van Reeuwijk, M. (2021) Drag distribution in idealized heterogeneous urban environments. *Boundary-Layer Meteorology*, 178, 225–248. https://doi.org/10.1007/s10546-020-00567-0.

Tang, Y., Lean, H.W. and Bornemann, F.J. (2013) The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorological Applications*, 20(4), 417–426. https://doi.org/10.1002/met.1300.

Wood, N., Brown, A.R. and Hewer, F.E. (2001) Parametrizing the effects of orography on the boundary layer: an alternative to effective roughness lengths. *Quarterly Journal of the Royal Meteorological Society*, 127(573), 759–777. https://doi.org/10.1002/qj.49712757303.

Xie, Z.-T., Coceal, O. and Castro, I.P. (2008) Large-eddy simulation of flows over random urban-like obstacles. *Boundary-Layer Meteorology*, 129(1), 1–23. https://doi.org/10.1007/s10546-008-9290-1.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

APPENDIX A. URBAN MORPHOLOGY CALCULATIONS

Consider the two-dimensional, horizontal building cross-section B at fixed height (Figure A1). Let $\mathbf{e_\theta} = (\cos\theta, \sin\theta)$ be the unit wind vector for a wind angle θ and $\mathbf{e_{\perp\theta}} = (\sin\theta, -\cos\theta)$ a vector perpendicular to it. The building width $b(B, \theta)$ (in direction $\mathbf{e_{\perp\theta}}$) is defined as the distance between two lines parallel to $\mathbf{e_\theta}$, such that B is entirely between these two lines (Moszynska, 2006). This definition is equivalent to projecting the building cross-section onto a vector perpendicular to $\mathbf{e_\theta}$ and calculating the length of the projected cross-section (Figure A1a). The building cross-section is enclosed by a unique minimal convex polygon (i.e., the boundary of the convex hull of B; red polygon in Figure A1a) with vertices V (red dots in Figure A1a). For calculation of the building width, it is sufficient to project the building vertices in V onto $\mathbf{e_{\perp\theta}}$ and to calculate the distance between the two outermost projected vertices. We may write this as

$$b(B, \theta) = \max\left(\|P\mathbf{x} - P\mathbf{y}\| : \mathbf{x}, \mathbf{y} \in V\right), \quad (A1)$$

with the projection matrix

$$P = \mathbf{e}_{\perp\theta}\mathbf{e}_{\perp\theta}{}^T = \begin{pmatrix} (\sin\theta)^2 & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & (\cos\theta)^2 \end{pmatrix}. \quad (A2)$$

Figure A1a shows building widths of a building cross-section for several wind angles and illustrates the projection of vertices from V. The shape of the building cross-section B may change with height. We will informally denote this by $B(z)$, such that we can express a height-dependent function $b(B(z), \theta)$ that characterises the vertical structure of a building. The frontal area for wind angle θ is derived by height integration of the building width, $\int_0^{z_{max}} b(B(z), \theta) \, dz$, which is consistent with the typical definition of a frontal area (e.g., Grimmond and Oke, 1999).

In realistic scenarios the wind angle varies frequently, therefore it is useful to consider the average width with respect to all wind directions. The building's mean width $\bar{b}(B)$ is given by (Moszynska, 2006)

$$\bar{b}(B) \equiv \frac{1}{2\pi} \int_0^{2\pi} b(B, \theta) \, d\theta = \frac{l(V)}{\pi}, \quad (A3)$$

where $l(V)$ is the perimeter of the convex polygon described by V. This relation states that the mean width is equivalent to the diameter of a circle with the same perimeter as the minimal convex polygon enclosing B (cf. $l(V)$ in Figure A1a and b). The calculation of the building mean width is computationally inexpensive and straightforward with Equation A3 and it is therefore useful as a basis for calculating vertical profiles and morphology parameters over urban areas; in this case for the grid boxes of the NWP model. For a grid box with buildings described by the vertical building cross-section profiles $B_n(z)$, we define the buildings' total (mean) width at height z as

$$L(z) = \sum_n w_n \bar{b}(B_n(z)), \quad (A4)$$

where w_n is a weighting equal to the fraction of building area within the grid box to account for buildings that are only partially contained within the grid box. The grid box total frontal area A_F is defined as a wind-directional averaged total frontal area, which is given by integrating the total width with respect to height (Equation 2). The normalised frontal area $\zeta(z)$ is given by Equation 1, and the frontal-area index is

$$\lambda_f = \frac{A_F}{A_T} = \frac{1}{A_T}\int_0^{z_{max}} L(z) \, dz. \quad (A5)$$

Note that, by summing up the building-width contributions of each building in the grid box in Equation A4, we do not account for reduced wind-facing surfaces from buildings next to one another or sheltering effects from buildings in close proximity. Frontal areas A_F calculated by this method will therefore tend to be larger than areas from a "skyline" method that calculates A_F from the projection of all building fronts onto one plane, as the density of buildings in the grid box is represented implicitly by our method. We define the mean building height of the grid box as

$$z_H = \frac{A_F}{L(0)}. \quad (A6)$$

This quantity resembles a weighted building-height average, where each building height is weighted by the building's (ground-level) mean width $\bar{b}(B(0))$. The maximum height of any building in the grid box is denoted as z_{max}. The plan-area index is calculated in the usual way by $\lambda_p = A_P/A_T$, where $A_P = \sum_n w_n A_{p,n}$ and $A_{p,n}$ is the building's plan area.

FIGURE A1 The width $b(B, \theta)$ of a building cross-section B. (a) Building widths $b(B, \theta)$ for wind angles $\theta = 0°, 120°$, and $255°$; vertices V (red dots) describing the minimal convex polygon that encloses B (red polygon) and has perimeter $l(V)$. (b) Building's mean width $\bar{b}(B)$ (red dotted line) and circle with perimeter $l(V)$ (red circle) [Colour figure can be viewed at wileyonlinelibrary.com]