Study of Realistic Urban Boundary Layer Turbulence with High-Resolution Large-Eddy Simulation

Mikko Auvinen [1,2,*,†] **, Simone Boi** [3,4] **, Antti Hellsten** [1] **, Topi Tanhuanpää** [5] **and Leena Järvi** [2,6]

1   Atmospheric Dispersion Modelling, Atmospheric Composition Research, Finnish Meteorological Institute, 00560 Helsinki, Finland; antti.hellsten@fmi.fi
2   Institute of Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, 00560 Helsinki, Finland; leena.jarvi@helsinki.fi
3   Department of Mathematics and Statistics, University of Helsinki, 00560 Helsinki, Finland; simone.boi@helsinki.fi
4   OP Financial Group, 00510 Helsinki, Finland
5   Department of Forest Sciences, University of Helsinki, 00790 Helsinki, Finland; topi.tanhuanpaa@helsinki.fi
6   Helsinki Institute of Sustainability Science, University of Helsinki, 00014 Helsinki, Finland
*   Correspondence: mikko.auvinen@fmi.fi; Tel.: +358-50-320 5515
†   Current address: Erik Palménin Aukio 1, 00560 Helsinki, Finland.

Abstract: This study examines the statistical predictability of local wind conditions in a real urban environment under realistic atmospheric boundary layer conditions by means of Large-Eddy Simulation (LES). The computational domain features a highly detailed description of a densely built coastal downtown area, which includes vegetation. A multi-scale nested LES modelling approach is utilized to achieve a setup where a fully developed boundary layer flow, which is also allowed to form and evolve very large-scale turbulent motions, becomes incident with the urban surface. Under these nonideal conditions, the local scale predictability and result sensitivity to central modelling choices are scrutinized via comparative techniques. Joint time–frequency analysis with wavelets is exploited to aid targeted filtering of the problematic large-scale motions, while concepts of information entropy and divergence are exploited to perform a deep probing comparison of local urban canopy turbulence signals. The study demonstrates the utility of wavelet analysis and information theory in urban turbulence research while emphasizing the importance of grid resolution when local scale predictability, particularly close to the pedestrian level, is sought. In densely built urban environments, the level of detail of vegetation drag modelling description is deemed most significant in the immediate vicinity of the trees.

Keywords: large-eddy simulation; turbulence; urban boundary layer; urban canopy; wavelet analysis; information entropy

1. Introduction

In a rapidly urbanizing world, it is important to improve our understanding of the turbulent dispersion and exchange processes influencing the microclimate and air quality of our cities. This poses a challenge because urban boundary layer (UBL) flows are complex flow systems that arise from the interaction between atmospheric boundary layer (ABL) turbulence and varying arrangements of urban obstacles [1]. While the general urban layout, characterized by the mean building (or building block) height, density, and orientation, governs the wind dynamics in a statistical sense, the actual aerodynamic properties of individual obstacles dictate the local flow conditions which bear the greatest significance to people residing or working in the area. Thus, it is important to aquire the ability to

predict these local scale flow system details for they strongly influence the local exchange rates of momentum, heat, and humidity, as well as the dispersion rates of pollutants [2–5].

Real cities have complicated morphologies due to the inherent variability in the shape and size of urban roughness elements (e.g., buildings, building blocks, individual street trees, and patches of park trees), and the frail appearance of topographic regularity is typically broken by the presence of openings (e.g., lawns, market places, and parking lots) that can significantly alter the local wind conditions. Parameterized approaches to UBL flows (e.g., References [6,7]) do not provide the means to identify and examine such complex, real world urban configurations in a targeted fashion. Only high-fidelity numerical or experimental techniques capable of capturing the relevant flow processes despite the complexity can offer a sufficiently accurate mechanism to study these flow systems. Unfortunately, the study of real urban climate is hindered by the difficulty and cost associated with performing thorough measurement campaigns in the field or in the wind tunnel. Luckily, the latest numerical modelling techniques, computational resources, and remote scanning datasets do offer tantalizing opportunities to examine realistic urban flow scenarios with models that describe the urban surface with an ever increasing level of detail. Particularly Large-Eddy Simulation (LES), a modelling approach in the field of Computational Fluid Dynamics (CFD) capable of temporally and spatially resolving the relevant turbulent processes governing the transport and exchange of momentum, energy, and materials between the urban surface and the ambient air, provides a well-established modelling framework with which the relevant turbulent processes in complex UBL flows can be simulated [8,9].

The research on UBL turbulence has progressed through an incremental course which initially, justifiably, exploited parametrized urban surfaces described by an arrangement of regular cuboids. The ability to resolve idealized urban canopy flows with the LES technique has been demonstrated via a series of studies involving comparisons with wind tunnel and direct numerical simulation (DNS) results [10–19], establishing the complex nature of fundamental urban canopy turbulence, its weak dependency on Reynolds number (at realistic, high Re conditions), the flow solutions' relative indifference to the choice of subgrid scale (SGS) turbulence model, and the general LES resolution requirements for capturing the relevant flow physics between and above regular urban roughness elements. The view on UBL flows is further augmented by a series of studies featuring real urban neighborhoods [2–5], providing additional quantification on the flow behavior and turbulent transport mechanisms within the roughness sublayer (RSL), the vertical range of the UBL within which the effect of individual urban obstacles are observed. These works employed rather restricted ABL turbulence, featured limited urban domains, and excluded urban vegetation. Urban vegetation plays an important role in UBL flows, explicated by Giometto et al. [20] in an LES study featuring a suburban setting. Furthermore, the validation study featuring real urban complexity (but without trees) documented in References [21,22] provides a relevant account of validation approaches, some of which are adopted and expanded upon in this work. Unfortunately, the resolution of their LES model was clearly insufficient to demonstrate the full capacity of LES in urban flow problems.

While the arbitrary complexity of real urban environments demands a high degree of accuracy from the LES simulations, it also brings to focus a central question concerning the necessary level of detail of the model inputs. Thus, considering the difficult nature of real urban flows, we must ask what is the meaningful level of detail with which the urban roughness elements, buildings, and particularly vegetation must be described within the model? This study examines issues relating to local scale predictability and the observability of changes in realistic UBL flows where the situation is further complicated by nonideal conditions. For instance, many major cities are situated near the coastline, which causes the boundary layer flow to abruptly go though a transition from smooth surface to rough surface flow (or vice-versa) as the wind crosses the shoreline. These transitions disturb the flow and create significantly large areas where developing boundary layer flow conditions prevail. Also, in real UBL flows, where the urban roughness interacts with real ABL turbulence, the divide between the smallest and the largest turbulent scales becomes particularly pronounced. Even under near neutral ABL conditions (most suitable for validation efforts) the flow contains both large-scale

and very large-scale turbulence motions (LSM and VLSM) of which the time scales are so long that it is unreasonable to expect local in situ measured time series to capture sufficiently converged statistics on them. Therefore, it is necessary to examine the complexities real urban flows impose and to seek methodologies to extract relevant information even concerning the local scale processes.

This numerical study employs state-of-the-art LES modelling techniques to examine a realistic UBL flow scenario designed to bear close correspondence with observable urban turbulence under neutral stratification. The PArallel Large-eddy Simulation Model (PALM) LES solver [23], with a newly added nested multi-scale system [24], is used to perform numerical simulations. The LES model domain features a highly detailed description of downtown Helsinki, a coastal capital city of Finland accommodating a variety of urban microclimate studies [25]. The considered city domain contains a significant amount of urban vegetation, which is included in the LES model as porous media. The study transparently documents the employed model configuration where fully developed ABL turbulence becomes incident with the urban surface, generating a developing UBL flow characteristic for coastal cities.

The numerical setup is designed to allow the ABL flow to develop and evolve very large-scale streamwise structures that are observed in high Reynolds number boundary layer flows (see e.g., Hutchins et al. [26], Fang and Porté-Agel [27], and references therein). These very large-scale motions (VLSM) of ABL, of which the length scales extend well beyond the boundary layer height δ, introduce difficulties in analyzing real urban turbulence statistics as the diurnally changing ambient conditions impose limits on the considered time series length. Thus, this work deliberately adopts a confined approach, where the considered time series do not yield sufficiently converged statistics on the VLSMs, and examines the sensitivity of the urban LES solution to central numerical modelling choices via comparative techniques. Particular emphasis is placed on the aerodynamic modelling of urban vegetation due to the embedded uncertainties in determining their spatial extent and aerodynamic attributes. The sensitivity of LES-resolved UBL turbulence to the alterations in vegetation model is examined in juxtaposition to its sensitivity to the modelled height of the ABL. This aspect is considered because the considered UBL flow is developing, having gone though a sharp transition at the leading edge of the city as the ABL flow becomes incident with the shoreline. Here, the inference that, under fully developed UBL conditions, the low altitude RSL turbulence is insensitive to outer layer dynamics is considered insecure, warranting confirmation in the chosen context. In addition, the considered sensitivities in local turbulence signals are also compared against the effects arising from an elementary reduction in grid resolution.

The comparative post-processing analysis starts by evaluating deviations in two-dimensional mean flow distributions over horizontal planes and proceeds by comparing the contents of virtual tower measurements extracted at different locations within the downtown area. The local tower measurements are subjected to joint time–frequency analysis with wavelets to inspect the makeup of the simulated turbulent fluctuations with some level of temporal coherence across different frequencies (or scales). In an effort to expose deviations in the local turbulent signals that arise due to the model modifications, concepts of information entropy and divergence are exploited in conjuction with wavelet analysis to carry out a deep probing comparison of LES-resolved urban canopy turbulence. These mathematical techniques are robustly introduced to establish a solid foundation for their utility in UBL turbulence research.

Due to the high degree of complexity of the UBL turbulence, this study aims to address what is the required level of detail for the urban surface description within high-resolution LES models. This information is relevant considering that the highest quality (vegetation resolving) light detection and ranging (LiDAR) scan datasets are not publically available for many potential study sites. This work is also considered a relevant precursor for future validation efforts, providing guidelines for the urban LES modeling and further identifying the bounds within which agreement against urban in situ measurements can be sought considering the modelling uncertainties and observation limitations.

2. Materials and Methods

2.1. Note on Nomenclature

Throughout this document, Cartesian vectors are denoted by boldface font and scalars are denoted by normal font. Thus, for instance, the coordinate vector and its components are denoted as $\mathbf{x} = (x, y, z)$. In the context of LES, the velocity vector obeys the following decomposition $\mathbf{U}(t, \mathbf{x}) = \mathbf{u}(t, \mathbf{x}) + \mathbf{u}''(t, \mathbf{x})$, where $\mathbf{U}$ is the unfiltered velocity field, $\mathbf{u}$ is the space filtered (grid resolved) velocity, and $\mathbf{u}''$ is the subgrid-scale velocity. The Cartesian velocity vector components are $\mathbf{u} = (u, v, w)$, whereas the flow components in a horizontally rotated coordinate system (x_1, x_2, x_3), where the x_1-axis is aligned with the mean flow's streamwise direction, are marked as $\mathbf{u} = (u_1, u_2, u_3)$, where u_1 is the component in the mean streamwise direction, u_2 is the component in the mean crosswind direction, and u_3 is the vertical component. These velocity components are computed as follows:

$$\alpha = \arctan(\overline{v}/\overline{u})$$
$$u_1 = u\cos(\alpha) + v\sin(\alpha)$$
$$u_2 = -u\sin(\alpha) + v\cos(\alpha)$$
$$u_3 = w$$

where α is the local horizontal mean flow direction. The Reynolds decomposition of the resolved velocity field follows the convention $\mathbf{u}(t, \mathbf{x}) = \overline{\mathbf{u}}(\mathbf{x}) + \mathbf{u}'(t, \mathbf{x})$, using overbar $\overline{\bullet}$ to indicate a temporal average and prime $(\bullet)'$ to specify fluctuations about the mean. By default, angled brackets $\langle\bullet\rangle \equiv \langle\bullet\rangle_{x,y}$ denote a horizontal average, while a subscript is used in other situations. All the spatial averages in this text are intrinsic averages (see Whitaker [28], Schmid et al. [29]), where all data points representing air or porous media (vegetation) are included while points within solid obstacles are excluded.

2.2. Large-Eddy Simulation Solver PALM

In this study PALM [23] LES model is utilized to study developing urban boundary layer flows. The solver employs finite-difference discretization on staggered Cartesian grid to numerically integrate the filtered, Boussinesq-approximated Navier–Stokes equations for incompressible flows, presented here using vector notation while omitting moisture- and Coriolis acceleration-related terms as they are not relevant in this study:

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{1}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -\nabla \cdot (\mathbf{u} \otimes \mathbf{u}) - \frac{1}{\rho_0}\nabla\pi^* + g\frac{\theta - \langle\theta\rangle}{\langle\theta\rangle}\hat{\mathbf{k}} - \nabla \cdot \left(\mathbf{T_{sgs}} - \frac{1}{3}\operatorname{tr}(\mathbf{T_{sgs}})\,\mathbf{I}\right) \tag{2}$$

$$\frac{\partial \theta}{\partial t} = -\mathbf{u} \cdot \nabla\theta - \nabla \cdot \mathbf{q_{sgs}}, \tag{3}$$

In Equation (2), g is the gravitational accelaration, $\hat{\mathbf{k}}$ is the unit vector in the z direction, $\mathbf{I}$ is the identity matrix, $\mathbf{T_{sgs}} = (\overline{\mathbf{u} \otimes \mathbf{u}} - \overline{\mathbf{u}} \otimes \overline{\mathbf{u}})$ is the subgrid-scale kinematic stress tensor as $\otimes$ denotes the outer vector product, $\pi^* = p^* + 1/3\rho_0 \operatorname{tr}(\mathbf{T_{sgs}})$ is the modified perturbation pressure (p^* being the perturbation pressure), and ρ_0 is the constant density of air at ground level. In the energy equation, Equation (3), θ is the potential temperature and $\mathbf{q_{sgs}} = \left(\overline{\mathbf{u}''\,\theta''}\right)$ is the subgrid scale temperature flux vector.

PALM solves the governing equations on a regular hexagonal domain $\Omega^{(n)}$ with a boundary $\partial\Omega^{(n)} = \{\Gamma_j^{(n)} : j \in B\}$, where the boundary set consists of $B = \{\texttt{l}, \texttt{r}, \texttt{n}, \texttt{s}, \texttt{t}, \texttt{b}\}$, which refer to the left, right, north, south, top, and bottom boundaries of the nth domain. The superscript n is the identification label of the domain, which is relevant when multiple domains are coupled via nesting technique (see Section 2.2.1 below).

For the discretization of the advection and diffusion schemes, the fifth-order Wicker–Skamarock and second-order central difference schemes are employed, respectively [30]. The solution strategy of PALM is based on a 3rd-order accurate Runge–Kutta time-stepping scheme and a predictor-corrector approach, where the divergence generated by the predictor step in the velocity field is subsequently removed by a corrector step utilizing the solution to a Poisson equation for π^*. An efficient iterative multigrid solver is used for the pressure solution. To attain closure in the final system of filtered equations, PALM implements the 1.5-order subgrid-scale (SGS) turbulence model by Deardorff [31] with modifications due to Moeng and Wyngaard [32] and Saiki et al. [33].

The PALM model system includes also a newly implemented surface energy model [34], which incorporates the principal surface energy budget processes occurring within urban environments. Since the objective of this study is to examine the impact of certain modelling choices and their observability in an urban flow system under neutral ABL conditions (when only the anthropogenic heat component is relevant), the functionality of the USM is not exploited here.

2.2.1. Nested LES Domain Approach

The PALM solver has been recently implemented with a self-nested domain approach, which allows multiple solution domains with decreasing domain sizes but increasing resolutions to be solved simultaneously in a coupled manner. This approach facilitates an efficient simulation of flow problems that contain a wide range of turbulent scales by making it possible to employ higher resolution only where necessary. In this context, a notational convention $\Omega^{(n,p)}$ is adopted for labeling domains, where the superscript $(n, p(n))$ includes a unique identifier n for every so-called child domain and an optional parent entry $p = p(n)$ to indicate whether domain n has a parent domain p within which it is fully contained. In situations where parent information is not critical, shorter notation $\Omega^{(n)}$ may be used. By convention, the main parent domain (which defines the outer boundaries of the LES model) is denoted $\Omega^{(1)}$, whereas the subsequent $n \in \{2, \ldots, M\}$ child domains become $\Omega^{(2,1)}, \ldots, \Omega^{(M,p(M))}$, noting that $M > p(M)$ always applies. The boundary notation is also adjusted accordingly such that, for instance, $\Gamma_t^{(3)}$ refers to the top boundary of domain $\Omega^{(3,p(3))}$.

The solutions of each parent and child pair are coupled either by one-way or two-way coupling schemes (see Hellsten et al. [24] for details). In this study, the LES model is composed of three one-way coupled domains forming a nested system $\Omega^{(3,2)} \subset \Omega^{(2,1)} \subset \Omega^{(1)}$. Table 1 lays out the relevant information on each computational domain. In the one-way coupling scheme, the parent domains $\Omega^{(p)}$ provide time-accurate Dirichlet boundary conditions for every prognostic variable on $\partial\Omega^{(n,p(n))}$, excluding the bottom boundaries $\Gamma_b^{(n)}$ where identical solid wall boundary conditions are employed on all domains. The one-way coupling scheme is particularly well suited for the considered advection-driven UBL flow study where it is desired that the solution of the coarser domains remain unaltered by the finer nested domain solutions (which is contrary in the two-way coupling mode.) The boundary conditions are further described in Section 2.4.1.

Table 1. Dimensions and grid specifications for the nested Large-Eddy Simulation (LES) domains.

		$\Omega^{(1)}$	$\Omega^{(2,1)}$	$\Omega^{(3,2)}$
Domain dimensions	L_x	7168 m	3072 m	1024 m
	L_y	1536 m	1024 m	512 m
	L_z	512 m	256 m	128 m
Cell size	$\Delta_x, \Delta_y, \Delta_z$	4 m	2 m	1 m
Domain node counts	N_x	1792	1536	1024
	N_y	384	512	512
	N_z	128	128	128
Total domain node count	$N_x N_y N_z$	88×10^6	101×10^6	67×10^6
Nested model node count	$\sum_i (N_x N_y N_z)^{(i)}$		256×10^6	

2.3. Urban Surface Model in LES

This study utilizes a numerical model that features a detailed description of a real urban area (Figure 1). The surface model features a $2.2 \times 1.5\,\text{km}^2$ area of downtown Helsinki, the coastal capital of Finland. This region contains a dense arrangement of perimeter blocks, an irregular distribution of vegetated openings and streets (boulevards), and a network of narrow street canyons plus a central multilane artery that divides two downtown districts. The chosen area encompasses universal elements of low built urban environments that are common particularly within European cities. The urban surface model is constructed from an open access LiDAR dataset (accessible via http://hel.fi/3D) provided by the City of Helsinki. For the purpose of this and future urban LES studies, the point cloud dataset has been processed to yield a series of detailed raster maps, including topography (i.e., digital elevation model (DEM)), orography (land surface elevation), vegetation height, and land cover distributions. These raster materials are also made openly accessible for the community [35].
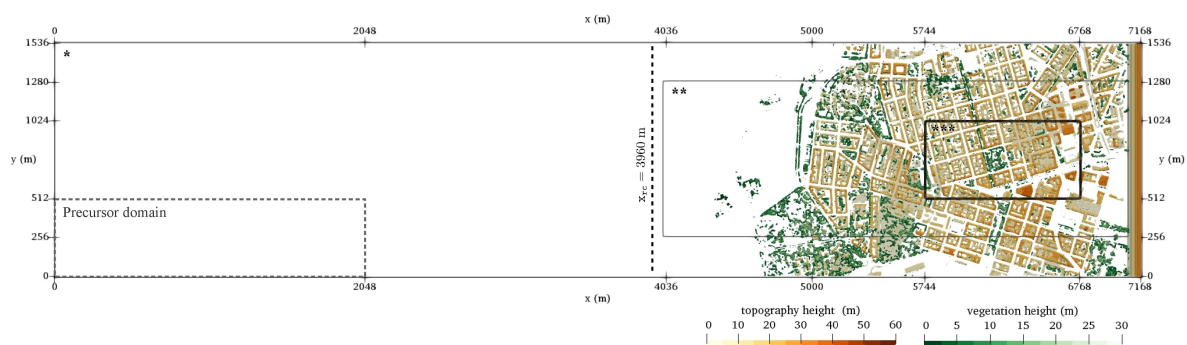

Figure 1. A two-dimensional visualization of the urban LES domain used in the study: The parent domain $\Omega^{(1)}$, domain $\Omega^{(2)}$, and $\Omega^{(3)}$ are marked with *, **, and ***, respectively, in the upper left corner of the domains' highlighted borders. The urban turbulence data is primarily extracted from within the $\Omega^{(3)}$ domain, which is discretized with uniform 1 m resolution ($\Delta = H/20$). The horizontal extent of the precursor simulation domain, used to initialize the urban LES, is also shown in the bottom left corner of the parent domain.

