

## Predicting Urban Surface Roughness Aerodynamic Parameters Using Random Forest

G. DUAN<sup>a</sup> AND T. TAKEMI<sup>a</sup>

<sup>a</sup>Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan

(Manuscript received 2 December 2020, in final form 16 May 2021)

**ABSTRACT:** The surface roughness aerodynamic parameters  $z_0$  (roughness length) and  $d$  (zero-plane displacement height) are vital to the accuracy of the Monin–Obukhov similarity theory. Deriving improved urban canopy parameterization (UCP) schemes within the conventional framework remains mathematically challenging. The current study explores the potential of a machine-learning (ML) algorithm, a random forest (RF), as a complement to the traditional UCP schemes. Using large-eddy simulation and ensemble sampling, in combination with nonlinear least squares regression of the logarithmic-layer wind profiles, a dataset of approximately  $4.5 \times 10^3$  samples is established for the aerodynamic parameters and the morphometric statistics, enabling the training of the ML model. While the prediction for  $d$  is not as good as the UCP after Kanda et al., the performance for  $z_0$  is notable. The RF algorithm also categorizes  $z_0$  and  $d$  with an exceptional performance score: the overall bell-shaped distributions are well predicted, and the  $\pm 0.5\sigma$  category (i.e., the 38% percentile) is competently captured (37.8% for  $z_0$  and 36.5% for  $d$ ). Among the morphometric features, the mean and maximum building heights ( $H_{ave}$  and  $H_{max}$ , respectively) are found to be of predominant influence on the prediction of  $z_0$  and  $d$ . A perhaps counterintuitive result is the considerably less striking importance of the building-height variability. Possible reasons are discussed. The feature importance scores could be useful for identifying the contributing factors to the surface aerodynamic characteristics. The results may shed some light on the development of ML-based UCP for mesoscale modeling.

**KEYWORDS:** Atmosphere-land interaction; Large eddy simulations; Classification; Regression

### 1. Introduction

Turbulence-resolving large-eddy simulation (LES) within a mesoscale model is typically run over subregions, while a substantial portion of the domain remains more coarsely resolved. Urban topographies are generally considered as surface roughness, which acts as drag forcing, flux sources or sinks (Muñoz-Esparza et al. 2014; Haupt et al. 2020). The microscale dynamics within the urban canopy layer are usually not explicitly resolved and the momentum exchanges with the boundary layers aloft are typically parameterized in mesoscale modeling (Brown 2000; Mirocha et al. 2014; Shen et al. 2019; Edwards et al. 2020).

A large amount of the literature has been devoted into improved parameterization of surface-layer impacts in mesoscale models, which led to the well-known Monin–Obukhov similarity theory (MOST; Monin and Obukhov 1954). Efforts including experiments and observations of the surface-layer micrometeorological processes (e.g., Haugen et al. 1971; Kaimal et al. 1976; Olesen et al. 1984), which has enabled the intercomparison of MOST with measurements, continue contributing to improving its accuracy (Nazarian et al. 2020; Maronga et al. 2020b; Bou-Zeid et al. 2020). Knowledge on the quantitative influence of urban surface roughness in mesoscale modeling has been greatly advanced with the emerging of techniques, which have made direct and reliable measurements of turbulent eddies and fluxes more conveniently achievable (e.g., Kaimal and Businger 1963; Hanafusa et al. 1982; Kaimal

and Wyngaard 1990; Cheng et al. 2020). Following MOST, the surface-layer turbulent fluxes are connected with the gradient of the mean profiles (Foken 2006), namely, taking the mean wind velocity  $U(z)$ , for example:

$$U(z) = \frac{u_*}{\kappa} \left[ \ln \frac{z-d}{z_0} + \psi(\zeta) \right] \quad \text{and} \quad \frac{\partial U(z)}{\partial z} = \frac{-\bar{u}'w'}{u_* \kappa (z-d)} \phi(\zeta), \quad (1)$$

where  $u_*$  is the friction velocity,  $\kappa$  is the von Kármán constant,  $\zeta = z/L$  accounts for the thermal stability, and  $z_0$  and  $d$  are the surface roughness aerodynamic parameters (i.e., the roughness length and the zero-plane displacement, respectively). For neutral stratification, the Obukhov length,  $L \rightarrow \infty$ , and  $\psi(\zeta) = 0$  and  $\phi(\zeta) = 1$ . The surface momentum flux,  $-\bar{u}'w'$  (Reynolds shear stress), can be obtained via vertical integration over the Prandtl layer (the constant-flux layer) and then specified as the surface boundary conditions (Maronga et al. 2020a).

The roughness length  $z_0$  and the zero-plane displacement height  $d$  represent two key components of the surface layer scaling [Eq. (1)]. Urban boundary layer flows are strongly perturbed by urban geometries, that is, buildings, which are unevenly distributed and anisotropic in the 3D details. The perturbations are typically represented using roughness elements in mesoscale models. The physical processes in the roughness surface layer (RSL) and within the urban canopy layer are of high complexity (Roth 2000; Britter and Hanna 2003; Bou-Zeid et al. 2020). Accurate knowledge of the roughness parameters is important to the understanding of the urban surface aerodynamic characteristics, and hence vital to the accuracy in modeling the surface boundary conditions (Foken 2006).

Corresponding author: G. Duan, g.duan@storm.dpri.kyoto-u.ac.jp

The roughness length may be assigned in accord with land-use categories (e.g., Wieringa 1992; Davenport et al. 2000); however, the applicability may be limited because of the fact that full-scale observations are not always readily obtainable for all terrain types and meteorological conditions. By contrast, the micrometeorological method and the morphometric method are perhaps more suitable in the urban boundary layer literature (Grimmond and Oke 1999): for the former,  $z_0$  and  $d$  are determined through least squares regression of the logarithmic wind profile (e.g., Cheng et al. 2007); for the latter, they are parameterized in terms of the surface morphometric features for the discrete roughness elements mounted on the ground (e.g., Macdonald et al. 1998; Ratti et al. 2002; Haghshima et al. 2009). This includes the building-packing indices

$$\lambda_f \equiv A_f/A_T \quad \text{and} \quad \lambda_p \equiv A_p/A_T \quad (2)$$

(the frontal-area index and the plan-area index, respectively, where  $A_f$  is the total frontal area,  $A_p$  the total plan area of the roughness obstacles, and  $A_T$  is the total lot area), the mean and maximum building heights  $H_{\text{ave}}$  and  $H_{\text{max}}$ , and associated building height variability  $\sigma_H$ . For convenience,

$$X^i \equiv \{\lambda_p, \lambda_f, H_{\text{ave}}, H_{\text{max}}, \sigma_H\}. \quad (3)$$

The micrometeorological approach requires knowledge of a well-defined logarithmic wind profile a priori, which may not be well suited to operational models. For the morphometric method, deploying tall towers for field observations is not required, instead  $z_0$  and  $d$  are directly computed from the morphometric statistics. The latter approach has proven successful generating encouraging estimations of the roughness aerodynamic parameters for realistic urban topographies (Kanda et al. 2013), and hence would be more promising than the first approach, especially given the ever-growing availability of high-resolution geophysical data and with the assistance of building-resolving computational fluid dynamics (CFD) techniques (Hanna et al. 2006; Nakayama et al. 2011; Park et al. 2015b; Chew et al. 2020). Similarly, let

$$Y^i \equiv \{z_0, d\}. \quad (4)$$

Among those approaches, one aspect that has attracted little attention is how important each morphometric index is on the parameterization of the roughness aerodynamic parameters. It is demanding to assign weights precisely to each of the indices to account for the differences in the contribution to the parameterization. Formulating new schemes for improved parameterization within the conventional urban canopy parameterization (UCP) framework remains mathematically challenging. By contrast, machine learning (ML) is capable of handling a large number of entries. For an urban canopy, this could include all the morphometric indices [Eq. (3)] that may exert influences on the surface aerodynamic characteristics [Eq. (4)], potentially in a complex manner (Martilli et al. 2002; Akinlade et al. 2004) and likely to be nonlinear (Duan and Takemi 2021). ML allows the important features to be identified by ranking the feature importance scores.

Exploring the potential of ML toward predicting the surface roughness aerodynamic parameters, briefly,

$$X^i \rightarrow Y^i, \quad (5)$$

may also assist the development of improved UCP schemes.

Recent developments of ML techniques have led to successful results in the studies of complex physical problems across climate networks, computer science, gene regulation, biology, and medicine (Zitnik et al. 2019; Gagne II et al. 2020; Tang et al. 2020; Biesbroek et al. 2020). ML has demonstrated great potential dealing with problems as complicated as turbulence modeling. The feasibility of representing the unresolved subgrid turbulent processes in place of the traditional subgrid-scale (SGS) parameterization has been evaluated for a global general circulation model in Rasp et al. (2018). Cheng et al. (2019) explored the application of ML for SGS modeling in LES for atmospheric flows of varying stability conditions. The SGS stresses produced from ML were shown to be more accurate than the conventional Smagorinsky SGS models (Smagorinsky 1963; Bardina et al. 1980). The value of ML has also been well recognized in studies of chaotic dynamical systems (Lellep et al. 2020) and hydrodynamic environments (Li et al. 2020).

This work explores the applicability of ML for the prediction of urban surface roughness aerodynamic parameters. Following the introduction, section 2a describes the urban topographies considered for the simulation of a neutrally stratified turbulent boundary layer flow using LES (section 2b, with the model validation given in appendix A). An ensemble sampling method is introduced in section 3a. In combination with the conventional micrometeorological approach, the nonlinear two-parameter least squares regression (section 3b), a dataset of the urban morphometric indices and the aerodynamic parameters,  $z_0$  and  $d$ , is established for the training of a random forest (RF) ML model (section 4a). Section 4b describes the measures for identifying the morphometric features that could be of predominant influence on the prediction of the aerodynamic parameters. The results are analyzed in section 5, wherein section 5a presents vertical distributions of the mean wind and turbulence statistics, and section 5b analyses the  $z_0$  and  $d$  statistics from the least squares fit against the most recently proposed UCP scheme. Predicted  $z_0$  and  $d$  using RF are analyzed in section 5c with both RF regression [section 5c(1)] and RF classification [section 5c(2)] considered. Informative morphometric indices are identified in section 5d. Robustness is discussed in section 5e. Summary and discussion are given in section 6.

## 2. Numerical experiments

### a. Urban topography

This study investigates a neutrally stratified turbulent boundary layer flow that developed over a realistic urban topography. Four residential regions  $\mathcal{D}$  within Osaka, Japan, are considered (Fig. 1). As in Takemi et al. (2020), the topography data of Osaka City are obtained from the digital surface model (DSM) and the digital elevation model (DEM) with a

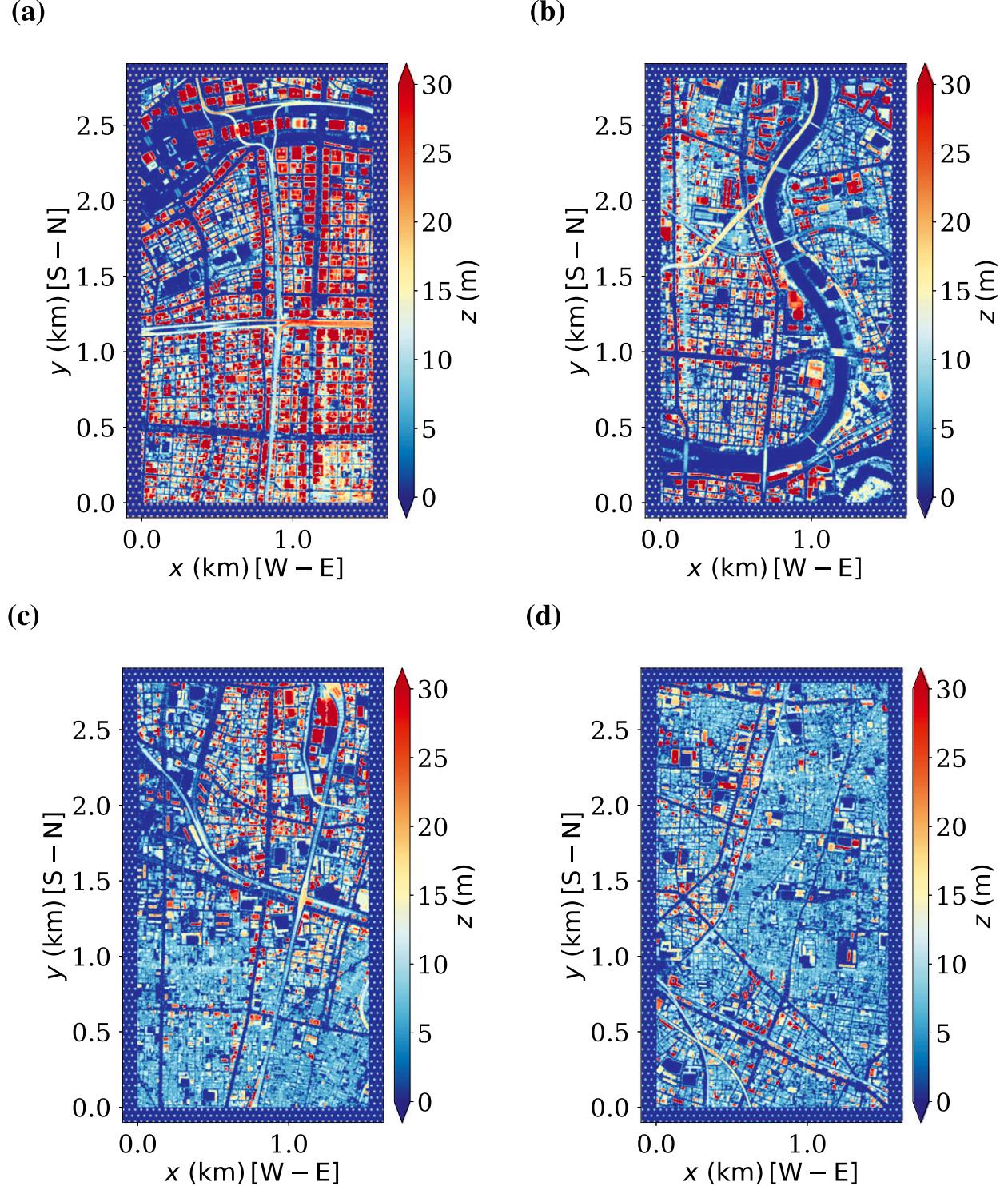


FIG. 1. Four urban areas within Osaka. The topographies are color shaded by the building heights. Shown are (a)  $H_{\text{ave},g} \pm \sigma_{H,g} = 13 \pm 10 \text{ m}$ , (b)  $H_{\text{ave},g} \pm \sigma_{H,g} = 11 \pm 7 \text{ m}$ , (c)  $H_{\text{ave},g} \pm \sigma_{H,g} = 16 \pm 16 \text{ m}$ , and (d)  $H_{\text{ave}} \pm \sigma_H = 24 \pm 19 \text{ m}$ . Herein  $H_{\text{ave},g}$  denotes the global average of the building height over the entire domain of each topography and  $\sigma_{H,g}$  is the associated building-height standard deviation. The analysis is performed for the four areas as a whole, hereinafter  $\mathcal{D}$ .

horizontal resolution of 2 m from the Kokusai Kogyo Co., Ltd. As a common practice in urban CFD, the ground elevation is excluded so as to be focused on the morphological characteristics associated with the roughness elements mounted on the ground surface. Each subregion possesses unique morphological profiles differing in building-packing densities, building-height variability and the roughness element arrangement, but we do not attempt to distinguish the results between the different subregions. Performing simulations for separate domains is simply to accommodate the computer power. The raw data are combined and analyzed as a whole.