Information concerning the surface morphometry of the $\Omega^{(2)}$ and $\Omega^{(3)}$ model domains is presented in Table 2. The relevant mean canopy height is assigned $H = 20$ m, which very closely matches the mean building height of the $\Omega^{(3)}$ domain and the upwind sector belonging to $\Omega^{(2)}$. The plan and frontal area fractions of obstacles are defined as $\lambda_P = A_P/A_{tot}$ and $\lambda_F = A_F/A_{tot}$, where A_P is the planar area of obstacles at ground level, A_F is the total frontal area (visible to the flow) of obstacles, and $A_{tot} = L_x \times L_y\ (\text{m}^2)$ is the total planar area of the domain. Using the flow regime categorization for idealized urban canopies according to Hussain and Lee [36], the obtained building area fractions classify the urban area as dense where skimming flow dominates [37]. However, due to the irregularity in building morphologies and variability in street canyon orientations and vegetation distributions, such flow regime characterization offers little information about the considered urban canopy flow.

The building and tree height distributions in Figure 2 illustrate the symmetric variability in building heights (except for the single outlier at $2.8H$) and the positive skew distribution of the vegetation heights. The vegetation height distribution closely resembles that of a suburban area reported by Giometto et al. [20], except that here the small plants and shrubs with $h_v < 2$ m are not explicitly described by the dataset. The effect of the low vegetation and other fine structural details (balconies, small chimneys, roof ventilator, etc.) are taken into account via surface roughness parametrization in the LES simulation (see Section 2.4 below).

Table 2. Summary of morphometric statistics for the urban area: Separate values are listed for the area within $\Omega^{(2)}$ but outside of $\Omega^{(3)}$ and within $\Omega^{(3)}$ alone.

		$\Omega^{(2)} \setminus \Omega^{(3)}$	$\Omega^{(3)}$
Mean	building heights	$\langle h_b \rangle = 19.2$ m	$\langle h_b \rangle = 20.8$ m
	tree heights	$\langle h_v \rangle = 11.8$ m	$\langle h_v \rangle = 10.2$ m
Characteristic canopy height		$H = 20$ m	
Std. of	building heights	$\sigma_{h,b} = 7.8$ m	$\sigma_{h,b} = 12.8$ m
	tree heights	$\sigma_{h,v} = 6.2$ m	$\sigma_{h,v} = 4.6$ m
Plan area fractions of	buildings	$\lambda_{P,b} = 0.32$	$\lambda_{P,b} = 0.51$
	trees	$\lambda_{P,v} = 0.12$	$\lambda_{P,v} = 0.08$
Frontal area fractions of	buildings	$\lambda_{F,b} = 0.19$	$\lambda_{F,b} = 0.29$
	trees	$\lambda_{F,v} = 0.09$	$\lambda_{F,v} = 0.07$
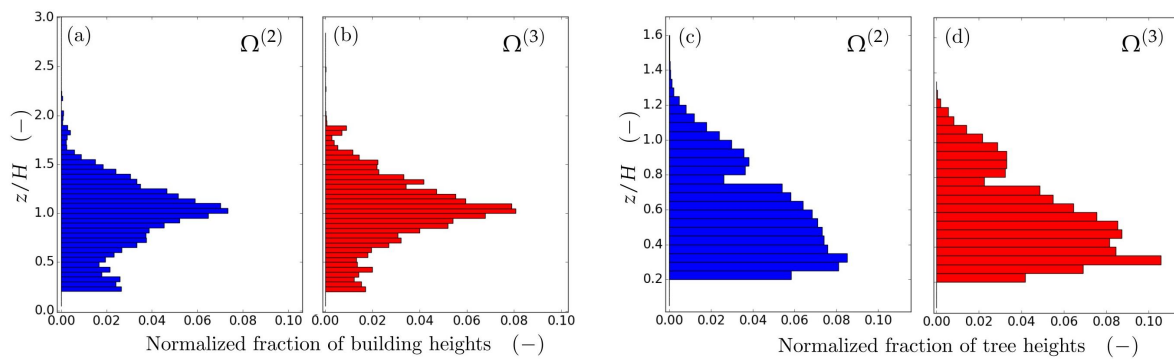
Figure 2. Normalized distributions of (**a**,**b**) building and (**c**,**d**) tree heights within $\Omega^{(2,1)}$ and $\Omega^{(3,2)}$ domains respectively.

2.4. LES Modelling

To facilitate this comparative investigation of local scale urban canopy turbulence, a baseline LES simulation setup is established to act as a reference onto which each individual alteration is subsequently introduced. The wind direction is held unaltered for the all the simulations because the structural complexity and irregular orientation of the considered city plan introduces a significant amount of directional and structural variation into the roughness sublayer turbulence.

2.4.1. Boundary Conditions and Initialization

The inlet boundary condition on $\Gamma_{\text{in}} \equiv \Gamma_l^{(1)}$ is designed to represent fully developed neutral ABL turbulence, which contains the appropriate range of length scales and remains unaffected by the urban roughness. This is achieved by employing the *turbulence recycling* technique (see References [23,38] for details), where the inlet boundary value is constructed from two contributions

$$\psi_{in}(x,y,z,t)|_{\mathbf{x} \in \Gamma_{\text{in}}} = \langle \overline{\psi} \rangle (z) + \psi'(x,y,z,t)|_{\mathbf{x} \in \Gamma_{\text{rc}}}.$$

Here, $\psi = \{u, v, w, \theta, ...\}$ is any prognostic variable such that the predetermined mean vertical inlet profile $\langle \overline{\psi} \rangle (z)$ is held fixed while the fluctuating components ψ' are recycled to the inlet from a specified recycling plane Γ_{rc} at $x = x_{\text{rc}}$ downstream (drawn in Figure 1). Only the fluctuations occuring for $z < 0.8 L_z$ are recycled, and therefore, $\delta = 0.8 L_z$ is utilized as the boundary layer height for the reference case from here on. The recycling height is modified in one of the documented simulations.

The inlet profiles $\langle \overline{\psi} \rangle (z)$ for the prognostic variables are generated by precomputing a fully developed ABL solution, called a precursor solution, on a $L_x = 2048$ m, $L_y = 512$ m, $L_z = 512$ m domain outlined in the bottom left corner of Figure 1, with periodic lateral boundary conditions and no-slip and free-slip conditions at the bottom and top, respectively. This flow is driven by a

prescribed negative pressure gradient, which is imposed for $z > 0.7\,L_z$. The magnitude and direction of the pressure gradient is adjusted such that $\langle \overline{u_{1\text{pre}}} \rangle \equiv \langle \overline{u_1} \rangle\,(z = 0.95\,L_z) \approx 10\;\text{m s}^{-1}$ and the flow angle with respect to the x-axis is $\alpha = \arctan(\langle \overline{v} \rangle / \langle \overline{u} \rangle) \approx 3°$. The recycling method yields a flow solution, in which the double-averaged profiles in Figure 3 closely resemble ABL flow in geostrofic balance, featuring a constant momentum flux layer at $z < L_z/2$. However, this setup exhibits faster convergence. The precursor simulation was initialized with a guessed mean profile and computed for $\Delta t = 290\,T_{\text{pre}}$, as $T_{\text{pre}} = L_z / \langle \overline{u_{1\text{pre}}} \rangle$, to obtain well-converged double-averaged profiles from the last $150\,T_{\text{pre}}$.
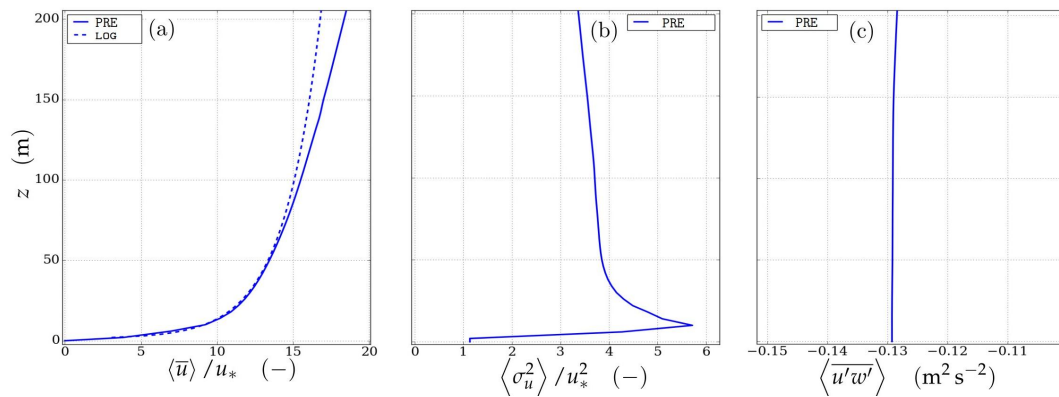
Figure 3. Double-averaged vertical profiles characterizing the precursor flow solution: (**a**) nondimensional u-component profile (solid blue line) juxtaposed with a reference logarithmic profile (dashed blue line); (**b**) nondimensional variance profile of u; and (**c**) vertical momentum flux profile.

The flow field of the urban LES simulation is, in turn, initialized by repetitively copying the precursor flow solution to the $\Omega^{(1)}$ domain. The prognostic variables within the $\Omega^{(2)}$ and $\Omega^{(3)}$ domains are initialized by interpolating the parent domain values onto their respective grids.

The relatively long frontal section representing ABL flow over sea surface is constructed to facilitate the development and evolution of large-scale and very-large-scale motions, termed LSM and VLSM, respectively, that are observed in pipe, channel, and boundary layer flows (see, for example, References [26,39–42], and references therein). The presence and relevance of these long streamwise structures, which affect the low altitude roughness sublayer flow via amplitude modulation, have been thoroughly established [43–46] in near neutral ABL flows, implying they ought to be considered a universal trait of UBL flows under similar conditions. Therefore, despite the statistical convergence difficulties they introduce, the inclusion of the large-scale boundary layer motions is deemed essential in the presented simulations as close compatibility is sought between modelled and (in the future) measured urban turbulence.

The small flow angle in the simulated ABL flow is important because the turbulence recycling setup is prone to numerically amplifying the elongated streamwise disturbances, causing the VLSM flow between the recycling planes to become correlated with the inlet [47]. The small $\langle \overline{v} \rangle$-component forces the LSM and VLSM structures to traverse the domain in the spanwise direction and to evolve without numerical amplification. This requires the distance between the inlet and the recycling plane to be sufficiently long. For this configuration, the distance $x_{\text{rc}} - x_{\text{in}} = 3960$ m $\approx 10\delta$ is found to be adequate when the mean flow angle results in $\delta/2$ displacement in the lateral direction as the flow moves from the inlet to the recycling plane.

The placement of the recycling plane is also sensitive to the disturbances that travel upstream from the leading edge of the urban surface. In this study, the disturbances generated by the real, unsymmetric shoreline of Helsinki gives rise to a weak perturbation at the recycling plane, which is subsequently transported all the way to the inlet. This perturbation manifests by adding a weak but observable systematic stripiness in the mean flow. This influence could be eliminated by substantially increasing the distance between the recycling plane and the leading edge of the city. However,

a computationally more efficient technique is utilized here instead. The effect of the unsymmetric disturbances is diminished by adding a single sparse array of narrow but tall spires (0.04δ wide "towers" reaching a height $\Delta z = L_z/2$, placed $\Delta y = L_z/2$ apart) right after the inlet plane. These structures introduce a minor but regular source of small-scale disturbances that do not have an observable impact on the generated ABL flow far downstream but make the turbulence recycling boundary conditions more immune to amplification of upwind travelling disturbances from the urban roughness.

In each computational domain, the no-slip wall boundary condition is imposed for all $\mathbf{x} \in \Gamma_{\mathbf{b}}$, including the building walls. The shear stress at the wall is computed in accordance to the logarithmic law of the wall with roughness, taking into account the effect of the roughness elements not resolved by the computational grid by assigning the subgrid-scale aerodynamic roughness length $z_{0,\mathrm{sgs}} = 0.03$ m. The thermal landscape of the city due to anthropogenic sources is taken into account in a simplified manner by assigning a Dirichlet condition $\theta|_{\mathbf{x} \in \Gamma_{\mathbf{u}}} = 1.005\,\theta_{\mathrm{pre}}$, where θ_{pre} is the constant potential temperature of air from the neutral precursor simulation and $\Gamma_{\mathbf{u}} \subset \Gamma_{\mathbf{b}}$ is the urban surface. The inclusion of anthropogenic heat is not relevant for this comparative study, but it is deemed an inherent ingredient of UBL flows under neutrally stratified atmospheric conditions.

2.4.2. Modeling of Trees as Porous Media

The difficulty of including trees within an urban LES model originates from the complex nature of their aerodynamic impact. The drag caused by an individual tree is much higher when it is isolated compared to being part of a forest canopy where the trees shelter each other. The earlier fundamental research on the effects of vegetation in ABL flows has considered continuous canopies where individual trees or vegetation elements are not identifiable objects with distinct spatial and aerodynamic properties or such that the canopy is constructed from a regular array of distict objects; see e.g., References [48–52].

Within densely built urban environments, the situation is complicated by the sporadic appearance of vegetation. Trees are typically planted individually along boulevards, esplanades, and other roads, whereas parks and other green areas may feature clusters of trees where the merging of foliage may form canopy-like patches. LiDAR measurements can provide spatially resolved information on the structural density of tree foliage, but their point cloud datasets may not provide reliable information concerning the inner structure of trees that have large crown sizes and dense leaf covers. The quality of this data depends on the amount of LiDAR beam reflections obtained from the inner part of the foliage. Even in the case of accurate structural probing of trees, their location with respect to the wind and each other strongly influence the aerodynamic properties they possess. For this reason, the modelling of vegetation within urban environments carries significant uncertainties of which the overall impact on the urban flow system needs to be established before detailed air quality and urban plan investigations featuring real urban sites can be undertaken.

In the PALM LES model, the aerodynamic drag due to vegetation is modelled as a momentum sink in the form of pressure drag (assuming the viscous drag contribution to be negligible in comparison) and is formulated as follows:

$$\mathbf{f}_v = -C_{Dv}\,LAD\,|\mathbf{u}|\mathbf{u} = -D_v\,|\mathbf{u}|\mathbf{u} \tag{4}$$

where C_{Dv} is a constant drag coefficient for vegetation and $LAD = LAD(\mathbf{x})$ (m^{-1}) is the leaf area density or, synonymously, plant area density of discretized tree elements (cells in the grid occupied by foliage). The LAD is the ratio of the plant's surface area (visible to the flow) to the volume it occupies. In this study, the uncertainty in the effective drag coefficient $D_v(\mathbf{x}) = C_{Dv}LAD(\mathbf{x})$ which combines the two model inputs is scrutinized by imposing variations in both C_{Dv} and LAD that are expected to span a relevant fraction of the uncertainty range associated with these inputs.

The formulation of Equation (4) simplifies a complex problem by splitting the user specified elements of D_v into a universal constant C_{Dv} and a plant's density measure, which is solely determined

by the spatial attributes of the tree. This is immediately problematic because trees are not only spatially complex objects but also elastic structures of which the structural and aerodynamic properties (e.g., branch rigidity, leaf, or needle shape reconfiguration under wind loads) can vastly vary between different species. Experimental drag coefficient measurements carried out by, for example, Mayhead [53], Rudnicki et al. [54], and Vollsinger et al. [55], reveal the wide scatter of result between different species. The situation is further complicated by the influence of wake interaction or "shading" [56,57] that occurs between trees within close proximity of each other. When this effect is included in the drag coefficient value, the implication is that the drag is caused by a population of trees and C_{Dv} is termed *bulk* drag coefficient. Due to the complexity of the urban vegetation system, the drag coefficient is interpreted to encompass both meanings.

Consider the choice for the drag coefficient C_{Dv} first. When ABL flows over continuous vegetated canopies are considered, $C_{Dv} \approx 0.1 \ldots 0.3$ is a typical range of values, and often, $C_{Dv} = 0.2$ is chosen [7,20,58]. This originates from Katul and Albertson [59], who found this to be the optimal value that minimized the discrepancy between two-dimensional Lagrangian stochastic turbulence model simulations and experimental measurements. However, this value is too small for urban situations where individual trees or closely packed populations of trees are sporadically distributed within the urban plan experiencing varying amounts of shading by buildings or other trees. The D_v values for unshaded individual trees in crosswind are measured to vary widely (as a function of wind speed and tree species), ranging between $D_v = 0.2 \ldots 1.2 \text{ m}^{-1}$ [53,55]. Clearly, specifying one value for the urban application is not optimal, but the suitable value for the considered urban application should lie somewhere between the two canopy regimes.

In this study, the drag coefficient value is approximated by first adopting the method by Raupach et al. [60], as utilized in Ghasemian et al. [61], where the drag coefficient for windbreaks is evaluated from

$$C_{Dv} = \frac{\Gamma_{b1} k}{k + \Gamma_{b1} k_1} \tag{5}$$

where $k = \Delta p_s / (1/2)\rho|\mathbf{u}|^2 d_v$ is the pressure loss coefficient, Δp_s is the drop in static pressure across vegetation of thickness d_v, Γ_{b1} is the bulk drag coefficient for a solid fence (assigned $\Gamma_{b1} = 1.07$ according to Jacobs [62]), and $k_1 = 1.5$ is an empirical constant. In accordance with Ghasemian et al. [61], choosing a representative pressure loss coefficient $k = 2$ (from the original measurement by Grunert et al. [63]), an initial value $C_{Dv,\text{ini}} = 0.59$ is obtained. As this value is a reasonable approximation for unsheltered tree arrays, the value is adjusted 20% smaller to account for the scattered sheltering within the urban area. Thus, the value $C_{Dv,\text{ref}} = 0.8 C_{Dv,\text{ini}} = 0.47$ is used as a reference in this study. To examine how the urban flow system is affected by this choice, the reference value is further scaled down (assuming a stronger impact from sheltering) by another 20% to get $C_{Dv,\text{mod}} = 0.8 C_{Dv,\text{ref}} = 0.376$. These two values represent a significant uncertainty margin in D_v, which provides the means to assess its necessary level of detail in relation to other modelling choices.

To probe the spatial variability of D_v, two alternative techniques are used to specify the *LAD* distributions. See Figure 4 for visual assistance. The first, higher level of detail method used as a reference relies on the work of Tanhuanpää et al. [64], who provided an (unpublished) upgraded register for the street trees in Helsinki. This register had been processed from the same open access Helsinki3D LiDAR dataset that has been employed here to generate the urban surface datasets. The register contains metrics for each $\sim$ 23,000 individual street trees within the city and includes percentile profiles of first pulse returns from the interior of the tree crowns. These vertical percentile plots provide a horizontally integrated plant area density profiles for each tree. This information alone cannot be used in the LES because the three-dimensional trees within the model are discretized by the LES grid cells and the required *LAD* information needs to be assigned via vertical profiles for each $\Delta_x \times \Delta_y \times (h_v(x_i, y_j) - h_0)$ grid column at location $x = x_i$ and $y = y_j$ that coincides with a tree crown. The tree height distribution $h_v(x, y)$ is obtained from the vegetation height raster dataset, and the starting height of the foliage is fixed at 4 m based on information from city officials stating that all

street and park trees are trimmed below 3–4 m. For the purpose of assigning *LAD* information to LES, a one-dimensional beta distribution profile $f(\alpha, \beta, z)$ (see Markkanen et al. [65]) with $\alpha = 2.4$ and $\beta = 1.7$ was manually optimized to yield, after horizontal averaging, the correct mean *LAD* profile shape of the three-dimensional street trees. The drag due to tree trunks is neglected in this study.

The second method is based on a low level of detail approximation that ignores all the vertical variabilities in *LAD*, thus employing a constant value in every grid cell within a tree crown. The value is derived from a reference, open grown urban deciduous tree (Tilia vulgaris), which is typical in Helsinki, of which the crown height is $h_v - h_0 = 11$ m and $LAI_{\texttt{ref}} = \int_{h_0}^{h_v} LAD\,dz = 13.0$ (note the branchless lower part of the trunk is not included), thus yielding $\langle LAD \rangle_z = LAI_{\texttt{ref}}/(h_v - h_0) = 1.18$ m^{-1} for the crown. This value falls within the normal range of urban trees, which grow higher density crowns than forest trees, and it is about half of the green ash shelterbelt crown *LAD* reported in Zhou et al. [66], although it is not clear wheather the reported values are two-sided or one-sided plant area densities. Typically *LAD* in Equation (4) is defined as *one-sided* leaf/plant area density [67].
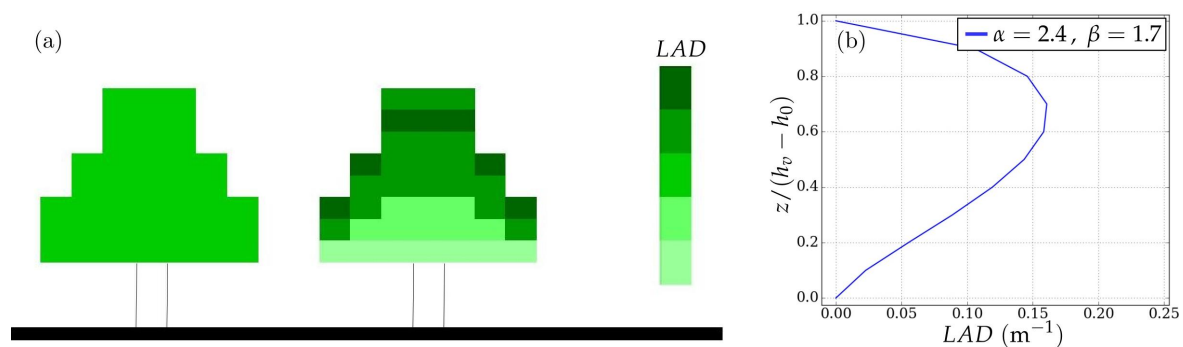
Figure 4. (**a**) A mock-up illustration of the two different leaf area density (LAD) distributions used in the study: In the low level of detail approach, every grid cell within a tree contains a fixed mean *LAD* value. In the higher level of detail model, each grid column respects the prescribed vertical profile processed from a point cloud dataset (**b**). The starting height of the tree crown $h_0 = 4$ m originates from a city policy that dictates that all branches below 3–4 m are trimmed. The two approaches always yield the same $LAI = \int_{h_0}^{h_v} LAD\,dz$.