The domain dimensions of the subregions are the same, that is,  $L_x = 1504 \times \Delta x = 3008$  m in the streamwise direction ( $x$ ),  $L_y = 864 \times \Delta y = 1728$  m in the spanwise direction ( $y$ ), and  $L_z \approx 500$  m in the vertical direction ( $z$ ). The spatial resolution is homogeneous in the lateral directions; that is,  $\Delta x = \Delta y = 2$  m, in accord with the DSM and DEM dataset. The grids are equally spaced in the vertical direction for the lower boundary layers [i.e.,  $\Delta z = \Delta z_{\min} = 2$  m for  $0 < 0 \leq 80$  m ( $\sim 4-8H_{\text{ave}}$ )] but are stretched at greater heights [i.e.,  $\Delta z_{k+1} = 1.08\Delta z_k$  for  $z > 80$  m (where  $k$  denotes the vertical grid index)], until  $\Delta z = \Delta z_{\max} = 16$  m, which is maintained up to the upper boundary of the domain, giving a total of 82 layers in the vertical direction. The ratio of the coarsest and finest vertical grid spacings considered in the current study is  $\Delta z_{\max}/\Delta z_{\min} = 8$ , which is smaller than Cui et al. (2004) and Michioka et al. (2019) by a factor of  $\sim 2$  and is smaller than Michioka and Sato (2012) by a factor of up to 4.75, allowing a much smaller grid deformation.

To avoid building-height discontinuities (that would normally arise from periodic boundary conditions for realistic urban surfaces), a buffer area filled with cubic roughness elements is set around each topography. Similar to Yoshida et al. (2018), the lateral dimensions (width and length) of the cuboids are fixed at 10 m. As in Duan and Takemi (2021), the height of the roughness elements is equal to the global average  $H_{\text{ave},g}$  of the building heights for each topography.

### b. Large-eddy simulations

The turbulent boundary layer flows are simulated using the Parallelized Large-Eddy Simulation Model (PALM) 6.0 (Maronga et al. 2020a), which is based on the implicitly filtered nonhydrostatic, incompressible Boussinesq equations and the 1.5-order Deardorff SGS scheme (Deardorff 1980). The setup of the 3D model essentially follows Duan and Takemi (2021), except that thermal effects are neglected. The model is initialized with the stationary solution from a Reynolds-average based turbulence parameterization and the flow is driven continuously by the same pressure gradient,  $\nabla p = (0.6 \times 10^{-3}, 0)$  Pa  $\text{m}^{-1}$ , for  $t_f = 7$  h (Gronemeier et al. 2021). As in many previous LES studies of urban turbulent boundary layer (TBL) flow (e.g., Lo and Ngan 2017; Takemi et al. 2020; Duan and Takemi 2021), Coriolis acceleration is neglected. This yields a moderate- to high-Reynolds-number TBL flow (Horiguchi et al. 2010; Wang and Anderson 2019),  $\text{Re} = U_\infty H_{\text{ave},g}/\nu = O(10^7)$  (where  $U_\infty$  is the mean freestream wind speed and  $H_{\text{ave},g}$  is the mean building height for  $\mathcal{D}$ ). The data for the final analysis are recorded every 10 s for the last 1 h of the simulation ( $t_f - t_0 = 3600$  s), which is in a statistically steady state.

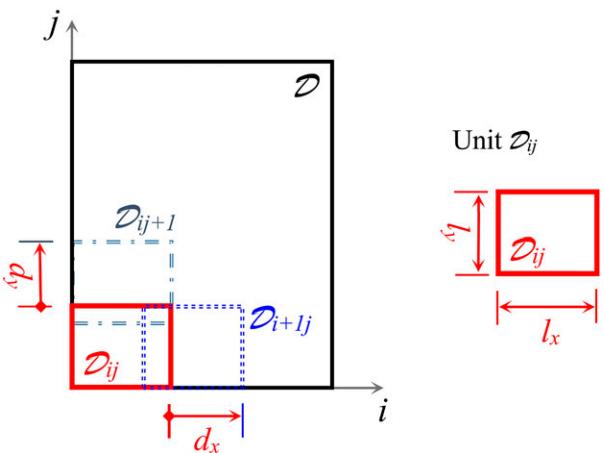


FIG. 2. Schematic illustration of the proposed sampling approach. The box outlined in red denotes the initial analysis area  $\mathcal{D}_{ij}$  of dimension  $l_x \times l_y$ . The updated analysis unit  $\mathcal{D}_{i+1,j}$  results from a displacement of  $\mathcal{D}_{ij}$  to the east by  $d_x$ . Similarly,  $\mathcal{D}_{ij+1}$  is a result of displacing  $\mathcal{D}_{ij}$  to the north by  $d_y$ . Note the overlap between neighboring analysis units.

The boundary conditions at the domain edges are the same: periodic in the lateral directions and Neumann at the top. At solid surfaces, there is no-slip for the velocity components and Neumann for scalars. The near-wall boundary conditions are parameterized following MOST: a Prandtl layer is assumed between the roughness height and the first grid level. As in Park et al. (2015b) for a full-scale LES of a realistic urban area, the roughness length for momentum,  $z_{0,m} = 0.1$  m. The occurrence of streak-like structures originating from the persistent periodic boundary conditions, also known as the spanwise locking of very large-scale structures (e.g., Hutchins and Marusic 2007; Fang and Porté-Agel 2015), is alleviated by applying shifted periodic boundary conditions in the streamwise direction, whereby the outlet boundary no longer align with the original inlet boundary but the one shifted by one grid spacing in the spanwise direction (Munters et al. 2016).

The governing equation in combination with the initial and boundary conditions are solved with the following numerical schemes. Third-order Runge–Kutta time-stepping scheme (Williamson 1980) is combined with a fifth-order finite differencing scheme for momentum and scalar advection (Wicker and Skamarock 2002). The Poisson equation for the pressure is solved with the multigrid method. The SGS eddy viscosity and diffusivity are parameterized by solving a prognostic equation for the SGS turbulent kinetic energy. The current LES model is validated by simulating a turbulent boundary layer flow over an array of roughness elements arranged in an identical manner as that implemented in the wind-tunnel experiments of Brown et al. (2000). The validation is presented in appendix A.

## 3. Dataset establishment

### a. Sampling

A dataset for the statistical analysis and the training of the random forest ML model in the latter sections is established

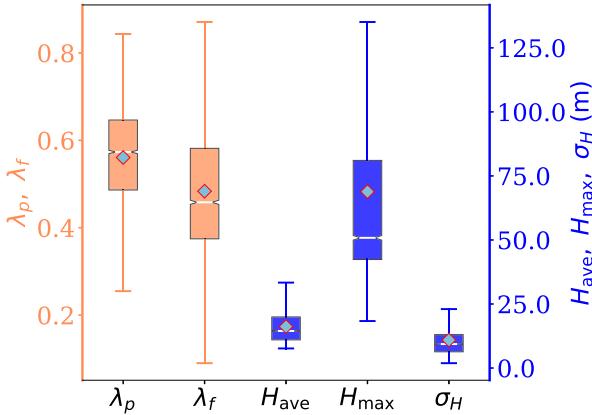


FIG. 3. Morphometric statistics of the studied domain (Fig. 1), applying the proposed sampling approach (Fig. 2). The red open diamonds denote the mean, and the short horizontal lines colored in white plot the median.

using an ensemble sampling approach after Duan and Takemi (2021). Briefly, the simulation data are analyzed for an ensemble of small urban units,  $\mathcal{D}_{ij}$ , within the domain,  $\mathcal{D}$  (see Fig. 1), that is,  $\mathcal{D}_{ij} \subset \mathcal{D}$  (see Fig. 2 for a schematic illustration). The analysis units  $\mathcal{D}_{ij}$  are defined by traversing a sampling square of specified lateral dimensions,  $l_x \times l_y$ , across  $\mathcal{D}$  with a displacement of  $d_x$  in the streamwise and/or a displacement of  $d_y$  in the spanwise directions. Herein, fixed unit dimensions  $l_x \times l_y = 250 \times 250 \text{ m}^2$  are combined with variable displacement distances,  $d_x \in \{100, 200, 300\} \text{ m}$  and  $d_y \in \{100, 200, 300\} \text{ m}$ . Relative to Yoshida et al. (2018), wherein  $d_x = l_x$  and  $d_y = l_y$ , the sample size in the current study is effectively increased, for example by a factor of 2<sup>2</sup> if  $d_x = (1/2)l_x$  and  $d_y = (1/2)l_y$ , allowing a more comprehensive characterization of the urban morphometric features.