2.5. Different Simulation Configurations

This numerical study consists of four urban LES simulations, which are based on an identical reference configuration but selectively altered in accordance with the stated objectives. The reference configuration, labelled `[R]`, adheres to the numerical method descriptions provided herein whereas the altered cases deviate from this reference only by the single declared model modification that is itemized in Table 3. The modified cases are labelled `[P]`, `[D]`, and `[B]` in accordance with the applied model changes in LAD [P]rofile, [D]rag coefficient, and [B]oundary layer height.

Table 3. A listing of all the LES cases included in this study and their respective modification presented in juxtaposition to the reference configuration.

Model Constituent	Reference Case	Modified Case
LAD profile:	$[R]: LAD(x_i, y_j, z) = f(\alpha, \beta, \frac{z - h_0}{(h_v - h_0)})$ for $h_0 \leq z \leq h_v(x_i, y_j)$ and $x_i, y_j \in \Gamma_{\texttt{b}}$	$[P]: LAD(x_i, y_j, z) = \langle LAD \rangle_{z, \texttt{ref}}$
Drag coefficient:	$[R]:$　$C_{Dv} = 0.47$	$[D]:$　$C_{Dv} = 0.376$
Neutral ABL height:	$[R]:$　$\delta \approx 20H$	$[B]:$　$\delta \approx 10H$

2.5.1. Data Acquisition from LES

After the initialization by precursor solution, the main simulations were run for $215T$, where $T = \frac{H}{u_*}$ is the turnover time for the eddies shed by the urban structures and an appropriate time scale for the urban canopy flow [12]. The precursor initialization method permits a relatively short $25T$ spin-up time for the nonequilibrium urban boundary layer flow. The allocated time is sufficient for flushing away the initial transients within the urban canopy, allowing the upstream ABL conditions to reestablish stationary conditions (adopted from the precursor solution).

All time series samples are recorded, and temporal averages are computed over $\Delta t_{avg} = 190T$, which is a sufficiently long time to accumulate well-converged statistics of low altitude urban canopy turbulence [15]. However, this study features a computational setup that allows the natural evolution of long streamwise velocity fluctuations, observed in ABL flows, and their interaction with the urban surface, introducing very long time scales that would necessitate up to 10 times longer averaging times to obtain stationary statistics for all the energy containing structures [26]. Limiting the sampling and averaging time in this study to $190T$ (which under conventional wind conditions translates to approximately 1.5–2 h) takes into account the issue of observability of local scale changes in real urban flows. Diurnally varying meteorology typically imposes restrictions on the length of the time series as changes due to alterations in the urban plan becoming muddled by the changes in meteorology. The temporal extent of the time series employed here with fixed meteorology is considered experimentally attainable (albeit rarely) under neutral ABL conditions. For example, Giometto et al. [5] employed urban tower data in 30 min blocks (here, $\sim 45T$) from the Basel Urban Boundary Layer Experiment (BUBBLE) experiment [68]. The use of longer time series, such as $650T$ by Castro et al. [19], $400T$ by Coceal et al. [12], and $300T$ by Xie and Castro [16] are difficult to justify when the objective is to establish the degree with which the changes to the street-canyon scale turbulence caused by the changes in the urban plan can be observed. Thus, the study of real urban boundary layer flows has to rely on time series that cannot provide sufficiently converged statistics for the very large scale motions of the ABL.

The following data are collected from each simulation for comparative analysis:

1. Horizontal 2D mean flow distributions are gathered from both the $\Omega^{(2)}$ and $\Omega^{(3)}$ domains such that $\overline{\mathbf{u}}(\mathbf{x})|_{z \in Z_{2d}}$, where $Z_{2d} = \{z_i : 0.35 \leq z_i/H \leq 7.2, i \in \mathbb{N} \text{ and } 1 \leq i \leq 12\}$. The planes at the last two elevations $z_{11}/H = 6.4$ and $z_{12}/H = 7.2$ are only extracted from $\Omega^{(2)}$.

2. A total of 11 virtual tower measurement stations is placed at locations $\mathbf{x}_l$, where $l \in \{\texttt{M1}, \texttt{M2}, \dots, \texttt{M11}\}$ refer to different station labels within $\Omega^{(3)}$. At each site, all prognostic variables are sampled at frequency $f = 80/T$ and vertical resolution of $\Delta_z = 1$ m. To tie resolution information into the comparison, the reference case [R] was used to gather two identical datasets, first, as described, from $\Omega^{(3)}$ and, second, from the solution of $\Omega^{(2)}$ with 2 m ($\Delta = H/10$) resolution. This data will be thus labeled [R2] throughout this document. The exact coordinate of [R2] virtual tower deviates from [R] by 0.5 m in both horizontal directions. This offset was examined to have a negligible effect on the results. The locations and labels, shown in Figure 5, are chosen by design to establish a representative sample of urban settings:

    *   M1: Narrow street canyon primarily aligned with the streamwise direction.
    *   M2: Same as above but with a different urban plan upstream.
    *   M3: Intersection at an opening with an isolated, dense formation of trees.
    *   M4: Narrow boulevard (street canyon with regular array of trees on both sides of the road) oriented perpendicular to the mean wind.
    *   M5: Upstream corner of a park with large deciduous trees.
    *   M6: Downstream corner of the abovementioned park.
    *   M7: Narrow streamwise aligned street canyon of which upstream is partially influenced by vegetation.

- M8: Wide street canyon with sporadic arrays of relatively small trees. Complex arrangement of city blocks upstream.
- M9: Further downstream location of the abovementioned street canyon. Upstream weakly influenced by trees.
- M10: T-intersection with a narrow street canyon upstream, a dense array of trees immediately behind the measurement station.
- M11: Center of a public square shielded by upstream buildings and trees.

All datasets are exposed to identical analysis, but in the interest of conciseness, the main text presents the analysis results of selected datasets.

topography height (m)   vegetation height (m)
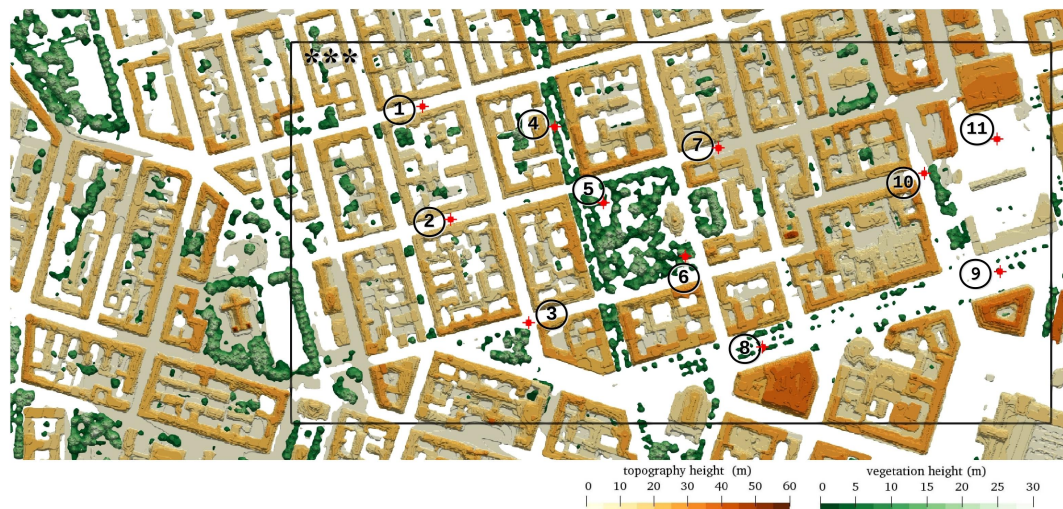
Figure 5. Locations of the flow data extraction sites (indicated by red dots next to circled identification number 1-11) within the $\Omega^{(3)}$ domain: In this document, the data sites are denoted by M1, M2, ... , M11. At each location, vertical profiles of prognostic variables are gathered at frequency $f = 80/T$. For conciseness, the main document only includes results from selected sites.

2.6. Analysis Methods

The different urban LES results are compared using methods in which the scope ranges from integral to local. At the most basic level, spatially and temporally averaged vertical profiles of elementary flow quantities are juxtaposed. This is followed by a statistical comparison of differences that arise between horizontal mean flow distributions. A particular attention is assigned to the comparative analysis of the location specific time series that are gathered from the network of virtual tower measurement sites. These datasets of LES-resolved turbulence are meticulously scrutinized to extract the coherent fluctuations they contain and to expose their respective frequency distributions. This is accomplished by the means of wavelet analysis, whereas the comparative analysis of the associated signal frequency content is achieved by exploiting the concepts of information entropy and divergence. The following sections introduce the relevant analysis methodologies used in the post-processing.

2.6.1. Statistical Evaluation of Differences in 2D Mean Flow Distributions

To gain a statistical insight of the level of differences that arise in the horizontal distributions of the $\overline{u}_1^+ = \overline{u}/u_*$ and $\overline{w}^+$ solutions due to the modelling choices, two statistical measures recommended by Chang and Hanna [69] are employed in the comparison. Root normalized mean square difference ($RNMSD$) gives a measure of mean difference that is composed of random scatter and systematic

bias, and fractional bias (*FB*) provides a specific measure for the systematic bias between two datasets. These metrics are defined as follows:

$$RNMSD\left(\psi\right) = \sqrt{\frac{\left\langle \left(\overline{\psi} - \overline{\psi}_{\mathrm{R}}\right)^2\right\rangle}{\left\langle \overline{\psi}\right\rangle \left\langle \overline{\psi}_{\mathrm{R}}\right\rangle}} \tag{6}$$

$$FB\left(\psi\right) = \frac{\left\langle \overline{\psi}\right\rangle - \left\langle \overline{\psi}_{\mathrm{R}}\right\rangle}{0.5\left(\left\langle \overline{\psi}\right\rangle + \left\langle \overline{\psi}_{\mathrm{R}}\right\rangle\right)}. \tag{7}$$

ψ is a generic variable from cases [P], [B], and [D], while ψ_{R} refers to the value from the reference simulation [R]. Because the double-averaged quantities for vertical velocities become very close to zero (particularly over $\Omega^{(3)}$ domain and at higher elevations), conditional evaluations $RNMSD\left(\psi|\psi > 0\right)$ and $FB\left(\psi|\psi > 0\right)$ are utilized instead where only the data points satisfying $\overline{\psi}_{\mathrm{R}} > 0$ are considered. The case $\overline{\psi}_{\mathrm{R}} < 0$ is also inspected but omitted for brevity as the results do not contribute new information.

In addition, a qualitative inspection of spatial dimension of mean flow differences is obtained via relative difference distributions defined for $\overline{\psi}(\mathbf{x})$ as

$$Rel\Delta\overline{\psi} = \left(\overline{\psi} - \overline{\psi}_{\mathrm{R}}\right)/|\overline{\psi}_{\mathrm{R}}|. \tag{8}$$

2.6.2. Joint Time–Frequency Analysis with Wavelets

A crucial step of our work is to analyze the time–frequency statistical properties in the fast-scale turbulent fluctuations of the velocity field:

$$\mathbf{u}'(\mathbf{x}_l, t) = \mathbf{u}(\mathbf{x}_l, t) - \overline{\mathbf{u}}(\mathbf{x}_l) \tag{9}$$

Here, $\overline{\mathbf{u}}(\mathbf{x}_l)$ is the time average of the flow at the chosen urban location $\mathbf{x}_l$ (see Section 2.5.1 for details). The objective is to gain insight about the statistics and frequency content of these fluctuations.

In the following subsections, we will illustrate the theoretical tools that will be deployed for our aim.

The first subsection will be dedicated to explaining the conventions about the Fourier transform and its discretization is adapted here; the Fourier filtering [70] methodology will be also explicated in detail. Such a procedure will be used to define the pure small-scale flow.

In the second subsection, we will introduce the wavelet transform. This mathematical instrument allows taking into account the intrinsic non-stationarity of the turbulent fluctuations, a common feature in inhomogeneous turbulence. Such a technique has already been used in environmental fluid mechanics literature [22]. Wavelet analysis will also make it possible to identify a threshold for the time scales of the flow under which we want to dedicate the present analysis. Consequently, all the scales above that threshold will be filtered out via Fourier analysis.

After we have obtained both the wavelet coefficients and the filtered fluctuations, we will analyze them from a statistical point of view. Unlike the approach presented in Hertwig et al. [22], where just sample kurtosis and skewness for wavelet coefficient histograms were considered, we will exploit other statistical indicators which provide a deeper view on the whole fluctuation probability density function. These indicators are Shannon and Rényi entropies as well as divergence functions and are commonly utilized in information theory [71–73]. Entropy functionals give us information about both the central region and tails of a given probability density function, while divergences make it natural to compare two density functions coming from different simulation configurations (e.g., boundary conditions, tree models, and resolution). By applying these tools to the time series, we will be precisely able to identify in which manner the simulation configurations affect the results with respect to city location, height, time instants, frequency, or time scales.

All the abovementioned quantities will be first introduced in their theoretical, continuous-time formulation. Therefore, it will be explained what conventions we will follow in the present paper to apply those functionals and transformations to our time series.

2.6.3. Fourier Transform: Filtering and Spectrum

Given the turbulent fluctuations $\mathbf{u}'(\mathbf{x}, t)$, an operation often performed is to analyze them by means of Fourier transform and antitransform. We refer to Appendix A for a thorough explanation about how to pass from continuous to discrete versions according to the conventions of the present paper and the rationale behind any following equation. Here, we will list their discrete expressions that are implemented to perform the data analysis of this work:

$$\widehat{\mathbf{u}'}(\mathbf{x}_i, f_n) \quad = \quad \sum_{k=0}^{N-1} \Delta t \, \mathbf{u}'(\mathbf{x}_i, t_k) \, e^{i\frac{2\pi kn}{N}} \tag{10}$$

$$\mathbf{u}'(\mathbf{x}_i, t_n) \quad = \quad \frac{1}{N\Delta t} \sum_{k=0}^{N-1} \widehat{\mathbf{u}'}(\mathbf{x}_i, f_k) \, e^{-i\frac{2\pi kn}{N}} = \frac{1}{N\Delta t} \sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor N/2 \rfloor} \widehat{\mathbf{u}'}(\mathbf{x}_i, f_k) \, e^{-i\frac{2\pi kn}{N}} \tag{11}$$

where $t_k = k\,\Delta t$, $\Delta f = 1/(N\Delta t)$, $f_n = n\,\Delta f$, and $\lfloor \bullet \rfloor$ stands for the floor function. Notice that, in the discrete framework of the Fourier analysis, the Δt is often dropped, as it represents a constant factor. However, here, it was retained so as to keep a direct contact to the continuous-time formulation and the wavelet transform that will be explained in the next section. As for the time series itself, all the information about it is contained in the index set $n \in \{0, \ldots, \lfloor N/2 \rfloor\}$, as the fluctuations are a real quantity.

When considering only small-scale fluctuations in the flow, of which the Fourier frequencies are higher than some cutoff f_c, slower Fourier modes will be removed via a plain, sharp high-pass filter. In practice, we are going to consider the small-scale fluctuations plus the average local offset. Therefore, the relationship implemented numerically to extract this quantity from the full fluctuation time series is as follows

$$\widetilde{\mathbf{u}}(\mathbf{x}_i, t_n) \quad = \quad \overline{\mathbf{u}}(\mathbf{x}_i) + \frac{1}{N\Delta t} \left[\sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{-n_c} \widehat{\mathbf{u}'}(\mathbf{x}_i, f_k) \, e^{-i\frac{2\pi kn}{N}} + \sum_{k=n_c}^{\lfloor N/2 \rfloor} \widehat{\mathbf{u}'}(\mathbf{x}_i, f_k) \, e^{-i\frac{2\pi kn}{N}} \right] \tag{12}$$

having denoted by n_c the index such that $n_c\Delta f = n_c/(N\Delta t) = f_c$. Again, we refer to Appendix A for the details behind the derivation of Equation (12) and its significance.

In the next sections, we are also going to consider another quantity related to the Fourier transform of the fluctuations, that is the *spectral density* [74] of a given velocity component u_i. Here, we calculate it as follows:

$$S_{u_i}(\mathbf{x}, f_k) = |\widehat{u_i}(\mathbf{x}, f_k)|^2 \Delta f \tag{13}$$

It is an even function of f as u_i is a real quantity. As a direct consequence of this, the mean squared velocity is related to the spectrum through the following identity:

$$\overline{(u_i)^2} = \sum_{k=0}^{N-1} S_{u_i}(\mathbf{x}, f_k)\Delta f \tag{14}$$

By looking at Equation (13), it is also apparent that the spectrum $S_{\widetilde{u}_i'}$ will be null for any $f < f_c$.

2.6.4. Wavelet Transform: Time–Frequency Joint Analysis

The limitation of the Fourier transform is that any information about the intrinsic non-stationarity of turbulence is lost. Indeed, the integration covers all the signal durations, and one loses any insight about variations in the flow spectrum with respect to time, be those large modes, vortices, coherent or

pulsating structures, intermittency, etc. Wavelet transform is the right instrument to overcome this issue, as it intuitively represents a *time-local* analysis with respect to different time-scales [75,76].

More precisely, a continuous wavelet transform (CWT) of the signal associated to a velocity component $u_i(\mathbf{x}, t)$, e.g., u_1, is defined as follows [75]:

$$W_{u_i}(\mathbf{x}, s, t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} u_i(\mathbf{x}, t)\Psi\left(\frac{\bar{t} - t}{s}\right) d\bar{t} \tag{15}$$

The graph of the function $W_{u_i}(s, t)$ on the plane (s, t) is called *scalogram*. The function $\Psi(\bar{t})$, which is complex in general, is called mother wavelet and generalizes the complex oscillating exponential functions in the Fourier transform. The mother wavelet generates a wavelet family by means of the dilation factor s and the time shift t in the sense that the analysis is carried out at time t of the velocity u_i. There exist several mother wavelets, any of which must satisfy the so-called admissibility criteria. One of these criteria is that the mother wavelet must have null mean value. Consequentially, the wavelet transform of u_i and u_i' are the same, as can be promptly noticed in Equation (15). The mother wavelet we will be utilizing in this work is a common simplified version of the complex *Morlet wavelet*:

$$\Psi_{\omega_{0\star}}(t_\star) = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t_\star^2} e^{i\omega_{0\star}t_\star} \tag{16}$$

where $t_\star = (\bar{t} - t)/s$ (see Equation (15)) is a non-dimensionalized, t-centered time and the dimesionless parameter $\omega_{0\star}$, which in turn generates a *family* of mother wavelets, can be freely chosen depending on the case considered. For the approximated form in Equation (16) it is common practice to choose it at least > 5, and we will set it to 6. The reasons behind these practices and choices are thoroughly explained in Appendix B. For our purposes, what is meaningful to pinpoint is that, thanks to this choice, the typical wavelength of the wavelet is simply $\sim s$. This biunivocal correspondence provides us with a prompt physical interpretation of the decomposition of the velocity u_i at time t in many wavelet *pseudo-harmonics* of scale s.

In reality, for our analysis, it will be expedient to consider the wavelet *power* scalogram $|W_{u_i}(\mathbf{x}_l, s, t)|^2$, which is defined as the graph of the square module of the wavelet transform with respect to the parameters (t, s). For the reasons we have illustrated previously, this graph can be promptly interpreted at a glance as the kinetic power of the velocity component concentrated at a certain scale s and location $\mathbf{x}_l$ thanks to the immediate connection between the parameter s and the time scales involved in the processes.

As for the discretization of Equations (15) and (16), we need to account for finite size effects and aliasing. Taking into account that $\Delta t = 0.5$ s is the time step in the time series $\mathbf{u}(\mathbf{x}_i, t_n)$, only time scales larger than 1.4 s can be resolved in the wavelet transform (for proof of this, see again Appendix B).

To calculate the wavelet transform in Equation (15), we just take the discrete convolution over a sequence of scales s_k:

$$W_{u_i}(\mathbf{x}_l, s_k, t_n) = \frac{1}{\sqrt{s_k}} \sum_{m=-\infty}^{\infty} u_i(\mathbf{x}_l, t_m)\Psi\left(\frac{t_m - t_n}{s_k}\right) \Delta t \tag{17}$$

with the convention that $u_i(\mathbf{x}_l, s_k, t_n) = 0$ when $n < 0$ or $n > N - 1$. The sum in Equation (17) thus is over finite number of elements in reality. Due to the 2s-wide spread of the Gaussian envelop in Ψ, this fact will result in finite-size effects. Namely, given a scale s_k, each time step t_n such that $t_n \leq 2s_k$ or $t_{N-1} - t_n \leq 2s_k$ will be affected by this finite-size effect in time. Such a distortion needs caution especially for large modes, where it may affect significantly the wavelet transform. This effect constrains the practical use of time–frequency analysis for large modes and, in practice, poses an admissible maximum for the scale s_k that is significantly lower than $t_{N-1}/2$. The distortion in time at the beginning and the end of the wavelet transform is however easily visible in the scalograms; therefore, it can be handled safely. Moreover, as already said previously in this work, we are going to

focus only on small-scale fluctuations, an aspect that makes these finite-size effects rather negligible in the present analysis.

2.6.5. Entropy and Divergence of Turbulent Signals

We illustrate the technique implemented here to estimate the Shannon entropy of the turbulent fluctuations and the related divergence between two different turbulent signal density functions. As we will see, these are natural tools to compare the turbulent fluctuations statistics between different simulation configurations.

Given a probability density function $\rho(a)$ for a continuous random variable a, we define Shannon entropy as the following nonlinear functional of ρ:

$$S_1(a) = - \int \rho(a) \ln \rho(a) \, da, \tag{18}$$

By looking at the Equation (18), the connection to the thermodynamic entropy definition from statistical mechanics is evident [77].

The definition in Equation (18) can be adapted to discrete contexts as well, and the latter is the formulation we will be using. The form it takes when a random variable b assumes N discrete values b_k, $k = 1, \ldots, N$ is as follows:

$$S_1(b) \quad = \quad - \sum_{k=1}^{N} P(b_k) \ln P(b_k) \tag{19}$$

By comparing this formula to Equation (18), it must be noticed that it does not correspond to the continuous limit when k becomes a continuous variable, that is, when the bin width $\Delta_k^* \to 0$, $N \to \infty$, taking into account that $P(b_k) = \rho(b_k)\Delta_k^*$. See Bonachela et al. [72] for further details.

Consider now a time series $a(t_1), \ldots, a(t_N)$ representing a time-sample of the random variable a (which in this study will be the time series of the squared wavelet coefficients of the small-scale fluctuations at each location $\mathbf{x}_i$). A histogram is constructed from the time series values by dividing the sample values into M nonuniform interval bins $\Delta_0^*, \ldots, \Delta_M^*$. These intervals will be chosen so that extreme events will be associated to larger intervals in order to capture better the small probability intervals and to improve the convergence of the histogram.

If the underlying quantity is theoretically continuous—be it the filtered velocity or the wavelet coefficients—then the probability of having an observation at a certain $\bar{t}$ within an interval Δ_k^* is given by the following:

$$P_k = P[a(\bar{t}) \in \Delta_k^*] = \int_{\Delta_k^*} \rho(a) da \tag{20}$$

This can be straightforwardly estimated by means of the associated discrete sample frequency $f^*(\Delta_k^*)$ in the given time series.

Once we measure the sample frequencies, we can plug them into Equation (19) with $f^*(\Delta_k^*)$ in the place of P_k:

$$\mathcal{H}_1(a) \quad = \quad - \sum_{k=1}^{M} f^*(\Delta_k^*) \ln f^*(\Delta_k^*) \tag{21}$$

The resulting quantity represents an estimator of the Shannon entropy. Such a simple frequentist estimator proves to be asymptotically unbiased [72,73]. In other words, although the expected value of the sample frequency $\mathbb{E}[f^*(\Delta_k^*)]$ is P_k, due to the nonlinearity of the entropy functional, $\mathbb{E}[\mathcal{H}_1[a]] \neq S_1(a)$; namely, there is a bias.

However, this bias goes asymptotically to zero as the sample size goes to infinity. Given that we herein deal with large samples, the asymptotic limit $\mathbb{E}[\mathcal{H}_1[a]] \to S_1(a)$ is recovered.

Finally, let us introduce the so-called Shannon entropy divergence functional for continuous density functions:

$$\mathcal{D}_1(\rho_1 \| \rho_2) \quad = - \quad \int \rho_1(a) \ln \frac{\rho_2(a)}{\rho_1(a)} \, da \tag{22}$$

Oftentimes, this functional is also called Kullback–Leibler divergence, albeit based on Shannon entropy. Divergence measures how much some new probability density function ρ_2 is different from a reference ρ_1. In a sense, it represents the increase of information that ρ_2 contains in comparison to ρ_1. A geometrical interpretation of this can be found in Xu and Erdogmuns [78]. When $\rho_1 = \rho_2$, the divergence is 0.

Similarly to the entropy case, we are going to estimate the divergence by first estimating the probabilities empirically. To do so, we will be constructing a histogram per different configuration of our numerical simulations by dividing up the value range of the time series into M bins $\Delta_1, \ldots, \Delta_M$. For the sake of consistency, we will shortly see that the bins of the histograms will need to be the same for all the configurations. Looking at Equation (20), we will now have M sample frequencies $f_n^*(\Delta_1^*), \ldots, f_n^*(\Delta_M^*)$ per simulation configuration $n \in \{ \text{[R], [P], [B], [D], [R2]} \}$. We will thus be able to estimate the discrete versions of Equation (22), by plugging the relative sample frequencies f^* in the place of the bin probabilities:

$$\mathcal{D}_1(f_1^* \| f_2^*) \quad = \quad -\sum_{k=1}^{M} f_1^*(\Delta_k^*) \ln \frac{f_2^*(\Delta_k^*)}{f_1^*(\Delta_k^*)} \tag{23}$$

By looking at Equation (23), it is now evident that this methodology is well defined only if the endpoints of the bins Δ_k^* are the same. Note also that, unlike the entropy functional, from Equation (23), the continuous limit in Equation (22) is recovered when we go to the continuous limit.

The current approach to estimating the divergence via the sample frequencies in the histograms is simple and consistent as we deal with fairly large samples, but certainly, it is not optimal as the convergence rate. Other more complex, sophisticated estimation methods are available in literature, making use of nearest-neighbor techniques or other machine learning algorithms [78,79]. However, the principal interest here does not lie in a precise quantitative estimation of entropy or divergence values but rather in a comparative analysis of the differences that arise in LES-simulated urban turbulence due to level of detail modifications in the model. The presented methodology will therefore be suitable to assess these differences with respect to frequency (or time scale) at different locations and for varying elevations.

3. Results and Discussion

3.1. Similarity Scaling Parameters

The comparison of results from different UBL flow realizations necessitates normalization with appropriate scaling parameters. In ABL flows with surface roughness and a well-defined ISL, this is typically achieved by employing the ABL adopted form of the logarithmic law of the wall

$$\frac{\overline{u}}{u_*} = \frac{1}{\kappa} \log \left(\frac{z - z_d}{z_0} \right), \tag{24}$$

where the friction velocity u_* specifies the characteristic velocity scale, the displacement height z_d sets the aerodynamic "zero level" for the logarithmic profile, and the roughness length z_0 provides a parametrization for the effect of drag due to surface roughness. The theoretical validity of this similarity solution for neutrally stratified boundary layers is roughly confined to elevation range $z_d < z < \delta/3$ [80], where the δ is either the boundary layer height or the channel centerline height.

Despite the nonequilibrium nature of the studied flow, the double-averaged (DA) vertical momentum flux profiles extracted from the $\Omega^{(3)}$ domain exhibit a linear section in the range $1.5 < z/H < 4$, indicating a region where, on the average, the vertical momentum flux is balanced by the negative mean pressure gradient. Even though this does not imply the existence of an ISL, this overall balance provides a robust mean to define $u_* = \left\langle \overline{u'w'}^2 + \overline{v'w'}^2 \right\rangle^{1/4}\Big|_{z \to z_d}$, which is obtained by extrapolating the linear slope in the flux profile down to the displacement height z_d. The obtained values for u_*, z_d, and z_0 are listed for each case in Table 4. The values for z_d and z_0 are obtained from a logarithmic fit to the narrow band $1.5 < z/H < 4$ in the streamwise velocity profile. The width of the band was determined to be most suitable for the logarithmic fit by trial and error.

Table 4. Similarity scaling parameters from the different computational cases.

Case	Friction Velocity u_* (m s^{-1})	Roughness ength z_0/H ($-$)	Displacement Height z_d/H ($-$)
[R]	0.539	0.063	1.08
[P]	0.521	0.063	1.08
[B]	0.505	0.067	1.05
[D]	0.535	0.066	1.04

In the scaling parameters, case [B] with the largest ABL eddies removed, brings about the most notable deviations. Compared to the reference case, friction velocity became approximately 6% lower, the roughness lengths became 6% higher, and the displacement height became 3% lower. Other small but observable deviations include case [D], with reduced canopy drag, causing a 4% decrease in displacement height and a 5% increase in roughness length. Simplifying the *LAD* profiles of trees resulted in negligible changes in the scaling parameters.

It is acknowledged that the use of log-law scaling parameters is problematic because the inherent assumptions concerning equilibrium conditions and the existence of an ISL above an urban canopy are generally violated. For sufficiently developed urban boundary layer flows, the logarithmic velocity profile manifests above the urban canopy only via horizontal averaging [10,12,15]. Therefore, it is difficult to obtain reliable log-law scaling parameters from urban field studies as horizontal averaging would require a dense network of towers reaching the elevation range $1.5 < z/H < 4$ that, for real urban flow scenarios (as in the considered study), falls entirely within the RSL. In this comparative study, the log-law scaling is found to provide the most robust mean to non-dimensionalize the flow data, but it is likely that other scaling techniques should be used in real urban validation studies.

3.2. Horizontal Mean Flow Distributions

As expected, the DA profiles can only provide information on how the chosen modeling alterations influence the urban flow system at an integral scale. Next, the mean flow distributions are compared in order to establish how the modeling choices embedded within the structural complexity of a real city alter the numerical predictions of urban wind. Figure 6 exhibits the horizontal distribution of $\overline{u}_1^+$ across the $\Omega^{(2)}$ domain at $z/H = 0.7$ obtained from case [R] with $\Delta = H/10$ resolution. This wind map highlights the heterogeneity and complexity of urban canopy flow within a real built environment featuring a network of street canyons of different widths, buildings and building blocks of varying dimensions, and openings and parks with different vegetation covers. Close inspection of these results would warrant a long report of interesting observations, but such examination falls outside the scope of this comparative study.
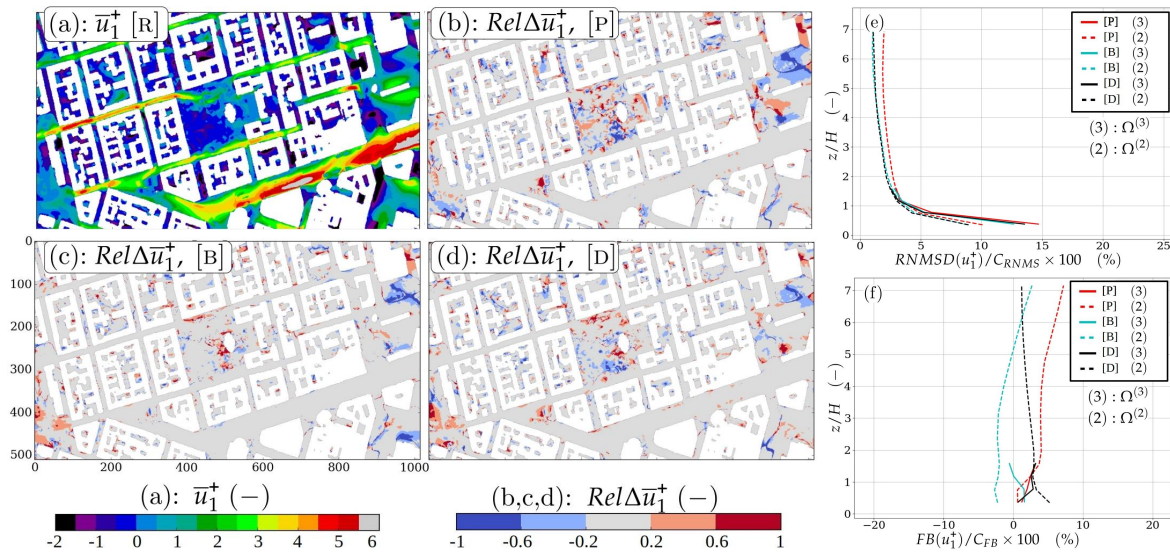
Figure 6. Horizontal distribution of $\overline{u}_1^+$ at $z/H = 0.7$ from case [R] across the urban area within the $\Omega^{(2)}$ domain featuring $\Delta = H/10$ resolution.

Qualitative and quantitative comparisons of $\overline{u}_1^+$ and $\overline{w}^+$ distributions across 2D planes described in Section 2.5.1 (see item 1) are depicted in Figures 7 and 8, respectively. Figures 7a and 8a display the velocity distributions of the reference [R] at elevation $z/H = 0.35$ across domain $\Omega^{(3)}$, whereas the Figures 7b–d and 8b–d show the relative difference distributions (see Equation (8)) between [R] and cases [P], [D], and [B], respectively. The quantitative comparison employs the statistical metrics *RNMSD* and *FB* (Section 2.6.1) of which the vertical variation for both $\overline{u}_1^+$ and $\overline{w}^+$ velocity components are plotted in Figures 7e,f and 8e,f. The metrics are evaluated across $\Omega^{(2)}$ for heights $0.35 < z/H < 7.2$ and across $\Omega^{(3)}$ for heights $0.35 < z/H < 1.6$, and they are presented relative to their recommended maximum limits C_{RNMS} and C_{FB} reserved for successful experimental validation [69]. Only the low altitude values from $\Omega^{(3)}$ are considered because the metrics become poorly defined when the double-averaged values approach zero (which occurs with $\langle w \rangle$ in $\Omega^{(3)}$ right above the canopy) and because the importance of improved resolution is particularly relevant within the urban canopy.

Despite emphasizing differences that occur when the reference values are small, $Rel\Delta\overline{u}_1^+$ and $Rel\Delta\overline{w}^+$ distributions successfully reveal the localized nature of the deviations that arise due to the model modifications. This is particularly notable with case [P], where the vertical *LAD* distributions of trees are simplified. The changes in the immediate proximity of trees within the densely vegetated park at the domain center demonstrate that the accuracy in *LAD* profiles becomes relevant if high precision is sought in the vicinity of vegetation. On the other hand, making a 20% error in specifying the vegetation drag coefficient (case [D]) leads to a less pronounced effect within the park, and the elimination of the largest ABL eddies (case [B]) has the least influence on the local details.

Juxtaposing $\overline{u}_1^+$ and $\overline{w}^+$ distributions and relative difference distributions conveys the inherent difficulty associated with predicting vertical velocity variations within a densely built urban environment. A real urban canopy gives rise to a flow system involving a complex labyrinth of interacting blunt-body wakes and street canyon vortices, which manifests as a rapidly sign-altering vertical mean flow distribution. Therefore, if the objective is to predict the up and down drafts within an urban canopy, all aspects concerning model accuracy must be carefully considered. The $Rel\Delta\overline{w}^+$ distributions in Figure 8b–d offer an instructive view of the uncertainty embedded in predicting vertical movements of the wind and the spatial dimension of wakes and vortecies within a real city.
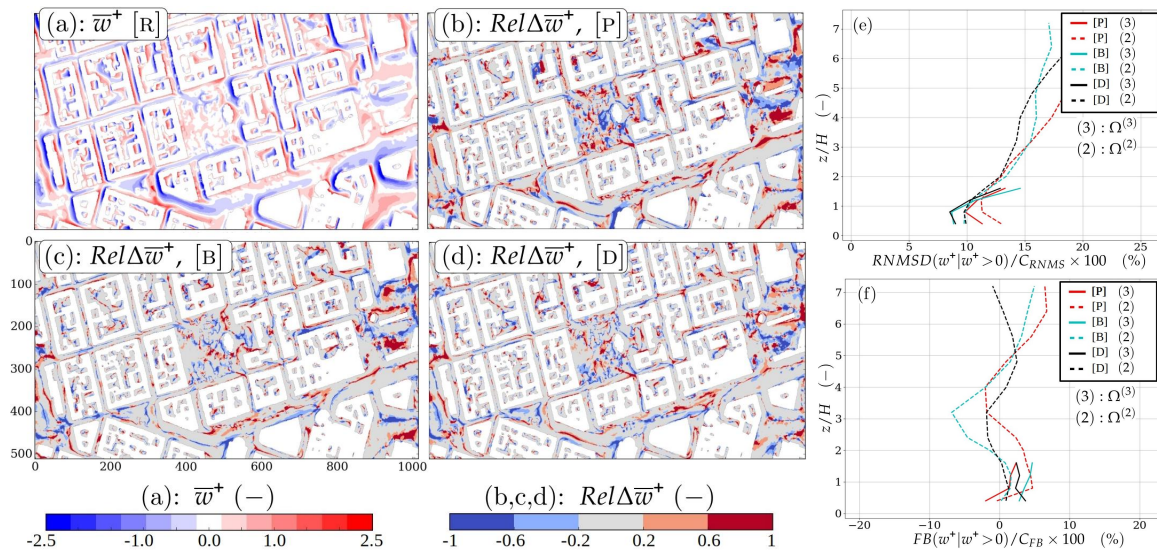
Figure 7. (**a**) Horizontal distribution of $\overline{u}_1^+$ at $z/H = 0.35$. (**b**–**d**) Relative difference distributions between cases [P], [B], and [D]. (**e**) Vertical variation of root-normalized-mean-square difference $RNMSD(\overline{u}_1^+)$ and (**f**) fractional bias $FB(\overline{u}_1^+)$ evaluated over 12 horizontal planes across $\Omega^{(2)}$ for heights $0.35 < z/H < 7.2$ and over 4 planes across $\Omega^{(3)}$ for heights $0.35 < z/H < 1.6$. Recommended criteria suggested by Chang and Hanna [69] are $RNMSD < C_{RNMS}$ ($= 1.22$) and $|FB| \leq C_{FB}$ ($= 0.3$), whereas zero is the ideal value for both.

　　Values from the larger but coarser $\Omega^{(2)}$ domain provides statistically convergent measures across the UBL, whereas values from $\Omega^{(3)}$ offer evidence on the level of deviations within the highly resolved urban canopy. On both domains, the first 4% and last 10% of the grid points in the streamwise direction are not used in the computations because they belong to one-way nesting border zones within which the finer solution adopts to the coarser one. Note that the entire domain $\Omega^{(3)}$, including its border zones, are included in the Figures 7 and 8a–d visualizations.

　　Figure 7e depicting the vertical profile of $RNMSD(\overline{u}_1^+)$ relative to its recommended maximum limit $C_{RNMS} = \sqrt{1.5}$ lays bare how modeling discrepancies in the horizontal wind become larger within the urban canopy close to the ground but diminishes rapidly above the roof height where mean wind speed increases. However, even the highest discrepancies yielded by the case [P] over $\Omega^{(3)}$ at low altitudes satisfy the criteria $RNMSD < C_{RNMS}$ with a significant margin, ensuring that the considered model uncertainties take up only a minor portion of the target margin. Similarly, the vertical profiles of $FB(\overline{u}_1^+)$, plotted relative to its recommended magnitude limit $C_{FB} = 0.3$ in (f), show only weak systematic bias between the simulations (consistently below 5% of desireable limits within range $0.35 < z/H < 5$). The relatively uniform fractional bias profile implies that the discrepancies reflected in $RNMSD(\overline{u}_1^+)$ close to the ground are principally random in nature [69]. Thus, in an integral sense, the presented solution for the horizontal wind exhibits robustness against the imposed modelling parameter changes away from the ground, but close to the ground (or pedestrian level), these sensitivities increase.

　　The conditionally sampled $RNMSD(\overline{w}^+)$ and $FB(\overline{w}^+)$ results in Figure 8e,f also reveal that the discrepancies in vertical velocity are primarily random in nature. The $RNMSD(\overline{w}^+)$ values within the urban canopy are confined to very modest $\sim 10\%$ levels; case [P] (with simplified LAD profile) again exhibits marginally higher deviations close to the ground than others. Nonetheless, the results are very acceptable given the complexity of the urban canopy flow. The steady increase of $RNMSD(\overline{w}^+)$ values with height reflects the difficulty associated with the normalized differences as the mean values approach zero.

Figure 8. (**a**) Horizontal distribution of $\overline{w}^+$ at $z = 0.35H$. (**b–d**) Relative difference distributions between cases [P], [B], and [D]. (**e**) Vertical variation of conditional root-normalized-mean-square difference $RNMSD(\overline{w}^+|\overline{w}^+>0)$ and (**f**) conditional fractional bias $FB(\overline{w}^+|\overline{w}^+>0)$ evaluated over 12 horizontal planes across $\Omega^{(2)}$ for heights $0.35 < z/H < 7.2$ and over 4 planes across $\Omega^{(3)}$ for heights $0.35 < z/H < 1.6$. Recommended criteria suggested by Chang and Hanna [69] are $RNMSD < C_{RNMS}$ $(= 1.22)$ and $|FB| \leq C_{FB}$ $(= 0.3)$, whereas zero is the ideal value for both.