Adopting this ensemble sampling approach, a dataset of various values for the surface roughness features, which include the building-packing indices ( $\lambda_f$  and  $\lambda_p$ ), the mean and maximum building heights ( $H_{\text{ave}}$  and  $H_{\text{max}}$ ) and associated building-height variability ( $\sigma_H$ ) [herein  $X$ , see also Eq. (3)], is created for the entire domain  $\mathcal{D}$  (Fig. 1). Given the displacement distances ( $d_x$  and  $d_y$ ) and the sampling-square dimensions ( $l_x$  and  $l_y$ ) considered in the current study, the dataset contains approximately  $4.5 \times 10^3$  analysis units. Figure 3 summarizes the morphometric statistics.<sup>1</sup> Except for  $H_{\text{max}}$ , which exhibits a strong positive skewness (implying the presence of skyscrapers), all of the other morphometric metrics are roughly centered around the mean.

The flow in the RSL is strongly perturbed by the surface roughness elements (Kastner-Klein and Rotach 2004). The dynamic statistics from LES, for example, flow and turbulence, should be processed before being used for the calculation of the roughness aerodynamic parameters (see section 3b) that

are subsequently incorporated into the dataset for ML. Following the common practice, a statistically representative vertical profile can be obtained via temporal and horizontal averaging. For convenience, angle brackets denote the horizontal averaging over an analysis unit,  $\mathcal{D}_{ij}$ , namely,

$$\langle \cdot \rangle \equiv \frac{1}{A_{ij}} \int_{A_{ij}} (\cdot) d\mathbf{x}, \quad (6)$$

where  $A_{ij}$  is the plan area of  $\mathcal{D}_{ij}$ . Each sampling unit contains  $(l_x/\Delta x) \times (l_y/\Delta y) = 125 \times 125$  grid points in a lateral plane, which is considered sufficient for statistically representative results through horizontal averaging. In Cheng and Castro (2002), the horizontal average was performed over 25 sampling profiles. It was also shown that a 4-profile average yielded very similar results to the 25-profile average, and a lateral average over  $25 \times 25$  grids was confirmed sufficient in comparison with a  $50 \times 50$  choice (Wang et al. 2018). The number of grid points in the horizontal plane (each grid point corresponds to one profile) is significantly greater than that used in Cheng and Castro (2002) and is larger by a factor of 6 in comparison with Wang and Ng (2018). As in Duan and Takemi (2021), an overbar denotes time averaging over the duration of the data collection ( $t_f - t_0$ ),

$$\bar{\cdot} \equiv \frac{1}{t_f - t_0} \int_{t_0}^{t_f} (\cdot)(t) dt. \quad (7)$$

The averaging time span is comparable to  $\sim 100\tau_{\text{eddy}}$ , where  $\tau_{\text{eddy}} = H_{\text{ave},g}/u_*$  denotes the eddy turnover time scale (see Figs. 3 and 7 for a rough estimation of the mean building height  $H_{\text{ave},g}$  and the friction velocity  $u_*$ , respectively) and should suffice for meaningful statistics (cf. Cheng and Liu 2011).

Ensemble averaging over all sampling units, denoted with a caret, will also be considered. The calculation follows Duan and Takemi (2021).

#### b. Roughness aerodynamic parameters, $z_0$ and $d$

The conventional micrometeorological approach estimating the surface roughness aerodynamic parameters, that is, the roughness length  $z_0$  and the zero-plane displacement height  $d$ , is generally performed through two-parameter regression of the logarithmic wind profile,

$$\langle \bar{u} \rangle = \frac{u_*}{\kappa} \ln \left( \frac{z - d}{z_0} \right), \quad (8)$$

by minimizing the following cost function (Cheng et al. 2007; Hagishima et al. 2009),

$$E = \sum_{z_i=H_{\log,l}}^{H_{\log,u}} \left[ \frac{\kappa \langle \bar{u} \rangle(z_i)}{u_*} - \ln \left( \frac{z_i - d}{z_0} \right) \right]^2, \quad (9)$$

where  $\kappa$  is the von Kármán constant, usually taken to be 0.4, and  $[H_{\log,l}, H_{\log,u}]$  define the vertical (grid) range of the logarithmic region.

The friction velocity reflects the balance between the flow and the underlying roughness surface, and has been an important velocity scale in urban canopy parameterization

<sup>1</sup> The data used for the random forest ML model are scaled to ensure zero mean and unit variance (see Figs. 7b,c).

(Britter and Hanna 2003). It can be calculated in two ways that differ in the physical meaning. One is based on the drag force exerted by the roughness elements (e.g., Cheng et al. 2007; Hagishima et al. 2009; Buccolieri et al. 2019),

$$u_* = \sqrt{\tau_p/\rho}. \quad (10)$$

Alternatively,  $u_*$  can be obtained directly from the eddy covariance statistics of the surface layer (e.g., Rotach 1993a,b; Oikawa and Meng 1995; Raupach et al. 1996),

$$u_* = \sqrt{|\langle -u'w' \rangle|}. \quad (11)$$

The former [Eq. (10)] requires knowledge of the pressure distribution along individual frontal and back surfaces of the roughness elements, which may be more suitable for idealized building obstacles (e.g., Hagishima et al. 2009; Kanda et al. 2013); however, the calculation is not straightforward for realistic urban topographies (cf. Kanda et al. 2013). The high inhomogeneity and irregular distribution of building heights as well as the nonuniform building facades preclude direct measurements of the drag force. By contrast, in the latter approach [Eq. (11)]  $u_*$  is directly derived from the Reynolds shear stress,  $-u'w'$ , which can be readily obtained from the resolved-scale wind speeds. While both methods have been applied in previous studies, it is argued that the second one reconciles well with field campaigns, wherein the eddy covariance statistics are more conveniently measurable than the drag. For convenience, the flux,  $-u'w'$  that averaged throughout the RSL, is used for the calculation of  $u_*$  (cf. Cheng and Castro 2002). Therefore,  $u_*$  serves as a characteristic velocity scale for the roughness surface layers. This yields a roughness Reynolds number,  $Re_* \equiv u_* z_0 / v = O(10^4)$ , which is comparable to the LES results of Kanda et al. (2013) (for  $\kappa \sim 0.4$ ) and represents an extremely rough surface condition (cf. Inagaki and Kanda 2008).

Applying the proposed sampling approach (section 3a) in combination with the two-parameter nonlinear least squares fitting, a dataset of approximately  $4.5 \times 10^3$  samples is established for  $z_0$  and  $d$  [herein  $Y$ ; see also Eq. (4)] and incorporated into the dataset created in section 3a, enabling the training of the ML model and allowing the results to be analyzed in a statistical manner in the latter sections.

## 4. Machine learning

### a. Random forest

A relatively less sophisticated ML model is employed, herein a random forest, which is nonetheless one of the most powerful supervised ML algorithms that is capable of both regression and classification (Breiman 1996). RF is an ensemble ML algorithm built from multiple decision trees that are trained independently on random subsamples of the training dataset, known as bootstrapping. It has been successfully applied in studies of land surface phenology (Rodríguez-Galiano et al. 2016) and atmospheric boundary layers (Bodini et al. 2020).

RF is able to handle a very large database, which could include statistics for all the morphometric indices and the roughness aerodynamic parameters (herein  $D_n$ ) established in section 3. Formally,

$$\mathcal{D}_n = \{(X^i, Y^i)\}_{i=1}^n, \quad (12)$$

where  $(X^i, Y^i)$  denotes an element pair of the feature set  $X [\lambda_p, \lambda_f, \sigma_H, H_{\text{ave}}, \text{ and } H_{\text{max}}]$ ; see Eq. (3)] and the response set  $Y [z_0 \text{ and } d, \text{ see Eq. (4)}]$ ; that is,  $(X^i, Y^i) \in (X, Y)$ . The objective is to predict  $Y^i$  for given  $X^i$  through an RF algorithm [see Eq. (5)] that trained on bootstrap samples of  $\mathcal{D}_n$ .

The final prediction is obtained through averaging (for regression) or voting the majority (for classification) over the ensemble trees, known as aggregating, making it more effective than any of the individual decision tree. A simplified mathematical illustration of ML using RF is given in appendix B.

### b. Feature importance measures

Identifying the most important features (e.g.,  $\lambda_p$ ,  $\lambda_f$ ,  $\sigma_H$ ,  $H_{\text{ave}}$ , and  $H_{\text{max}}$ ) that affect the target variables ( $z_0$  and  $d$ ) is not straightforward in the formulation of conventional UCP. By contrast, RF is capable of differentiating informative predictors from noninformative ones. This can be done by ranking the importance metrics associate with the features, which include the decrease of the node impurity on each feature (herein  $I_G$ ) and the change of model performance scores before and after shuffling the values (permutation) of a feature (herein  $I_P$ ) (Menze et al. 2009). See appendix C for a brief description of  $I_G$  and  $I_P$ .

Large values of  $I_G$  and  $I_P$  indicate that the associated features are informative, while small ones imply that the features are of less importance. The feature importance metrics will be analyzed in section 5d for each of the morphometric indices, wherein the ones that could be of substantial influence on the prediction of the surface roughness aerodynamic parameters are identified.

## 5. Results

### a. Mean and turbulence profiles

Figure 4 plots the temporally and laterally averaged wind profiles for the ensemble of the urban units. While the profiles are noticeably affected by the surface roughness parameters,  $H_{\text{ave}}$  and  $\sigma_H$ , the general pattern follows a familiar picture: the wind speed increases away from the surface roughness, followed by a likely exponential (within  $z/H_{\text{ave}} \sim 5\text{--}20$  for high  $\sigma_H$  and  $z/H_{\text{ave}} \sim 10\text{--}30$  for low  $\sigma_H$ ), and eventually converges toward the freestream velocity at greater heights.

A well-defined inertial sublayer (ISL) may not be always expected for a vertical (grid) range, although the wind profiles plotted in Fig. 4 exhibit a strong exponential. Figure 5 plots the vertical profiles for the second-order turbulence statistics. The general pattern of the Reynolds shear stress (Fig. 5a) shows a common feature with previous studies (Raupach et al. 1996). The fluxes reveal a great inhomogeneity in the vertical: there is a rapid decay toward the urban surface and the overall

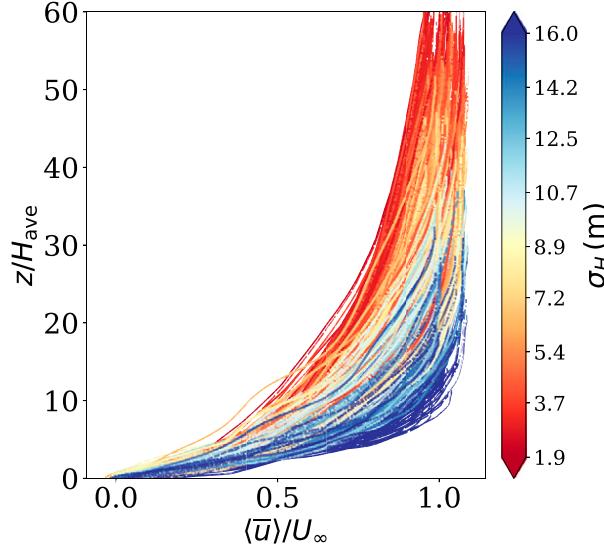


FIG. 4. The mean wind profiles associated with the ensemble of the analysis urban units. The overbar denotes the temporal average, and the angle brackets indicate the horizontal average [Eq. (6)]. The profiles are shaded by  $\sigma_H$ .

strength is weak within the canyons ( $z/H_{\text{ave}} \lesssim 1$ ); peak values are obtained above the urban canopy layer ( $z/H_{\text{max}} \lesssim 1$ ) and the profiles converge to small values as the height approaches the freestream level. There is a similar pattern for the turbulent kinetic energy (Fig. 5b). The results agree well with the wind-tunnel measurements of Kastner-Klein and Rotach (2004) for a scaled 3D realistic urban topography.

Contrary to the field campaigns of Rotach (1995) and Oikawa and Meng (1995), wherein the turbulence statistics peaked around 1–1.5 times the canopy height, the maxima herein are observed at greater heights,  $z/H_{\text{ave}} \sim 3$ –10. This is perhaps because the canopy morphology was less of a concern in the field campaigns: the former was performed for a close-to-unity street canyon and the latter for a flat space downstream a residential area. For a vast portion of the profiles, the turbulence statistics attain the maximum around  $H_{\text{max}}$  (indicated in the inset “ $H_{\text{max}}/H_{\text{ave}}$ ” as short horizontal lines). This is consistent with the previous finding of Kanda et al. (2013) for realistic urban areas within Tokyo and Nagoya, Japan; however, we note that this is less so for regions of small building-height variability.

While a very narrow ISL is vaguely seen for  $z/H_{\text{ave}} \sim 10$  and low  $\sigma_H$  (overlapped by the high  $\sigma_H$  profiles), the fluxes are not even close to constant, instead exhibiting a quick convergence to small values for  $z/H_{\text{ave}} \gtrsim 10$ , and this is doubly so for areas of high  $\sigma_H$ . Nevertheless, it will be seen in the latter sections that the lack of a well-defined ISL does not preclude the existence of a logarithmic layer and hence the two-parameter nonlinear regression for the surface roughness aerodynamic parameters,  $z_0$  and  $d$ .

#### b. $z_0$ and $d$ , conventional micrometeorological approach

This section presents  $z_0$  and  $d$  obtained through the conventional micrometeorological approach, that is, least squares regression of the logarithmic wind profile (see section 3b). We show in Fig. 6 that all of the logarithmic profiles are well regressed and the regression errors are small ( $\lesssim 5\%$  for both  $z_0$  and  $d$ ). Note that the vertical (grid) range for the plotting is extended beyond the logarithmic region so that the profiles

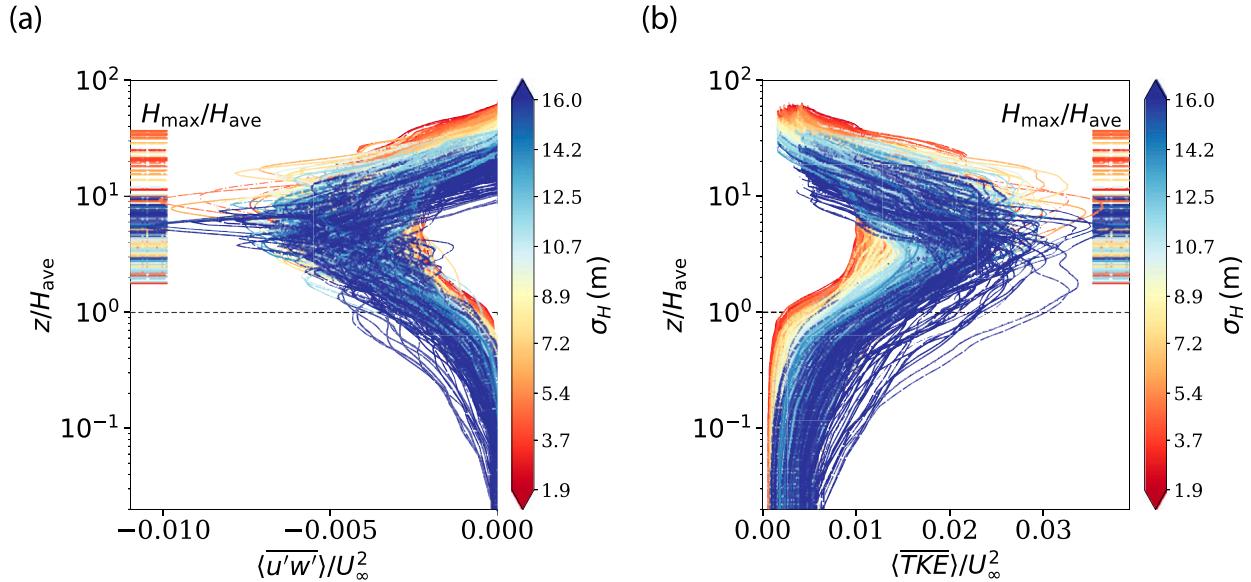


FIG. 5. As in Fig. 4, but for the second-order turbulence statistics: (a) turbulent momentum flux  $u'w'$  and (b) turbulent kinetic energy (TKE). Note the logarithmic scale on the ordinate axis. The inset figure legend plots the normalized maximum building height  $H_{\text{max}}/H_{\text{ave}}$  that is associated with each profile.