3.3. Local Turbulence Profiles

The practical limitations in conducting measurement campaigns within real urban environments assures that locally obtained time series data remain an essential point of contact between model results and experimental measurements. For this reason, it is important to examine how observable the changes caused by the model modifications are in the virtual tower measurements described in Section 2.5.1. The structural composition of real cities is generally so complex that no individual measurement station can provide street canyon flow information that is representative of the whole city, but neighborhood scale representativeness is attainable.

3.3.1. Mean Flow

The vertical profiles of dimensionless streamwise-rotated velocity, obtained from all 11 virtual tower locations shown in Figure 5, are depicted in Figure 9. The set of figures also includes the DA profile $\langle \overline{u}_1^+ \rangle$ from domain $\Omega^{(3)}$ to mark the contrast between local and spatially averaged profiles. The profiles also contain bootstrapped 95th confidence intervals (computed using 1000 iterations) in grey color, but they remain within the line width. The local profiles now also include the case [R2], where the flow data is collected from the reference simulation in an identical manner as data labeled [R] but from $\Omega^{(2)}$ featuring $\Delta = H/10$ resolution. Thus, [R2] provides an exceptional view of how resolution affects the local model results. This evidence is essential for deciding to what degree the nesting technique should be utilized in urban LES simulations.
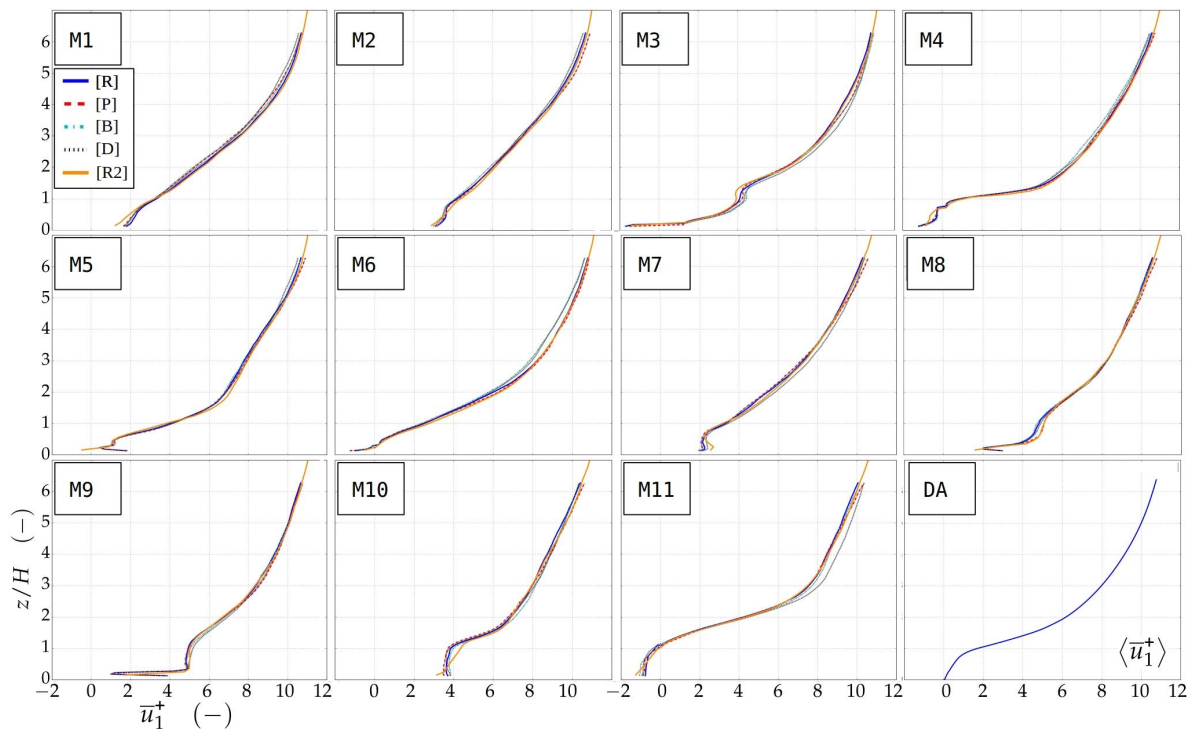
Figure 9. Normalized streamwise-rotated velocity profiles in the vertical direction obtained from each modified model simulation and all 11 virtual tower locations: A double-averaged profile is included in the bottom-right-hand corner for reference.

The comparison of these local profiles confirm that the imposed model changes bring about only small deviations in the mean flow. Even at stations in the immediate vicinity of tall vegetation (e.g., M3, M4, M5, M6, and M10), the agreement between the $\Delta = H/20$ resolved cases remains strong and the observable deviations do not reveal any systematic trends. Even stations M9 and M11, that are in the last 10% of the $\Omega^{(3)}$ domain where the solution is affected by the one-way nesting treatment at the outflow border, feature only minor differences, where $z/H < 2.5$ despite the dramatically different flow conditions.

On the other hand, the profiles from the $\Delta = H/10$ resolved [R2] case consistently exhibit deviations close to the ground and near high gradients. Thereby, we can conclude that, if local predictability in mean wind is sought, the effect of resolution outweighs the level of detail variations considered herein.

The dimensionless vertical velocity profiles (with their bootstrapped 95th confidence intervals) of four selected stations are displayed in Figure 10. The profiles reflect the difficulty in predicting the swirls and vortices within the urban canopy. At the shown scale, the deviations appear notable, but again, the $\Delta = H/20$ resolved profiles closely adhere to each other whereas the $\Delta = H/10$ resolved [R2] profiles depict the most discernible deviations, reaffirming the relevance of sufficient resolution in urban flows. A careful inspection reveals that making a consistent 20% mistake in vegetation drag (case [D]) has a weaker impact on the mean results than simplifying the LAD profile (case [P]) within the urban canopy. Also, eliminating the largest ABL eddies (case [B]) brings about only minor and random deviations compared to the reference case.
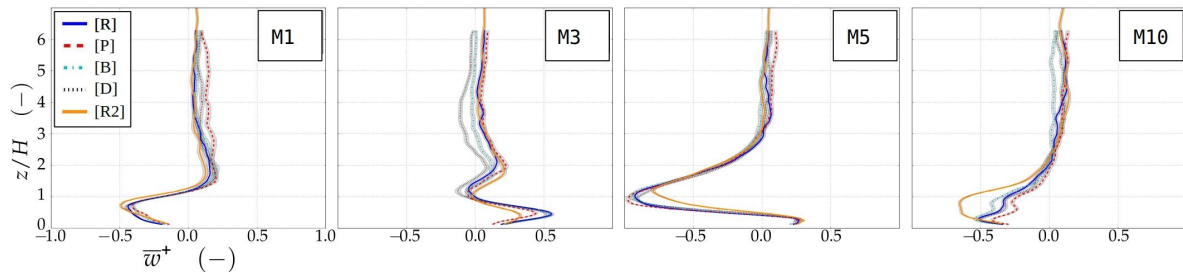
Figure 10. Normalized vertical velocity profiles obtained from each modified model simulation at selected virtual tower locations.

3.3.2. Velocity Variance

Mean wind profiles provide a very limited view of the interaction between atmospheric boundary layer turbulence and urban environment. Therefore, this study places emphasis on examining the dynamics of the LES-predicted urban canopy turbulence. Here, the issue of observability is also considered relevant as real urban canopy flows always contain diurnal variation, which complicates the sampling and analysis of time series sufficiently long to provide convergent statistics. For this reason, the sampling period of $190T$ used here in, or equivalently $150T_\delta$ as $T_\delta = \frac{\delta}{\langle \overline{u} \rangle (z=\delta)}$, is considered an appropriate compromise between seeking convergent statistics and maintaining relevance in connection to real observations. (This time period translates to ~ 1.5 h of real time under typical wind conditions.)

The local vertical profiles of dimensionless variance plotted in Figure 11 (excluding stations M9 and M11 as they reside very close to the nesting boundary zone) for resolved u_1 lay bare that the variance values within the roughness sublayer where $z > H$ have not converged properly. The drastic scatter in local σ_{u1}^2 / u_*^2 profiles occurring between measurement stations implies that the urban roughness sublayer does contain the slowly evolving VLSM structures of which the time scale is much longer than T or T_δ. Recall that the computational setup has been designed to allow these long streamwise perturbations, observed in ABL measurements, to evolve and subsist but without observable model-related amplification.

Although each simulation is identically initialized, the feedback from the urban surface to the recycling plane, no matter how weak, results in different flow realizations for each computational case. The resulting scatter in variance profiles offers evidence that the slowly evolving modes are localized, as made apparent by the local discrepancy, for example, in case [D] profiles at stations M1 and M4 versus M7. In general, the scatter diminishes as statistical convergence improves within narrow street canyons (stations M1, M2, M4, M7, and M10) and densely vegetated parks (stations M5 and M6) where the surrounding structures and obstacles effectively convert larger turbulent structures into smaller ones, enhancing the energy transfer toward the small scale eddies and hence also dissipation. The wider street canyons, particularly if parallel to the flow, allow the energetic large scale structures to penetrate the lower altitudes (see station M8), and therefore, poor convergence of variance is observed also within such street canyon, similar to the higher altitudes.
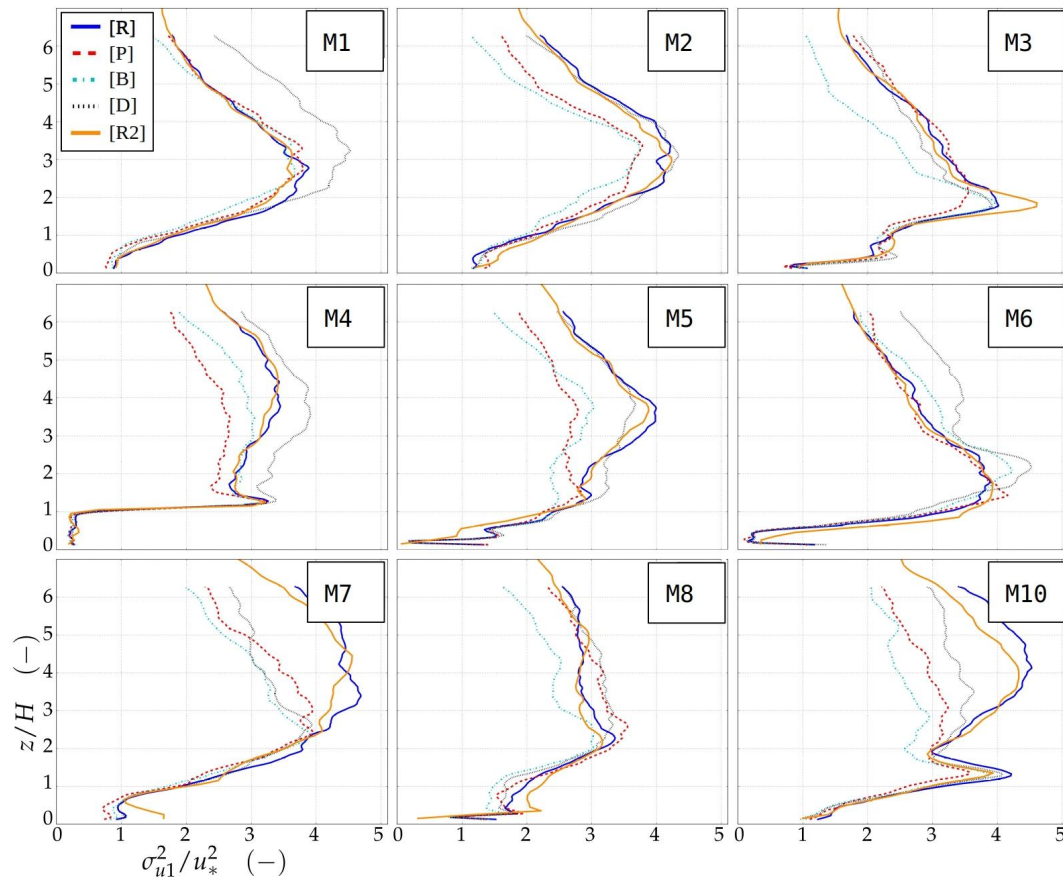
Figure 11. Local vertical plots of streamwise velocity variance σ_{u1}^2/u_*^2 from selected measurement sites: The contribution of statistically insufficiently described very-large-scale turbulence motions is distinctly visible above the urban canopy.

The profiles of case [B], with the largest ABL eddies removed, when compared to other cases for $z/H > 1$ exhibit consistently lesser variance values and the least amount of local variation between stations. Therefore, it is natural to conclude that the inclusion of the largest ABL eddies in the numerical model is conducive to the existence of the largest streamwise modes. However, even in case [B], these modes exist as seen when juxtaposing M3 and M6 profiles in the upper part of the roughness sublayer.

The variance results lay bare the imminent complications that arise when seeking agreement between simulated and measured UBL flow data; in the absence of idealized conditions or a monotonous urban plan, the observed *TKE* within the roughness sublayer is likely to contain low-frequency fluctuations of which the spatial characteristics may be difficult to uncover from tower measurements. It is thereby considered vitally important to employ analysis techniques which expose the dynamic makeup of the urban flow system. In this study, time–frequency analysis with wavelets is exploited in the dissection of the obtained time series in a manner that is also considered highly relevant when experimental urban flow data is considered. Through wavelet analysis, the comparative study can be broadened to encompass also the frequency content of the turbulence signal.

3.4. Time–Frequency Analysis

Each time series for $u_1(t, z)$ obtained from the virtual towers stations is subjected to continuous wavelet transform (Equation (15)) using complex Morlet wavelet with center frequency $\omega_0 = 6$ to generate a power scalogram $|W_{u_1^+}(t,s)|^2$ for each z-coordinate. Example power scalograms obtained at elevations $z/H = 0.6$ and $z/H = 1.7$ from the M3 station are shown in Figure 12. The scalogram features time t (in seconds) on the horizontal axis and scales $s = \frac{\omega_0}{2\pi f}$ (in seconds) on the vertical axis. Due to the

convenient choice for ω_0, there is a close inverse relationship between scales and frequency, i.e., $s \approx f^{-1}$. The figure excludes scales $s/T > \approx 4$, thus, excluding wave lengths $\lambda_w > \langle \overline{u_1} \rangle / f = \langle \overline{u_1} \rangle \, 4T$ that exceeds δ at elevation $z/H = 2$.
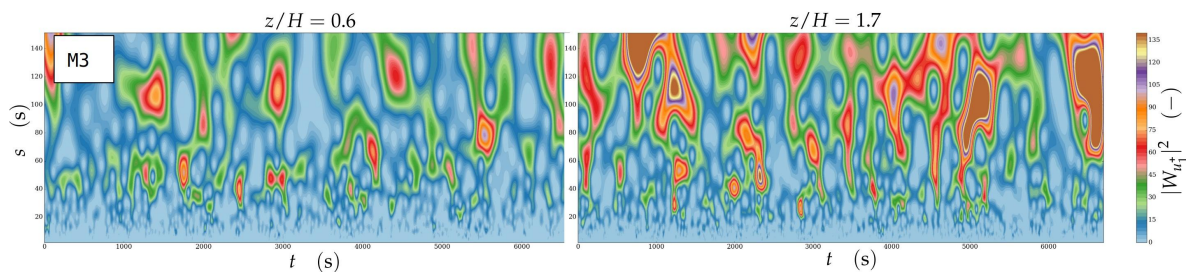
Figure 12. Example power scalograms of streamwise velocity time series from station M3 at $z/H = 0.6$ (**left**) and $z/H = 1.7$ (**right**), collected from the reference simulation [R].

The power scalogram $|W_{u_1^+}(t,s,z)|^2$ provides a time-resolved breakdown of the fluctuation energy content of u_1 into contributions from different frequency ranges. It offers a thorough view into the energy composition of the obtained turbulence signal and can be subjected to further analysis via different approaches.

In this comparative study, the power scalograms are first temporally averaged to uncover how the variance values are constructed from different frequency contributions. (Power spectrum via Fourier analysis could offer the same information.) This procedure is applied to u_1 time-series profiles obtained from the virtual tower measurements. The results from cases [B], [R], and [R2] at stations M1, M3, and M6 are depicted in Figure 13. The figure contains $\overline{|W_{u_1^+}|^2}(s,z)$ contour maps that are juxtaposed with corresponding σ_{u1}^2/u_*^2 variance plots (variance corresponds to the integral of $\overline{|W_{u_1^+}|^2}$ over all scales). Note the different range on the z-axis for [R2] results obtained from domain $\Omega^{(2)}$.

The role of the LSM and VLSM ($s > 2.5T$) contributions in variance plots for $z/H > 1$ becomes immediately apparent from the displayed $\overline{|W_{u_1^+}|^2}$ distributions. Particularly, attention is drawn to the contributions within range $1 < z/H < 4$ that highlight a confined energetic region, which arises from the interaction between the ABL turbulence and the urban canopy layer. The amplification of coherent structures within this elevation range is substantiated by the [R2] contours and variance profiles, which reveal how the effect of the urban surface vanishes for $z/H > 5$, providing an estimate $\sim 5H$ for the UBL height. These bulging UBL regions of elevated $\overline{|W_{u_1^+}|^2}$ values appearing in both [R] and [B] cases display remarkable similarity despite the complexity in the distributions. This provides evidence that the amplification of coherent large-scale structures just above the urban canopy height does not (significantly) involve the larger ABL eddies that scale with the boundary layer height.

The effect of LES resolution ([R] vs. [R2]) is distinctly observable for $z < H$ in results from stations M3 and M6. The coarser solution allows large-scale modes to energize the flow at near-ground elevations, whereas the improved resolution facilitates smaller eddies and higher velocity gradients which, in turn, lead to greater kinetic energy dissipation. The effect of resolution on variance within the park trees (station M6) is particularly notable.

Establishing the spatial nature and dynamic behavior of these amplified coherent large-scale structures, possibly characteristic to developing UBL flows, deserves to be pursued further, but this falls beyond the scope of this work. Within the context of this comparative investigation, these slow turbulent modes near the urban canopy present difficulties as they smear all the statistical differences that arise from the model variations within the roughness sublayer. Thus, guided by the power scalogram results, the statistical analysis will be confined to the sufficiently converged frequency/scale range that nonetheless covers all the turbulent length scales that are directly influenced by the vegetation model LOD variations.
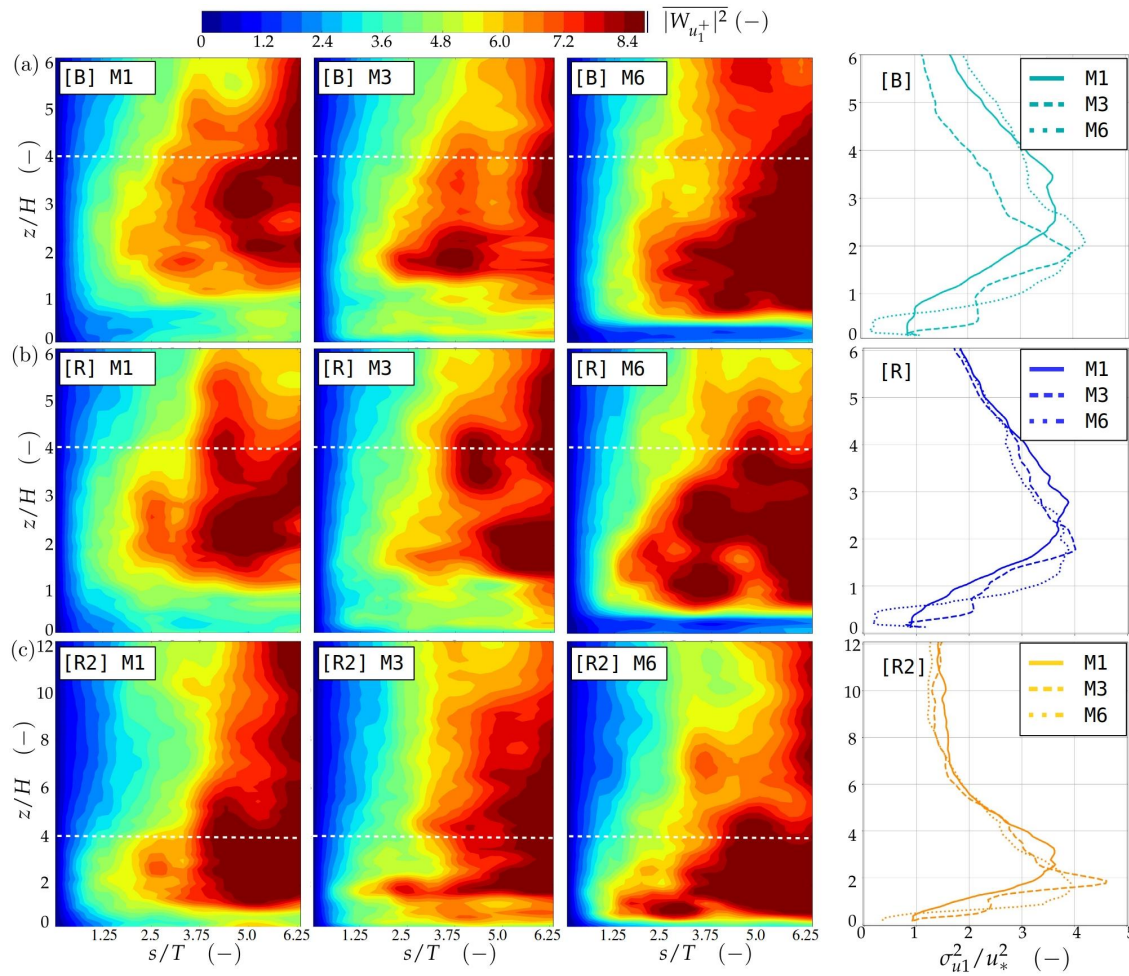
Figure 13. Temporally averaged power scalograms of u_1^+ time series obtained from stations M1, M3, and M6, collected from (**a**) case [B] with the largest ABL eddies removed, (**b**) reference case [R], and (**c**) coarser resolution reference case [R2], where the data is coincidentally sampled from domain $\Omega^{(2)}$. Note the vertical extent of the $\Omega^{(2)}$ results is $\sim 12H$, whereas $\Omega^{(3)}$ results are limited to $\sim 6H$. White dashed line highlights $z/H = 4$ for convenience.