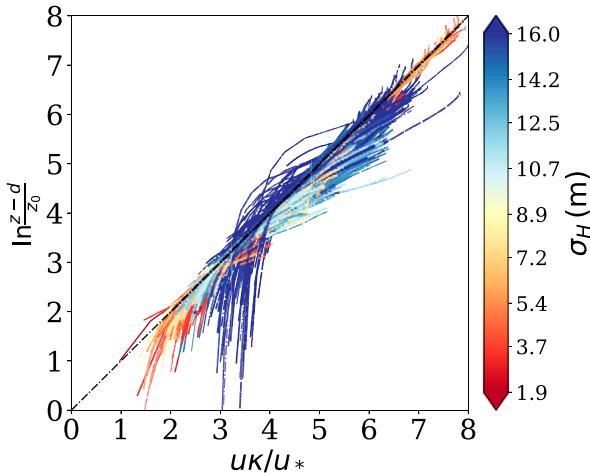


FIG. 6. Log-law fitting. The vertical (grid) range for the plotting is extended beyond the logarithmic region so that the profiles associated with different analysis units may be better distinguished.

associated with the different analysis units may be better distinguished. The profiles collapse well in the logarithmic layer with almost the same slope, and intercept at different vertical levels with the theoretical line (cf. Blocken et al. 2007). Despite a few scattered ones, the profiles are considerably well regressed and even for the sampling units of extremely high building-height variability (e.g.,  $\sigma_H \geq 15$ , indicated in the colorbar), the regression is in most cases well accommodated (see the  $R^2$  squared in Fig. 7a). Samples that do not yield satisfactory regression results (largely attributable to the narrow vertical grid range) will be simply discarded given the sufficiently large sample size (see section 3). Alternatively, one can always tweak the vertical (grid) range manually and visually observe if the logarithmic layer is well located (e.g., Kanda et al. 2013).

Figure 7 plots a summary of the statistics for the normalized roughness aerodynamic parameters ( $z_0/H_{\text{ave}}$  and  $d/H_{\text{ave}}$ ) that obtained from the two-parameter nonlinear least squares regression and the quality of the fitting ( $R^2$ ). While  $R^2$  (Fig. 7a) varies among the sampling units, the values are overall above 0.98. The mean is substantially lower than the median, indicating that most values of  $R^2$  are higher than the average, which suggests good quality of the regression ( $R^2 > 0.99$ ). As compared with the roughness length ( $z_0/H_{\text{ave}}$ ), the spread of the displacement height ( $d/H_{\text{ave}}$ ) exhibits a strong positive skewness, generally implying the presence of high-rise buildings (cf. Ratti et al. 2002). A heavier right tail is seen for both  $z_0$  (Fig. 7c) and  $d$  (Fig. 7d). The probability density functions (pdf) follow well a skewed Gaussian distribution (RMSE < 0.1): the skewness and kurtosis for  $z_0$  are 1.0 and 4.3, respectively, and are 1.5 and 4.8 for  $d$ .

Figures 8a and 8b compare  $z_0/H_{\text{ave}}$  and  $d/H_{\text{ave}}$  versus the plan-area index  $\lambda_p$  for the least squares regression against that from the parameterization after Kanda et al. (2013). While the UCP scheme has proven an overall good performance (reported correlation coefficients squared  $R^2$  were 0.55 for  $z_0$  and 0.78 for  $d$ ), nonnegligible discrepancies are seen between

measured (LES) and predicted (UCP) values. This can be partially attributed to the friction velocity  $u_*$ , which differed between the two studies. In principle, one could directly compare the regressed  $z_0$  and  $d$  with the predicted values; however, the qualitatively different methods for the calculation of  $u_*$  preclude pointwise comparison (see section 3b for a brief discussion). Nevertheless, we try to demonstrate this by comparing the current  $u_*$  (denoted as  $u_{*,\text{RSL}}$  in Fig. 8c) with values obtained from another method that is based on the peak  $-u'w'$  (denoted as  $u_{*,\text{peak}}$ ). The latter approach is also often used in previous studies of urban TBL flow (e.g., Rotach 1999; Kastner-Klein and Rotach 2004).

Despite some scattered points for large  $H_{\text{max}}/H_{\text{ave}}$ , a vast majority of the data are well clustered for areas of  $H_{\text{max}}/H_{\text{ave}} \sim 5$  and the relative discrepancies between  $u_{*,\text{peak}}$  and  $u_{*,\text{RSL}}$  are roughly within  $\pm 20\%$ . Figure 8c may be interpreted in the following manner. Large abscissa values ( $H_{\text{max}}/H_{\text{ave}}$ ) in combination with low building-height variability ( $\sigma_H$ , indicated in the colorbar) imply the presence of skyscrapers, which is in contrast to a cluster of high-rise buildings. For the latter,  $H_{\text{max}}/H_{\text{ave}}$  is generally small, while  $\sigma_H$  is high. Scenarios, wherein both  $H_{\text{max}}/H_{\text{ave}}$  and  $\sigma_H$  are small, essentially indicate a more homogeneous surface. It is noticeable that large discrepancies of  $u_*$  arise for high  $H_{\text{max}}/H_{\text{ave}}$  ( $\geq 20$ ) and relatively low  $\sigma_H$ , that is, regions that include skyscrapers, and the associated  $u_*$  discrepancy is  $>40\%$ . This is consistent with the previous finding of Kanda et al. (2013). For brevity,  $u_*$  is used instead of the notation  $u_{*,\text{RSL}}$  in other places of the text.

Plotting  $z_0/H_{\text{ave}}$  versus  $\lambda_p$  and so for  $d/H_{\text{ave}}$  (Figs. 8a,b, respectively) helps to highlight the discrepancies that may not be easily resolved using the conventional UCP, herein the overestimation of the scatter for the former and the underestimation for the latter. It was indicated in Kanda et al. (2013) that the parameterization needs to be improved for areas with high-rise buildings centered in the domain or a vast of misalignment of building facades with the dominant wind direction. The influence of tall bluff elements on  $z_0$  and  $d$  was also noted in Hagishima et al. (2009) though for idealized uniform and nonuniform building arrays. Obviously more geometric factors affecting the surface roughness heterogeneity and the uneven distribution of the roughness elements may need to be considered or the weights quantifying the importance of each contributing factor be adjusted for improved performance of the parameterization.