Although the amplification of large-scale modes within the urban roughness sublayer have not been substantiated experimentally, the presence of low frequency ABL structures in real urban canopy flows should be anticipated. For this reason, the importance of conducting time–frequency analysis of urban flow data is strongly emphasized here.

3.4.1. Filtering Out the Largest Atmospheric Boundary Layer Modes

To perform a targeted comparison of turbulence statistics within a frequency/scale range that is representative of urban canopy turbulence and free of the observed VLSMs, the streamwise-oriented velocity time-series are subjected to a sharp Fourier filter. To specify the appropriate cutoff frequency/scale, both $\overline{|W_{u_1^+}|^2}$ distributions and spectral energy density plots are examined. Figure 14a displays an example spectral energy density curve obtained from case [R] and station M6 at two heights $z = 0.5H$ and $z = 3H$. The M6 station is located within a park of dense vegetation. The spectra reveals that the LES-resolved u_1 signal at $z/H = 3$ successfully captures the theoretical $-5/3$ slope energy cascade well over an order of magnitude frequency range. As expected, within the vegetation covered park at $z/H = 0.5$, the enhanced dissipation due to trees is clearly observed.
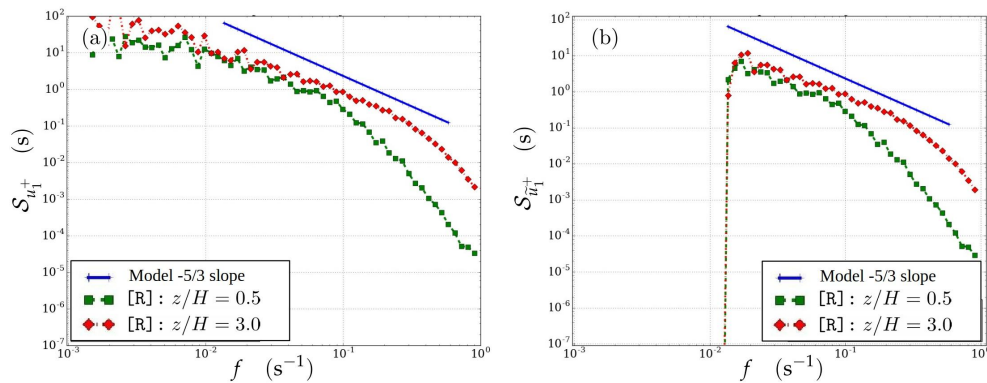
Figure 14. Example spectral energy densities ($\mathcal{S}$) of the streamwise-oriented velocity u_1 signal extracted at two heights from simulation `[R]` and location `M6` which cuts through a park with densely packed trees: (**a**) Spectral energy density of the original velocity signal and (**b**) spectral energy density of the filtered $\widetilde{u}_1$ signal (cutoff frequency $f_c = 0.011\,\mathrm{s}^{-1}$ or scale $s_c = 2.3T$).

A sharp cutoff frequency $f_c = 0.011\,\mathrm{s}^{-1}$ or scale $s_c = 2.3T$ for the Fourier filtering is specified to eliminate the fluctuating modes in which wave lengths ($\lambda_w > \langle \overline{u_1} \rangle \, s_c \approx 9H$) exceed the length of two charactristic city blocks ($\sim 4H/$block) within the domain. Figure 14b illustrates the spectral energy density of the filtered streamwise velocity, labeled $\widetilde{u}_1$. For completeness, the vertical velocity component is also subjected to identical Fourier filtering, thus yielding $\widetilde{w}$.

A selection of newly computed dimensionless variance profiles are plotted in juxtaposition with their original profiles in Figures 15 and 16 for $\widetilde{u}_1$ and $\widetilde{w}$, respectively. The plots for $\sigma^2_{\widetilde{u}1}/u^2_*$ and $\sigma^2_{\widetilde{w}}/u^2_*$ feature tighter ranges on the axes to allow improved differentiation of differences. The removal of $s > 2.3T$ contributions reduced the original variance of u_1 up to 75% (the original peak value $\max(\sigma^2_{u1}) \approx 4u^2_*$ appearing in both `M5` and `M10` profiles approximately at height $z/H = 4$ reduced to $\sigma^2_{\widetilde{u}1} \approx 1u^2_*$ after filtering) whereas the variance of w was reduced by approximately 20%. This is expected given the horizontal degrees of freedom of the LES model.

The influence from the VLSM modes is still indirectly contained in the high-frequency (low scale) fluctuations via amplitude modulation, and the observed scatter in profiles for $z/H > 3$ offers an indication for its magnitude. The filtered variance profiles also feature 95th confidence intervals computed via bootstrapping technique to confirm that the resulting scatter does not arise from poorly converged statistics.
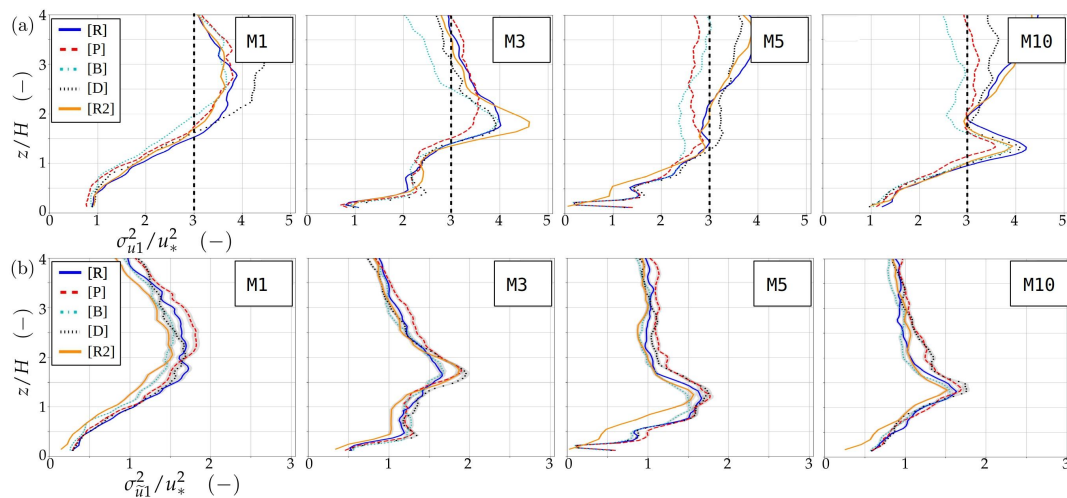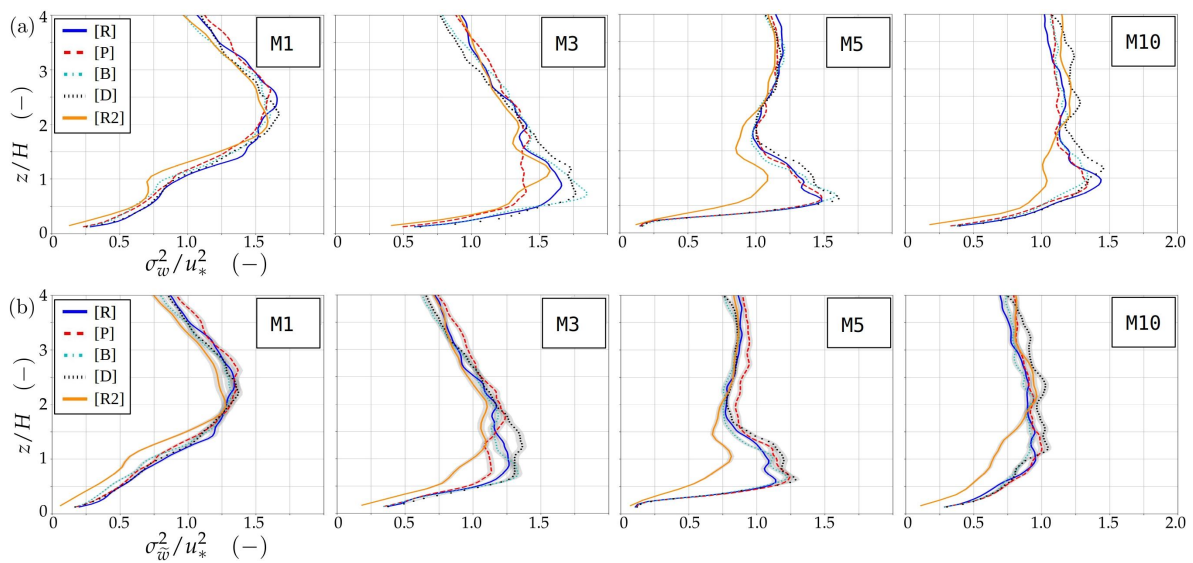

Figure 15. A juxtaposition of original (**a**) and filtered (**b**) dimensionless streamwise-aligned velocity variance profiles at selected measurement locations: The range on (**b**) axes has been modified to allow better differentiation of the differences. Black dashed lines have been added to (**a**) to indicate the variance range of (**b**) graphs.

The dissimilarity between the shown $\sigma^2_{\tilde{u}1}$ profiles at different locations reiterates how the local urban plan influences the local scale turbulence. The different model variations featuring $\Delta = H/20$ resolution ([R], [P], [B], and [D]) bring about only modest changes as the profiles maintain their characteristic location-specific shape for $z/H < 4$. Deviations due to simplified tree LAD profiles ([P]) or systematically reduced vegetation drag coefficient ([D]) at low altitudes are generally minor and difficult to distinguish from the prevailing scatter, but these two cases also exhibit almost identically at stations. Variance of $\tilde{w}$, however, does reveal notable differences at M3 (an open intersection with dense patch of trees next to the measurement site) between the [R] and cases [P] and [D]. The local flow pattern features a swirl (see Figure 8a) in which the exact location in the LES solution is influenced by the $D_v = C_{Dv}LAD$ distribution of the trees.

The close agreement of u_1 and w variance profiles between case [B], with largest ABL eddies removed, and case [R], solidifies the conclusion that the nature of the local scale urban canopy turbulence is dictated by the arrangement of the surrounding structures and not by the conditions at the upper part of the ABL.

The most significant deviations are again introduced by the coarser [R2] solution at low altitudes. This is particularly notable in the $\sigma^2_{\tilde{w}}$ profiles in Figure 16b where the vertical fluctuations are clearly subdued close to the ground and most notably within the vegetation at location M5. The narrow streamwise-oriented street canyon site M1 forms an exception in the shown profiles as all model variations agree on the profile shape that reaches its maximum within range $2 < z/H < 3$ which is notably higher than at the other shown site. Sites M2 and M7 exhibit similar variance profiles.


Figure 16. A juxtaposition of original (**a**) and filtered (**b**) dimensionless vertical velocity variance profiles at selected measurement locations.

3.4.2. Comparing Coherent Structure Content Using Entropy and Divergence

The final approach used to investigate the sensitivity of LES modeled urban canopy turbulence to modeling choices involves a detailed examination of the probability density distributions of coherent local u'_1 fluctuations, conveyed via power scalograms $|W_{u_1^+}|^2(t, s, z)$ evaluated for each urban location. The objective is to compare the energy distribution of turbulent fluctuations occuring at a specific time scale s_i, location $\mathbf{x}_m$, and height z_j. Here, the concepts of information entropy and, more specifically, Shannon entropy (see Equations (19) and (21)) and the associated divergence (Equations (22) and (23)) are exploited to perform a quantitative comparison. Information entropy provides a measure for how broadly the magnitudes of the coherent turbulent fluctuations are distributed, as quantified by a discrete probability density function $\rho^* \equiv \rho^*(|W_{u_1^+}|^2)$. The information divergence, on the other hand, provides a measure of how two probability density functions or sample frequency distributions

$f^* = \rho^* \Delta^*$ differ from each other. These entropy and divergence evaluations will be carried out over a time scale range $0.025 < s/T < 1.5$ considered most relevant in this study and over an elevation range $0 < z/H < 4.5$, which encompasses the essential part of the urban RSL.

Figure 17 demonstrates the discrete probability density distributions of power scalogram coefficients $\rho^* \equiv$ featuring selected case pairs, time scales, and elevations. The Shannon entropy (see Equation (21)) of the reference case is denoted as $\mathcal{H}_1(f_A^*)$ and the modified case is denoted as $\mathcal{H}_1(f_B^*)$. The divergence between the two sample frequency distributions is denoted as $\mathcal{D}_1(f_A^*, f_B^*)$.
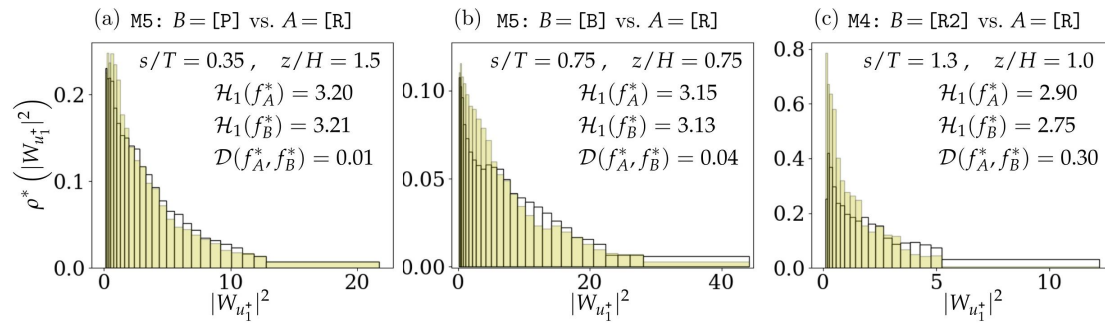
Figure 17. Example discrete sample frequency density distributions $\rho^*(\bullet) = f^*(\bullet)/\Delta^*$ used to construct the entropy and divergence distributions on (s, z) plane as presented in Figure 18: White (colorless) bars are from [R] reference simulation, while the yellow bars are from one of the modified cases. (**a**) Nearly equivalent histograms, both resulting in relatively high entropy values while yielding very low divergence. (**b**) Moderately deviating histrograms which give rise to relatively low divergence. (**c**) Clearly deviating histograms that result in very high divergence. Note the merging of bins at the tail end of the distributions.

The generation of the discrete sample frequencies f^* for Equations (21) and (23) must be performed with care because the evaluation of divergence requires that both discrete probability distributions, f_A^* and f_B^*, feature exactly the same bins. Therefore, the construction of histograms of the reference case $A = $[R] and the selected modified case B at specific time scale and elevation must be carried out simultaneously such that the bin distributions can be made suitable for both. This is achieved by first establishing the common min/max bounds for both histograms and then by redefining a quadratic distribution of bin widths from min to max using M initial bins. In this work, $M = 40$ was determined to provide a sufficiently resolved initial distribution of bins. To ensure each bin represents converged information, every bin Δ_k, where the number of samples $n_k < 50$ in either of the histograms, is merged with Δ_{k-1}. This process is repeated until the bin distribution has converged. The final histograms typically feature between 20 and 34 bins. In Figure 17, all histograms contain multiple merged bins at the tail end of the distributions.

Figure 18 presents the Shannon entropy distributions $\mathcal{H}_1(f^*) \equiv \mathcal{H}_1(f^*(|W_{u_1^+}|^2))$ for the reference case [R] (signified by subscript A) in juxtaposition with all the modified cases [P], [B], [D], and [R2] (subscript B) together with the corresponding divergence distributions $\mathcal{D}_1(f_A^*, f_B^*)$. The individual distributions are all labeled at the top-left corner, and the divergence plots are placed at the B vs. A cross sections. A complete assembly of results is included from stations M4, M5, and M6.

The obtained $\mathcal{H}_1(s, z)$ distributions embody the complexity of the simulated RSL turbulence: within a given monitoring station, the prevailing broader patterns depict distinct similarities between all $\Delta = H/20$ resolved cases [R], [P], [B], and [D] whereas the comparison between different monitoring stations reveal systematic differences. Some fine-detail differences particularly at larger scales ($s/T > 1$), that are also observable as non-systematic spots in the divergence distributions, are assumed evidence for the intermittent coherence that stems from the slowly evolving LSM and VLSM modes, which had their unique realizations in each simulation (except for [R] and [R2]). However, in this context, it is not possible to determine the root cause of the sporadic divergence appearances occuring at larger time scales ($s/T > 1$) or higher elevations ($z/H > 2$)

because the footprint of the flow phenomena extends far upstream. Nonetheless, it is informative to observe the magnitude range of these disordered divergence instances in relation to where systematic divergence appears.

In general, the divergence between the reference and the other $\Delta = H/20$ resolution simulations remained systematically very low for $s/T < 1$ throughout the RSL. A noteworthy exception occurs in Figure 18c M6: [P] vs. [R], where, below the vegetation, the coherent fluctuation distributions from [P] yielded consistently lower entropy which manifests as a narrow band of elevated divergence at $z/H = 0.3$ (between the ground and the lower part of the foliage). The top-heavy LAD profile in [R] allowed stronger fluctuations to penetrate into the lower elevations. This indicates that the significance of LAD-profile accuracy becomes relevant if the near-ground wind conditions in vegetated areas are of specific interest. Relevant application examples include the exchange processes with the top layer of soil and the dispersion of traffic-related pollutants on urban esplanades and boulevards. Note that the similar difference between [P] vs. [R] is not observable at the M5 station which is close to the upwind edge of the park.
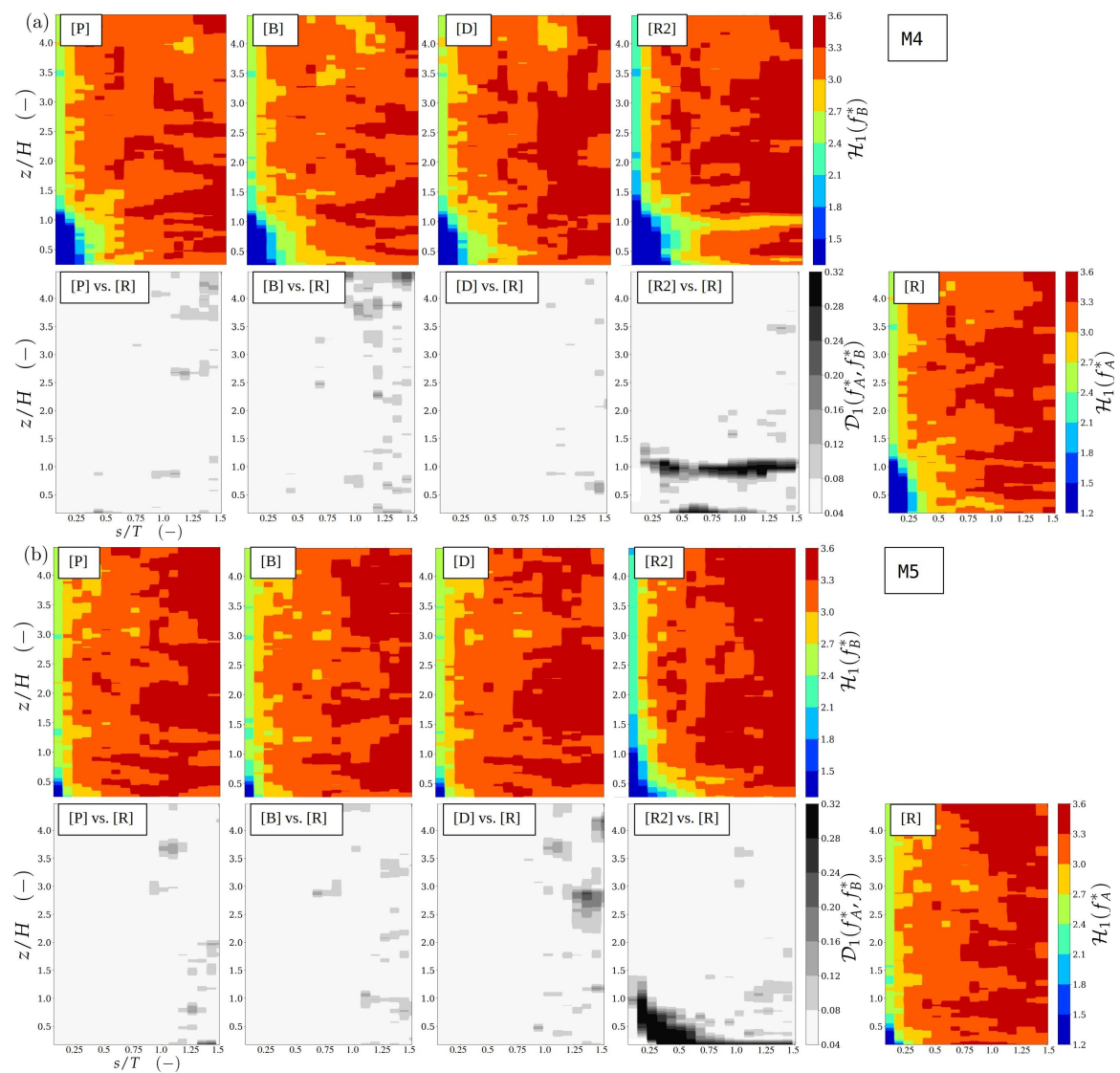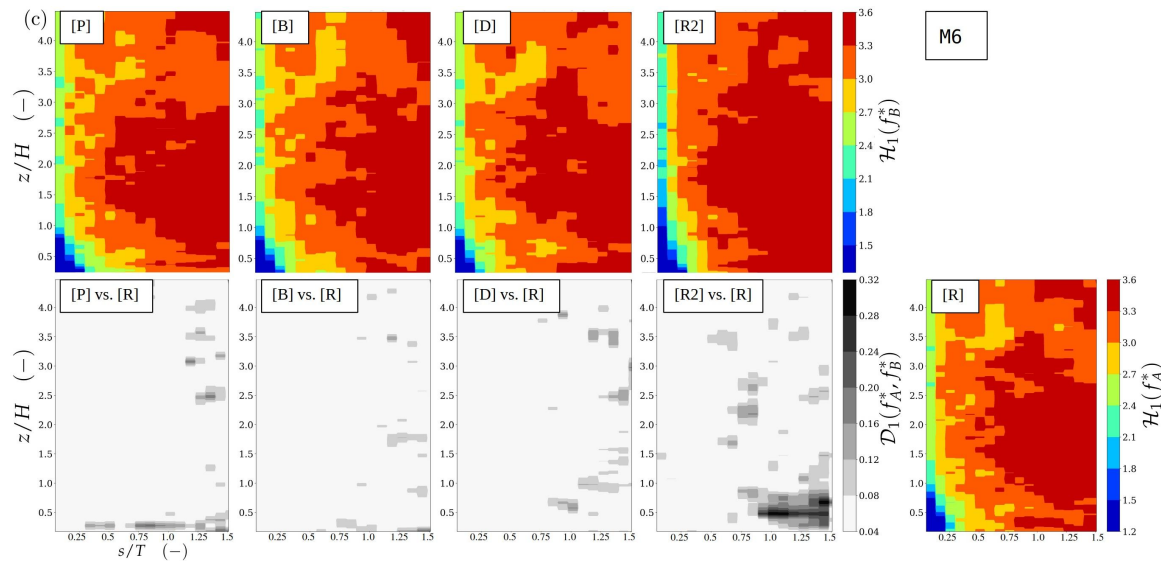

Figure 18. *Cont.*

Figure 18. Distributions of Shannon entropy $\mathcal{H}_1$ and divergence $\mathcal{D}_1$ on (s,z)-plane evaluated from power scalograms $|W_{u_1^+}|^2$ obtained from selected measurement stations (**a**) M4, (**b**) M5, and (**c**) M6. In each subfigure, the $\mathcal{H}_1$ distribution for the reference case [R] is shown on the right, while the distributions for the modified [P], [B], [D], and [R2] cases are aligned on the top row. Each Shannon entropy value is computed from discrete sample frequency distributions f^* of which subscript A signifies [R] and B one of the modified cases. The divergence distributions $\mathcal{D}_1(f_A^*, f_B^*)$ are positioned at the B vs. A cross sections.

Comparing the divergence distributions between the different $\Delta = H/20$ resolved cases against the divergence between [R] and [R2] puts their magnitudes into perspective. Within the urban canopy, in close proximity of buildings and vegetation, the effect of reducing resolution from $\Delta = H/20$ to $\Delta = H/10$ outweighs all the effects that arise due to the considered LOD modifications. In Figure 18a M4: [R2] vs. [R], the region around the roof height, where the mean wind flows perpendicular to the narrow vegetated street canyon resulting in a skimming flow with high local vertical gradient in velocity, the resolved turbulent fluctuations differ considerably. Although the case [R2] yielded highly comparable mean (Figure 9) and variance (Figure 11) profiles for the streamwise velocity, the lower resolution results in notably lower information entropy in the resolved turbulence than in the [R] solution. As Figure 17c illustrates, the rare and stronger bursts of coherent turbulence are far less frequent in the coarser [R2] than in the [R] solution. This is a general feature that separates the turbulence in [R2] from [R]: in the vicinity of the urban roughness elements, the coarse solution fails to resolve the stronger bursts of coherent fluctuations. For instance, in the scale range $0.45 < s/T < 1.2$, the near-ground turbulence (between the ground and the foliage) displays marked levels of divergence for the same reason.

Moreover, the resolution-induced divergence also dominates in the results for M5 and M6 stations. However, there is a constrast in how the turbulence solutions differ: at station M5, which is near the upwind edge of the park, the divergence is systematically elevated within the foliage at low scales ($s/T < 0.75$), whereas at M6, which is further downstream, the divergence is accumulated at higher time scales ($s/T > 1$). At M5, in addition to the infrequency of strong bursts in [R2] solution, the small amplitude coherent fluctuations of the high-resolution [R] solution adapts to the vegetation drag faster, leading to a significantly reduced peak at the left end of the distribution as well. At station M6, the wind is significantly subdued throughout the depth of the vegetation canopy and the two solutions principally differ because the coarser solution contains a reduced count of weak large-scale fluctuations.

4. Conclusions

In this comparative study, a realistic urban boundary layer flow scenario is numerically investigated with a high-resolution LES model. The scenario involves neutral ABL flow developed over sea surface that becomes incident with a coastal city where topographic description within the LES model is obtained from a high-resolution LiDAR dataset. The simulations are repeated with modified LES models to investigate the sensitivity of the urban flow system to modelling choices that carry uncertainties. This work employs a number of techniques in quantifying the magnitude and nature of the deviations that stem from uncertainties in specifying the aerodynamic properties of urban vegetation and the vertical extent of the ABL. The results and their comparative analysis lead to the following conclusions and subsequent recommendations for urban LES modelling.

The applied LOD modifications to the urban surface or boundary layer depth resulted in minor changes in the log-law similarity scaling parameters computed from double-averaged profiles. The elimination of the largest ABL eddies (case [B]) had the most observable impact, reducing friction velocity u_* by 6%, increasing roughness length z_0 by 6%, and reducing displacement height z_d by 3%.

The vertical distributions of $RNMSD(\overline{u}_1^+)$, the root-normalized-mean-square differences (with respect to the reference case) for streamwise velocity evaluated over horizontal planes, obtain their highest values near the ground but diminish rapidly with elevation. Case [P], featuring simplified LAD-profiles of trees, yielded marginally higher $RNMSD(\overline{u}_1^+)$ values than other model modifications, but even its maximum remained at only 10% of the recommended maximum level desired when model results are validated against experimental measurements. Thus, the model uncertainties leave a significant margin for successful experimental model validation. On the other hand, the conditional $RNMSD(\overline{w}^+)$ values for vertical velocity (evaluated using points where $\overline{w}^+ > 0$ in the reference case) remain consistently around 10% from the recommended maximum throughout the canopy height again case [P] recording the highest discrepancy levels close the ground. Fractional bias (FB) values for both velocity components remained well below 10% of the recommended limit throughout the canopy and above, indicating that the model modifications generate very low levels of systematic deviations.

Overall, the ability to quantify differences in urban canopy flows becomes difficult close to the level of pedestrians and vehicles, particularly in areas where the wind speeds are reduced due to the density of the surrounding urban plan and/or drag caused by dense vegetation. Predictions concerning $\overline{w}$ within a complex urban canopy are endowed with greater uncertainty as the network of street canyons and blunt obstacles gives rise to a complex distribution of vortices and wakes. This is in contrast with the horizontally streamwise-aligned velocity, which carries most of the momentum and of which the distribution is largely dictated by energetic effects.

The $\overline{u}_1^+$ profiles obtained at 11 different virtual tower positions depict very high degrees of adherence between the cases, demonstrating that the street-canyon flow within a city section having a high plan and frontal area fraction of buildings is, nonetheless, primarily dictated by the solid obstacles. On the other hand, the profiles obtained from the reference case but extracted from the $\Delta = H/10$ resolved nested domain reveal the pronounced sensitivity of the solution to grid resolution. The deviations arising due to coarser grid resolution systematically outweighed the deviations due to imposed model modifications. This outcome was particularly pronounced in the comparison of local $^+$ profiles although [P] (with simplified LAD-profiles) also showed deviations where vegetation caused changes in mean vertical velocity.

Comparison of local variance profiles is complicated by the presence of the slowly evolving very-large-scale ABL modes which interact with the urban roughness and comprise up to 75% of the observed variance in the streamwise velocity component. The use of sharp Fourier filter to eliminate the direct influence of the VLSM turbulence facilitated a meaningful comparison where, again, the case with reduced spatial resolution yielded the most notable differences within the urban canopy. Case [P] also exhibited deviations in variance profiles at locations where vegetation is in the immediate vicinity. The effects in local variances caused by reducing the vegetation drag or eliminating the largest ABL eddies were minor.

The energy distribution of the turbulent fluctuations at different frequencies are examined and compared by the means of wavelet analysis exploiting the concept of information entropy to provide a measure for the diversity of coherent fluctuation amplitudes within the signal. A measure of the difference between two signals' information entropy content, in turn, is given by information divergence. The divergence distributions, where the reference signal is compared against the modified cases over different frequency and elevation values, reveal that, within the canopy in areas where strong shear layers exists or where the flow adjusts to the surrounding vegetation, the change in grid resolution clearly leads to the highest divergence values. Changing the *LAD* profiles also results in systematic appearances of divergence within parks below the tree crowns. However, these changes in turbulence were weaker than the changes caused by reduced grid resolution.

Recommendations

- Nested domain approach is essential in realistic UBL flows due to the size requirement for the computational domain.
- Sufficient grid resolution is critical when urban canopy flow is of primary interest. Grid spacing of $\Delta = H/20$ or better is recommended in the vicinity of the urban roughness elements and close to the ground.
- The accuracy of the *LAD* profile is relevant when detailed information is sought close to the ground in proximity of trees.
- The shape of vegetation should be modelled as accurately as possible, but the flow system within a densely arranged urban environment is not very sensitive to the precision of the constant drag coefficient C_{Dv} in the vegetation drag model.
- The grid resolution on the top part of the domain can be relaxed as the largest eddies do not influence the urban canopy turbulence to an observable degree. However, it is beneficial to employ sufficiently tall computational domains, with $L_z \approx 20H$, particularly when the transition from fully developed ABL flow to a developing UBL flow is resolved.

Author Contributions: Conceptualization, M.A., S.B., A.H., and L.J.; methodology, M.A., S.B., and A.H.; software, M.A., S.B., and A.H.; formal analysis, M.A. and S.B.; investigation, M.A. and S.B.; resources, M.A. and T.T.; data curation, M.A. and T.T.; writing—original draft preparation, M.A. and S.B.; writing—review and editing, L.J.; visualization, M.A.; supervision, A.H. and L.J.; project administration, M.A., A.H., and L.J.; funding acquisition, M.A., A.H., and L.J. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ABL	Atmospheric boundary layer
DA	Double average
ISL	Inertial sublayer
LAD	Leaf area density
LAI	Leaf area index
LES	Large-Eddy Simulation
LSM	Large-scale motions
RSL	Roughness sublayer
UBL	Urban boundary layer
VLSM	Very-large-scale motions

Appendix A. Fourier Filtering: Continuous and Discrete Version

Given the turbulent fluctuations $\mathbf{u}'(t, \mathbf{x})$, where time is ideally a continous variable, the Fourier transform and antitransform read as follows:

$$\widehat{\mathbf{u}'}(f, \mathbf{x}) \quad = \quad \int_{\mathbb{R}} dt\, \mathbf{u}'(t, \mathbf{x})\, e^{i2\pi ft} \tag{A1}$$

$$\mathbf{u}'(t, \mathbf{x}) \quad = \quad \int_{\mathbb{R}} df\, \widehat{\mathbf{u}'}(f, \mathbf{x})\, e^{-i2\pi ft} \tag{A2}$$

Since our time series always have constant time steps (i.e., $t_k = k\,\Delta t$), the convention to discretize the Fourier transform here will be $\Delta f = 1/(N\Delta t)$, $f_n = n\,\Delta f$, and the following:

$$\widehat{\mathbf{u}'}(f_n, \mathbf{x}_i) \quad = \quad \sum_{k=0}^{N-1} \Delta t\, \mathbf{u}'(t_k, \mathbf{x}_i)\, e^{i\frac{2\pi kn}{N}} \tag{A3}$$

$$\mathbf{u}'(t_n, \mathbf{x}_i) \quad = \quad \frac{1}{N\Delta t} \sum_{k=0}^{N-1} \widehat{\mathbf{u}'}(f_k, \mathbf{x})\, e^{-i\frac{2\pi kn}{N}} = \frac{1}{N\Delta t} \sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{\lfloor N/2 \rfloor} \widehat{\mathbf{u}'}(f_k, \mathbf{x})\, e^{-i\frac{2\pi kn}{N}} \tag{A4}$$

The second equality in Equation (A4) is justified because, from Equation (A3), it turns out that $\widehat{\mathbf{u}'}(f_n, \mathbf{x}_i)$ is N-periodic with respect to the index n. It is thus convenient to consider the period of the frequency series between indexes $-\lfloor(N-1)/2\rfloor$ and $\lfloor N/2 \rfloor$, with $\lfloor \bullet \rfloor$ being the floor function. Such a convention renders it easier to associate the the continuous Fourier transform to its discrete version. All the information about the time series is contained in the positive half of the frequency indexes $0, \ldots, \lfloor N/2 \rfloor$, as the fluctuations are a real function.

Now, let us suppose we recognize a cutoff frequency f_c in the fluctuations $\mathbf{u}'(t, \mathbf{x})$ under which the flow can be considered small-scale turbulence and above which the flow consists of slow modes only, that is large-scale fluctuations. This split can be done via a sharp Fourier filter:

$$\widetilde{\mathbf{u}'}(t, \mathbf{x}) \quad = \quad \int_{\substack{[-\infty\,,\,-f_c] \\ \cup\,[f_c\,,\,\infty]}} df\, \widehat{\mathbf{u}'}(f, \mathbf{x})\, e^{-i2\pi ft} \tag{A5}$$

$$\mathbf{u}'_{LM}(t, \mathbf{x}) \quad = \quad \int_{[-f_c\,,\,f_c]} df\, \widehat{\mathbf{u}'}(f, \mathbf{x})\, e^{-i2\pi ft} \tag{A6}$$

As a consequence of the reality of the fluctuation function, we also have that $\widehat{\mathbf{u}'}(f, \mathbf{x}) = \widehat{\mathbf{u}'}^{\star}(-f, \mathbf{x})$, the symbol * standing for complex conjugation. The small-scale fluctuations are represented by $\widetilde{\mathbf{u}'}(t, \mathbf{x})$, whereas $\mathbf{u}'_{LM}(t, \mathbf{x})$ constitutes the large-mode fluctuations. It is the former subset that is analysed in the present work. However, to do so, it is expedient to consider the small-scale fluctuations plus the average local offset:

$$\widetilde{\mathbf{u}}(t, \mathbf{x}) = \widetilde{\mathbf{u}'}(t, \mathbf{x}) + \overline{\mathbf{u}}(\mathbf{x}_i) \tag{A7}$$

In practice, our time series will be the velocity fluctuations plus the offset computed on a time sequence $t_1, \ldots, t_N$, and as a result of this, what will be implemented is a discrete combination of Equations (A5)–(A7):

$$\widetilde{\mathbf{u}}(t_n, \mathbf{x}_i) \quad = \quad \overline{\mathbf{u}}(\mathbf{x}_i) + \frac{1}{N\Delta t} \left[\sum_{k=-\lfloor\frac{N-1}{2}\rfloor}^{-n_c} \widehat{\mathbf{u}'}(f_k, \mathbf{x}_i)\, e^{-i\frac{2\pi kn}{N}} + \sum_{k=n_c}^{\lfloor N/2 \rfloor} \widehat{\mathbf{u}'}(f_k, \mathbf{x}_i)\, e^{-i\frac{2\pi kn}{N}} \right] \tag{A8}$$

having denoted by n_c the index such that $n_c \Delta f = n_c/(N\Delta t) = f_c$.

Appendix B. Wavelet Transform, Morlet Family, and Its Approximated Form

In this Appendix, we intend to illustrate more precisely and to justify the assumptions and choices which have been made about the wavelet analysis in the main text.

Let $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ be the spaces of functions over the real numbers of which the absolute value and squared absolute value are integrable, respectively. The continuous wavelet transform (CWT) of a function $\varphi(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is defined as follows [75]:

$$W_\varphi(s,t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} d\bar{t}\, \varphi(\bar{t}) \Psi\left(\frac{\bar{t}-t}{s}\right) \tag{A9}$$

The graph of the function $W_\varphi(s,t)$ on the plane (s,t) is called *scalogram*. The function $\Psi(\bar{t})$, which is complex in general, is called mother wavelet and generalizes the complex oscillating exponential functions in the Fourier transform. The mother wavelet generates a wavelet family by means of the dilation factor s and the time instance t in the sense that the analysis is carried out at time t of the function φ. There exist several mother wavelets, any of which must satisfy the so-called admissibility criteria, i.e., they must be square integrable, and the following quantity must be well defined:

$$C_\psi = \int_0^\infty \frac{\left|\widehat{\Psi}(\tau)\right|^2}{|\tau|}\, d\tau$$

An immediate consequence from the convergence of this integral is that the time average of the mother wavelet $\widehat{\Psi}(0)$ must be zero.

The wavelet antitransform is as follows [75]:

$$\varphi(t) = \frac{1}{C_\psi} \int_0^\infty \frac{ds}{s^2} \int_{-\infty}^{\infty} W_\varphi(s,\bar{t}) \frac{1}{\sqrt{s}} \Psi\left(\frac{\bar{t}-t}{s}\right) d\bar{t} \tag{A10}$$

From Equation (A10), it can be promptly noticed that there are two continuous and unbounded parameters in the wavelet antitransform to integrate over. This fact means the wavelet family generated from the mother is a redundant basis. Similarly to any other basis for a square-integrable functions space, a numerable subset can indeed be extracted from the two continuous parameters $s, \bar{t}$ without losing the invertibility of the transform [76].