### c. $z_0$ and $d$ , random forest

#### 1) RF REGRESSION

Figure 9 compares the roughness aerodynamic parameters,  $z_0$  (Fig. 9a) and  $d$  (Fig. 9b), obtained from the RF regression (“predicted”) with those from the conventional least squares regression (“measured”; see sections 3b and 5b). A typical 75%–25% training-validation ratio (e.g., Were et al. 2015) is used for the random split of the original dataset. Despite a few outliers in Fig. 9a, the performance of the prediction using RF ML reaches up to  $R^2 = 0.94$  (RMSE = 0.15 m) for  $z_0$  and  $R^2 = 0.99$  (RMSE = 2.33 m) for  $d$  as compared with the

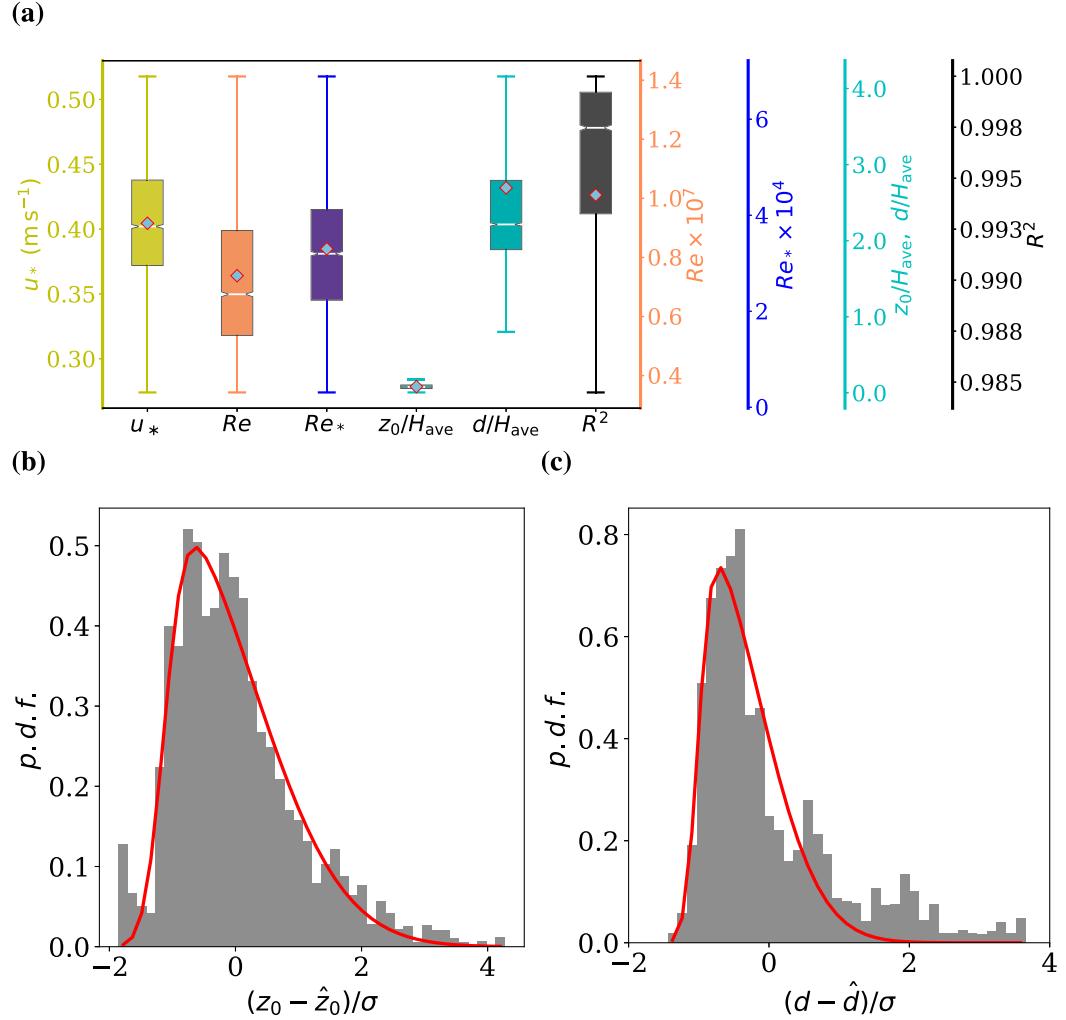


FIG. 7. (a) A summary of the dynamic characteristics, (b) pdf for  $z_0$ , and (c) pdf for  $d$ . The  $R$  squared in (a) reflects the quality of the two-parameter nonlinear least squares regression of the wind profiles (Fig. 4) toward the theoretical logarithmic profile (Fig. 6). For compactness, statistics of  $Re$  (see section 2b),  $u_*$ , and  $Re_*$  (see section 3b) that are associated with the ensemble of the analysis units are plotted in the same panel with the roughness aerodynamic parameters. As in Fig. 3, the red open diamonds denote the mean and the short horizontal lines colored in white plot the median. The red curves in (b) and (c) denote the fitting to skewed Gaussian distributions. The caret denotes the ensemble average over all sampling units (see Duan and Takemi 2021), and  $\sigma$  is the associated standard deviation. The data of  $z_0$  and  $d$  are normalized to ensure zero mean and unit variance.

conventional UCP algorithm: see Kanda et al. (2013), wherein  $R^2 = 0.55$  for  $z_0$  and  $R^2 = 0.78$  for  $d$ . Once the relationship is established, one can conveniently predict the roughness aerodynamic parameters for given surface topographical features. We have confirmed that a 80%–20% random split of the dataset for training and testing (e.g., Ghorbani et al. 2020) does not affect the results qualitatively.

The exact performance of the model depends on various factors, which include the number of trees (or estimators) in a forest, the fraction of the original dataset allocated for the training of an individual tree and the number of features to find the best split, the minimum required number of samples in a node to allow further splitting and the maximum depth that a

tree can be grown. The optimum combination of the hyperparameters is obtained through cross validation.

## 2) RF CLASSIFICATION

As compared with the regression, which predicts values, classification is useful in distinguishing between urban surfaces of differential roughness and aerodynamic characteristics in accord with the aerodynamic parameters for given topographic entries. To do so, values of the aerodynamic parameters need to be separated into discrete classes that label different value ranges, herein 10 for  $z_0$  ( $z_{01} - z_{10}$ ) and 9 for  $d$  ( $d_1 - d_9$ ). The bounds for neighboring ranges may be defined in a practical manner following Grimmond and Oke (1999); however, for