The mother wavelet we will be considering here is the complex *Morlet wavelet*:

$$\Psi_{\omega_{0\star}}(t_\star) = c_{\omega_{0\star}} \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t_\star^2}\left(e^{i\omega_{0\star}t_\star} - \kappa_{\omega_{0\star}}\right) \tag{A11}$$

where $t_\star$ is some rescaled, dimensionless time and $\kappa_{\omega_{0\star}} = e^{-\frac{1}{2}\omega_{0\star}{}^2}$ is defined through the null-average condition, and the normalization constant $c_{\omega_{0\star}}$ comes from setting the integral of the square modulus equal to 1:

$$c_{\omega_{0\star}} = \left(1 + e^{-\omega_{0\star}{}^2} - 2e^{-\frac{3}{4}\omega_{0\star}{}^2}\right)^{-\frac{1}{2}} \tag{A12}$$

It is also relevant to consider the Fourier transform of the Morlet wavelet with respect to the corresponding, dimensionless frequency $f_\star$ (A11):

$$\widehat{\Psi}_{\omega_{0\star}}(f_\star) = c_{\omega_{0\star}} \pi^{-\frac{1}{4}} \left(e^{-\frac{1}{2}(\omega_{0\star}-2\pi f_\star)^2} - \kappa_{\omega_{0\star}} e^{-\frac{1}{2}(2\pi f_\star)^2}\right) \tag{A13}$$

The dimesionless parameter $\omega_{0\star}$, which in turn generates a *family* of mother wavelets, can be freely chosen depending on the case considered. It is usually chosen as an integer, as it corresponds to the number of pseudo-periods within the two-sided Gaussian width $2s$. Here, we will set it to 6 for the

following reasons. First, we recall that it is common practice to consider the following approximation whenever $\omega_{0\star} \geq 5$ [22]:

$$c_{\omega_{0\star}} \simeq 1 \qquad \& \qquad \kappa_{\omega_{0\star}} \simeq 0$$

As for the Fourier transform, we get a Gaussian centered in $\omega_{0\star}/(2\pi)$:

$$\hat{\Psi}_{\omega_{0\star}}(f_\star) \simeq \pi^{-\frac{1}{4}} e^{-\frac{1}{2}(\omega_{0\star} - 2\pi f_\star)^2} \tag{A14}$$

We can therefore notice from Equation (A11) that the Morlet wavelet can be well approximated as follows:

$$\Psi_{\omega_{0\star}}(t_\star) \simeq \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t_\star^2} e^{i\omega_{0\star}t_\star} \quad , \tag{A15}$$

We can also interpret this wavelet as the atom of the joint time–frequency analysis associated to the dimensionless frequency $\frac{\omega_{0\star}}{2\pi}$ at time 0. By virtue of Equation (A11), plugging the rescaled time with respect to the time scale s into the wavelet function $\Psi\left(\frac{\bar{t}-t}{s}\right)/\sqrt{s}$ involved in Equation (A9) reveals the spirit of time–frequency joint transform, as its envelop is a Gaussian function centered around the time t. The generated wavelets have the following Fourier transform with respect to $\bar{t}$:

$$\sqrt{s}\,\hat{\Psi}_{\omega_{0\star}}(sf) \simeq \pi^{-\frac{1}{4}}\sqrt{s}\, e^{i2\pi f t} e^{-\frac{[f-\omega_{0\star}/(2\pi s)]^2}{2/(2\pi s)^2}} \tag{A16}$$

We can notice the oscillating complex factor due to the time translation t. Also, the Gaussian is centered around the frequency $\omega_{0\star}/(2\pi s)$, with its variance widening by a factor $1/s^2$ as a consequence of the uncertainty principle [75]. Each positive value of the rescaling time s biunivocally corresponds to a wavelet ordinary frequency $f = \omega_{0\star}/(2\pi s) \sim 1/s$ because of the choice $\omega_{0\star} = 6$. The typical wavelength (or pseudoperiod) of the wavelet consequentially is simply $\sim s$. This biunivocal correspondence allows defining the *spectrogram* as the graph of the wavelet transform in the plane (f, t) instead of (s, t).

Finally, we need to establish how Equations (A9) and (A15) will be discretized, taking into account finite size effect and aliasing. With $\Delta t = 0.5$ s being the time step in the time series $\mathbf{u}(t_n, \mathbf{x}_i)$, we already know that time scales close to $\simeq 1$ s are affected by aliasing [81]. Wavelet resolution also imposes constraints on the smallest time scale that can be analyzed. From Equation (A16), the wavelet bandwidth can be identified as $[\omega_{0\star}/s + 3/s]/(2\pi)$. Given that the sample frequency is $f_s = 2$ Hz, the fulfillment of the following constraint would be sufficient in order to avoid aliasing [81]:

$$2\left(\frac{\omega_{0\star}}{2\pi s} + \frac{3}{2\pi s}\right) < f_s$$

As a result of this, only time scales sufficiently larger than $(\omega_{0\star} + 3)/(2\pi) \simeq 1.4$ s can be well resolved in the wavelet transform, as stated in the main text.

References

1. Britter, R.E.; Hanna, S.R. Flow and Dispersion in Urban Areas. *Annu. Rev. Fluid Mech.* **2003**, *35*, 469–496. [CrossRef]
2. Tseng, Y.H.; Meneveau, C.; Parlange, M.B. Modeling flow around bluff bodies and predicting urban dispersion using large eddy simulation. *Environ. Sci. Technol.* **2006**, *40*, 2653–2662. [CrossRef] [PubMed]
3. Bou-Zeid, E.; Overney, J.; Rogers, B.D.; Parlange, M.B. The Effects of Building Representation and Clustering in Large-Eddy Simulations of Flows in Urban Canopies. *Bound.-Layer Meteorol.* **2009**, *132*, 415–436. [CrossRef]
4. Xie, Z.T.; Castro, I.P. Large-eddy simulation for flow and dispersion in urban streets. *Atmos. Environ.* **2009**, *43*, 2174–2185. [CrossRef]

5.　Giometto, M.; Christen, A.; Meneveau, C.; Fang, J.; Krafczyk, M.; Parlange, M. Spatial Characteristics of Roughness Sublayer Mean Flow and Turbulence Over a Realistic Urban Surface. *Bound.-Layer Meteorol.* **2016**, *160*, 1–28. [CrossRef]

6.　Santiago, J.L.; Martilli, A. A Dynamic Urban Canopy Parameterization for Mesoscale Models Based on Computational Fluid Dynamics Reynolds-Averaged Navier–Stokes Microscale Simulations. *Bound.-Layer Meteorol.* **2010**, *137*, 417–439. [CrossRef]

7.　Krayenhoff, E.; Santiago, J.L.; Martilli, A.; Christen, A.; Oke, T. Parametrization of drag and turbulence for urban neighbourhoods with trees. *Bound.-Layer Meteorol.* **2015**, *156*, 157–189. [CrossRef]

8.　Tominaga, Y.; Stathopoulos, T. CFD simulation of near-field pollutant dispersion in the urban environment: A review of current modeling techniques. *Atmos. Environ.* **2013**, *79*, 716–730. [CrossRef]

9.　Buccolieri, R.; Hang, J. Recent Advances in Urban Ventilation Assessment and Flow Modelling. *Atmosphere* **2019**, *10*, 144. [CrossRef]

10.　Cheng, H.; Castro, I.P. Near Wall Flow over Urban-like Roughness. *Bound.-Layer Meteorol.* **2002**, *104*, 229–259. [CrossRef]

11.　Kanda, M.; Moriwaki, R.; Kasamatsu, F. Large-Eddy Simulation of Turbulent Organized Structures within and above Explicitly Resolved Cube Arrays. *Bound.-Layer Meteorol.* **2004**, *112*, 343–368. [CrossRef]

12.　Coceal, O.; Thomas, T.; Castro, I.; Belcher, S. Mean flow and turbulence statistics over groups of urban-like cubical obstacles. *Bound.-Layer Meteorol.* **2006**, *121*, 491–519. [CrossRef]

13.　Coceal, O.; Thomas, T.; Belcher, S. Spatial variability of flow statistics within regular building arrays. *Bound.-Layer Meteorol.* **2007**, *125*, 537–552. [CrossRef]

14.　Letzel, M. High Resolution Large-Eddy Simulation of Turbulent Flow around Buildings. Inst. für Meteorologie und Klimatologie, 2007. Available online: www.muk.uni-hannover.de/institut/dissertationen.htm (accessed on 11 February 2020).

15.　Xie, Z.; Castro, I. LES and RANS for turbulent flows over arrays of wall-mounted obstacles. *Flow Turbul. Combust.* **2006**, *76*, 291–312. [CrossRef]

16.　Xie, Z.T.; Castro, I.P. Efficient generation of inflow conditions for large eddy simulation of street-scale flows. *Flow Turbul. Combust.* **2008**, *81*, 449–470. [CrossRef]

17.　Letzel, M.; Krane, M.; Raasch, S. High resolution urban large-eddy simulation studies from street canyon to neighbourhood scale. *Atmos. Environ.* **2008**, *42*, 8770–8784. [CrossRef]

18.　Anderson, W.; Li, Q.; Bou-Zeid, E. Numerical simulation of flow over urban-like topographies and evaluation of turbulence temporal attributes. *J. Turbul.* **2015**, *16*, 809–831. [CrossRef]

19.　Castro, I.P.; Xie, Z.T.; Fuka, V.; Robins, A.G.; Carpentieri, M.; Hayden, P.; Hertwig, D.; Coceal, O. Measurements and Computations of Flow in an Urban Street System. *Bound.-Layer Meteorol.* **2017**, *162*, 207–230. [CrossRef]

20.　Giometto, M.; Christen, A.; Egli, P.; Schmid, M.; Tooke, R.; Coops, N.; Parlange, M. Effects of trees on mean wind, turbulence and momentum exchange within and above a real urban environment. *Adv. Water Resour.* **2017**, *106*, 154–168. [CrossRef]

21.　Hertwig, D.; Patnaik, G.; Leitl, B. LES validation of urban flow, part I: Flow statistics and frequency distributions. *Environ. Fluid Mech.* **2017**, *17*, 521–550. [CrossRef]

22.　Hertwig, D.; Patnaik, G.; Leitl, B. LES validation of urban flow, part II: Eddy statistics and flow structures. *Environ. Fluid Mech.* **2017**, *17*, 551–578. [CrossRef]

23.　Maronga, B.; Gryschka, M.; Heinze, R.; Hoffmann, F.; Kanani-Sühring, F.; Keck, M.; Ketelsen, K.; Letzel, M.O.; Sühring, M.; Raasch, S. The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: Model formulation, recent developments, and future perspectives. *Geosci. Model Dev.* **2015**, *8*, 2515–2551. [CrossRef]

24.　Hellsten, A.; Ketelsen, K.; Barmpas, F.; Tsegas, G.; Moussiopoulos, N.; Raasch, S. Nested Multi-scale System in the PALM Large-Eddy Simulation Model. In *Air Pollution Modeling and Its Application XXV*; Mensink, C., Kallos, G., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 287–292.

25.　Wood, C.R.; Järvi, L.; Kouznetsov, R.D.; Nordbo, A.; Joffre, S.; Drebs, A.; Vihma, T.; Hirsikko, A.; Suomi, I.; Fortelius, C.; et al. An Overview of the Urban Boundary Layer Atmosphere Network in Helsinki. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1675–1690. [CrossRef]

26. Hutchins, N.; Chauhan, K.; Marusic, I.; Monty, J.; Klewicki, J. Towards reconciling the large-scale structure of turbulent boundary layers in the atmosphere and laboratory. *Bound.-Layer Meteorol.* **2012**, *145*, 273–306. [CrossRef]

27. Fang, J.; Porté-Agel, F. Large-Eddy Simulation of Very-Large-Scale Motions in the Neutrally Stratified Atmospheric Boundary Layer. *Bound.-Layer Meteorol.* **2015**, *155*, 397–416. [CrossRef]

28. Whitaker, S. *The Method of Volume Averaging*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 13.

29. Schmid, M.F.; Lawrence, G.A.; Parlange, M.B.; Giometto, M.G. Volume Averaging for Urban Canopies. *Bound.-Layer Meteorol.* **2019**, *173*, 349–372. [CrossRef]

30. Wicker, L.J.; Skamarock, W.C. Time-Splitting Methods for Elastic Models Using Forward Time Schemes. *Mon. Weather. Rev.* **2002**, *130*, 2088–2097. [CrossRef]

31. Deardorff, J. Stratoculumus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteorol.* **1980**, *18*, 495–527. [CrossRef]

32. Moeng, C.; Wyngaard, J. Spectral analysis of large-eddy simulations of the convective boundary layer. *J. Atmos. Sci.* **1988**, *45*, 3573–3587. [CrossRef]

33. Saiki, E.; Moeng, C.; Sullivan, P. Large-eddy simulation of the stably stratified planetary boundary layer. *Bound.-Layer Meteorol.* **2000**, *95*, 1–30. [CrossRef]

34. Resler, J.; Krč, P.; Belda, M.; Juruš, P.; Benešová, N.; Lopata, J.; Vlček, O.; Damašková, D.; Eben, K.; Derbek, P.; et al. PALM-USM v1.0: A new urban surface model integrated into the PALM large-eddy simulation model. *Geosci. Model Dev.* **2017**, *10*, 3635–3659. [CrossRef]

35. Auvinen, M. RasterH3D: Raster Maps of Helsinki City Processed from Helsinki 3D Open Access LiDAR Point Cloud Dataset. 2019. Available online: https://doi.org/10.5281/zenodo.2538073 (accessed on 13 February 2020).

36. Hussain, M.; Lee, B. A wind tunnel study of the mean pressure forces acting on large groups of low-rise buildings. *J. Wind. Eng. Ind. Aerodyn.* **1980**, *6*, 207–225. [CrossRef]

37. Grimmond, C.S.B.; Oke, T.R. Aerodynamic Properties of Urban Areas Derived from Analysis of Surface Form. *J. Appl. Meteorol.* **1999**, *38*, 1262–1292. [CrossRef]

38. Auvinen, M.; Järvi, L.; Hellsten, A.; Rannik, U.; Vesala, T. Numerical framework for the computation of urban flux footprints employing large-eddy simulation and Lagrangian stochastic modeling. *Geosci. Model Dev.* **2017**, *10*, 4187–4205. [CrossRef]

39. Kim, K.; Adrian, R. Very large-scale motion in the outer layer. *Phys. Fluids* **1999**, *11*, 417–422. [CrossRef]

40. Balakumar, B.; Adrian, R. Large-and very-large-scale motions in channel and boundary-layer flows. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2007**, *365*, 665–681. [CrossRef]

41. Chung, D.; McKeon, B.J. Large-eddy simulation of large-scale structures in long channel flow. *J. Fluid Mech.* **2010**, *661*, 341–364. [CrossRef]

42. Guala, M.; Hommema, S.; Adrian, R. Large-scale and very-large-scale motions in turbulent pipe flow. *J. Fluid Mech.* **2006**, *554*, 521–542. [CrossRef]

43. Hutchins, N.; Marusic, I. Evidence of very long meandering features in the logarithmic region of turbulent boundary layers. *J. Fluid Mech.* **2007**, *579*, 1–28. [CrossRef]

44. Hutchins, N.; Marusic, I. Large-scale influences in near-wall turbulence. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2007**, *365*, 647–664. [CrossRef]

45. Mathis, R.; Hutchins, N.; Marusic, I. Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers. *J. Fluid Mech.* **2009**, *628*, 311–337. [CrossRef]

46. Anderson, W. Amplitude modulation of streamwise velocity fluctuations in the roughness sublayer: Evidence from large-eddy simulations. *J. Fluid Mech.* **2016**, *789*, 567–588. [CrossRef]

47. Fishpool, G.M.; Lardeau, S.; Leschziner, M.A. Persistent Non-Homogeneous Features in Periodic Channel-Flow Simulations. *Flow Turbul. Combust.* **2009**, *83*, 323–342. [CrossRef]

48. Raupach, M.; Thom, A.S. Turbulence in and above plant canopies. *Annu. Rev. Fluid Mech.* **1981**, *13*, 97–129. [CrossRef]

49. Finnigan, J. Turbulence in Plant Canopies. *Annu. Rev. Fluid Mech.* **2000**, *32*, 519–571. [CrossRef]

50. Yue, W.; Parlange, M.B.; Meneveau, C.; Zhu, W.; Van Hout, R.; Katz, J. Large-eddy simulation of plant canopy flows using plant-scale representation. *Bound.-Layer Meteorol.* **2007**, *124*, 183–203. [CrossRef]

51. De Langre, E. Effects of wind on plants. *Annu. Rev. Fluid Mech.* **2008**, *40*, 141–168. [CrossRef]

52. Schlegel, F.; Stiller, J.; Bienert, A.; Maas, H.G.; Queck, R.; Bernhofer, C. Large-eddy simulation of inhomogeneous canopy flows using high resolution terrestrial laser scanning data. *Bound.-Layer Meteorol.* **2012**, *142*, 223–243. [CrossRef]

53. Mayhead, G. Some drag coefficients for british forest trees derived from wind tunnel studies. *Agric. Meteorol.* **1973**, *12*, 123–130. [CrossRef]

54. Rudnicki, M.; Mitchell, S.J.; Novak, M.D. Wind tunnel measurements of crown streamlining and drag relationships for three conifer species. *Can. J. For. Res.* **2004**, *34*, 666–676. [CrossRef]

55. Vollsinger, S.; Mitchell, S.J.; Byrne, K.E.; Novak, M.D.; Rudnicki, M. Wind tunnel measurements of crown streamlining and drag relationships for several hardwood species. *Can. J. For. Res.* **2005**, *35*, 1238–1249. [CrossRef]

56. Raupach, M.R. Drag and drag partition on rough surfaces. *Bound.-Layer Meteorol.* **1992**, *60*, 375–395. [CrossRef]

57. Nepf, H.M. Drag, turbulence, and diffusion in flow through emergent vegetation. *Water Resour. Res.* **1999**, *35*, 479–489. [CrossRef]

58. Gromke, C.; Blocken, B.; Janssen, W.; Merema, B.; van Hooff, T.; Timmermans, H. CFD analysis of transpirational cooling by vegetation: Case study for specific meteorological conditions during a heat wave in Arnhem, Netherlands. *Build. Environ.* **2015**, *83*, 11–26. [CrossRef]

59. Katul, G.; Albertson, J. An Investigation of Higher-Order Closure Models for a Forested Canopy. *Bound.-Layer Meteorol.* **1998**, *89*, 47–74. [CrossRef]

60. Raupach, M.; Woods, N.; Dorr, G.; Leys, J.; Cleugh, H. The entrapment of particles by windbreaks. *Atmos. Environ.* **2001**, *35*, 3373–3383. [CrossRef]

61. Ghasemian, M.; Amini, S.; Princevac, M. The influence of roadside solid and vegetation barriers on near-road air quality. *Atmos. Environ.* **2017**, *170*, 108–117. [CrossRef]

62. Jacobs, A.F.G. Flow around a Line Obstacle. Ph.D. Thesis, Wageningen Agricultural University, Wageningen, The Netherlands, 1983.

63. Grunert, F.; Benndorf, D.; Klingbeil, K. Neuere Ergebnisse zum Aufbau von Schutzpflanzungen. *Beitrage fur die Forstwirtschaft* **1984**, *18*, 108–115.

64. Tanhuanpää, T.; Vastaranta, M.; Kankare, V.; Holopainen, M.; Hyyppä, J.; Hyyppä, H.; Alho, P.; Raisio, J. Mapping of urban roadside trees—A case study in the tree register update process in Helsinki City. *Urban For. Urban Green.* **2014**, *13*, 562–570. [CrossRef]

65. Markkanen, T.; Rannik, Ü.; Marcolla, B.; Cescatti, A.; Vesala, T. Footprints and Fetches for Fluxes over Forest Canopies with Varying Structure and Density. *Bound.-Layer Meteorol.* **2003**, *106*, 437–459. [CrossRef]

66. Zhou, X.; Brandle, J.; Takle, E.; Mize, C. Estimation of the three-dimensional aerodynamic structure of a green ash shelterbelt. *Agric. For. Meteorol.* **2002**, *111*, 93–108. [CrossRef]

67. Weiss, M.; Baret, F.; Smith, G.; Jonckheere, I.; Coppin, P. Review of methods for in situ leaf area index (LAI) determination: Part II. Estimation of LAI, errors and sampling. *Agric. For. Meteorol.* **2004**, *121*, 37–53. [CrossRef]

68. Rotach, M.W.; Vogt, R.; Bernhofer, C.; Batchvarova, E.; Christen, A.; Clappier, A.; Feddersen, B.; Gryning, S.E.; Martucci, G.; Mayer, H.; et al. BUBBLE—An Urban Boundary Layer Meteorology Project. *Theor. Appl. Climatol.* **2005**, *81*, 231–261. [CrossRef]

69. Chang, J.C.; Hanna, S.R. Air quality model performance evaluation. *Meteorol. Atmos. Phys.* **2004**, *87*, 167–196. [CrossRef]

70. Hellsten, A.; Zilitinkevich, S. Role of convective structures and background turbulence in the dry convective boundary layer. *Bound.-Layer Meteorol.* **2013**, *149*, 323–353. [CrossRef]

71. Sanei Tabass, M.; Mohtashami Borzadaran, G.R.; Amini, M. Renyi entropy in continuous case is not the limit of discrete case. *Math. Sci. Appl. E-Notes* **2016**, *4*, 113–117.

72. Bonachela, J.A.; Hinrichsen, H.; Munoz, M.A. Entropy estimates of small data sets. *J. Phys. A Math. Theor.* **2008**, *41*, 202001. [CrossRef]

73. Roulston, M.S. Estimating the errors on measured entropy and mutual information. *Phys. D Nonlinear Phenom.* **1999**, *125*, 285–294. [CrossRef]

74. Frisch, U.; Kolmogorov, A.N. *Turbulence: The Legacy of AN Kolmogorov*; Cambridge University Press: Cambridge, UK, 1995.

75. Gao, R.X.; Yan, R. *Wavelets Theory and Applications for Manufacturing*; Springer: New York, NY, USA, 2011.

76. Wojtaszczyk, P. *A Mathematical Introduction to Wavelet*; Cambridge University Press: Cambridge, UK, 1997.
77. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–666. [CrossRef]
78. Xu, D.; Erdogmuns, D. Renyi's Entropy, Divergence and Their Nonparametric Estimators. In *Information Theoretic Learning*; Springer: New York, NY, USA, 2010; pp. 47–102.
79. Krishnamurthy, A.; Kandasamy, K.; Póczos, B.; Wasserman, L.A. Nonparametric Estimation of Renyi Divergence and Friends. In Proceedings of the ICML, Beijing, China, 21–26 June 2014.
80. Pope, S.B. *Turbulent Flows*; Cambridge Univ. Press: Cambridge, UK, 2001.
81. Shannon, C.E. Communication in the presence of noise. *Proc. Inst. Radio Eng.* **1949**, *37*, 10–21. [CrossRef]