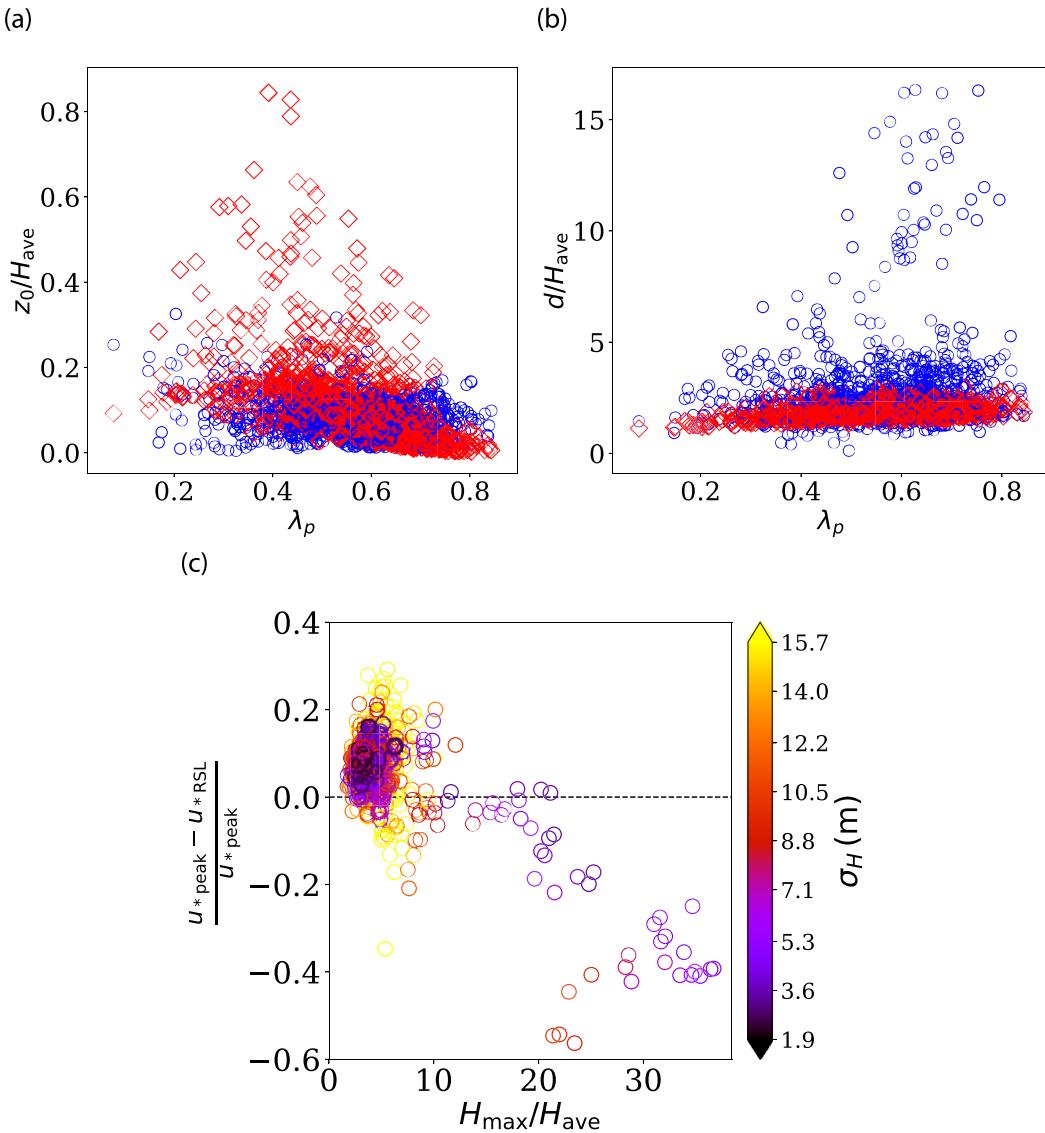


FIG. 8. Normalized (a) roughness length  $z_0$  and (b) zero-plane displacement height  $d$  vs the plan-area building packing density  $\lambda_p$ . Red open diamonds denote values obtained using the UCP scheme after Kanda et al. (2013), and blue open circles are from the least squares regression of the current LES data. (c) Comparison of the friction velocities estimated through two different approaches:  $u^*_{\text{peak}}$ , which is based on peak  $-u'w'$  (e.g., Rotach 1999; Kastner-Klein and Rotach 2004), and  $u^*_{\text{RSL}}$ , which is based on averaged  $-u'w'$  throughout the RSL (the current study).

simplicity they are assigned in a statistical way, that is in terms of the distance (in units of the standard deviation) to the ensemble mean (see Table 1 for a summary of the discrete classes and the corresponding value ranges).

The performance of the classification is quantified using confusion matrix in Fig. 10. Except for some negligible amount of confusion, for example, when attempting to distinguish  $z_04$  from  $z_08$  and  $d9$  from  $d6$  (highlighted in the dash-dotted blue squares), the overall performance of the classification is exceptional. The fraction of predictions that the model correctly classified is as high as 0.97 for both  $z_0$  and  $d$ . There is a well-defined off-diagonal pattern—the distribution of correct predictions

roughly follows a bell shape, which preserves the nature of the original dataset (see Fig. 7), and the  $\pm 0.5\sigma$ , that is, the 38% percentile (see Table 1), is well captured: 37.8% for  $z_0$  and 36.5% for  $d$  (see the intersectional portion of the horizontal and vertical bands,  $z_04$  for Fig. 10a and  $d3$  for Fig. 10b).

#### d. Feature importance identification

Figure 11 plots the model performance scores (definitions see section 4b),  $I_G$  (Figs. 11a,b) based on the reduction of node impurity upon each feature and  $I_P$  (Figs. 11c,d) based on the decrease of model performance, for the RF regression. The surface morphometric features exhibit greatly differing

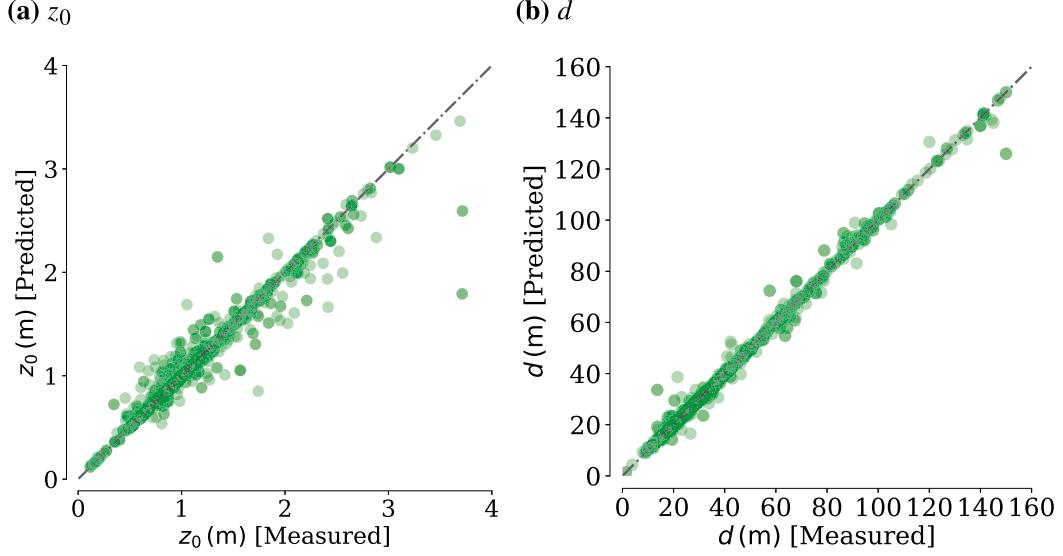


FIG. 9. The RF regression of (a)  $z_0$  and (b)  $d$ . The dash-dotted lines denote an exact 1:1 relationship.

influence on the roughness aerodynamic parameters. For  $I_G$  (Figs. 11a,b), the average and maximum building heights ( $H_{ave}$  and  $H_{max}$ ) account for up to 60% of the influence on the prediction of  $z_0$  (Fig. 11a). Although the building-height variability  $\sigma_H$  is relatively less informative in comparison with the building packing indices ( $\lambda_p$  and  $\lambda_f$ , ~30%), the information contribution is nonnegligible (~10%). By contrast, the displacement height,  $d$  (Fig. 11b), is shown to be almost exclusively determined by  $H_{max}$ : the  $I_G$  score is up to 90%. The  $I_P$  scores (Figs. 11c,d) exhibit a rough agreement with  $I_G$ . While the orders of the feature indices for  $I_P$  and  $I_G$  do not exactly coincide—for example, that for  $\lambda_p$  and  $\lambda_f$  and also  $H_{ave}$  and  $H_{max}$  are swapped in Fig. 11c as compared with Fig. 11a and  $\lambda_f$  and  $\sigma_H$  are swapped in Fig. 11d as compared with Fig. 11b, the general trends of the feature importance agree within the error bars, and the agreement is markedly better for  $d$ .

Actually, the importance of the maximum building height,  $H_{max}$ , was also emphasized in Kanda et al. (2013). A rough but direct evidence is the inclusion of  $H_{max}$  for the parameterization of  $z_0$  and  $d$  in Kanda et al. (2013) as compared with the original formulas of Macdonald et al. (1998). Despite that  $H_{max}$  does not appear explicitly in the parameterization of  $z_0$ , the influence is implicitly accounted for through the building-height variability,  $\sigma_H$ . Distinct aerodynamic characteristics were also observed behind high-rise buildings in a densely built-up urban area in Seoul (Park et al. 2015a).

Figure 12 plots the  $I_G$  (Figs. 12a,b) and  $I_P$  (Figs. 12c,d) scores for the RF classification. It is surprising that the  $I_G$  scores imply a qualitatively different effect of the surface morphometric features on the classification of the roughness aerodynamic parameters than on the regression (cf. Fig. 11). Taking  $z_0$  (Figs. 12a,c) for example, the  $I_G$  scores (Fig. 12a) suggest that the morphometric features considered herein are almost equally important for the classification, while the  $I_P$  scores (Fig. 12c) support the argument that the morphometric indices are unevenly informative (also note the swapped feature

order). This is perhaps because the labeling of classes for  $z_0$  and  $d$  (see Table 1) artificially reduces the randomness, which consequently biases the feature importance scores for the classification.

Discrepancies of the feature importance scores between  $I_G$  and  $I_P$  may be also attributed to the difference in the nature of the two feature importance measures (introduced in section 4b), namely  $I_G$  is biased because of the correlation between the impurity reduced by one feature and that reduced by the other as the tree being grown deeper, as opposed to  $I_P$ , which does not have the correlation issue and is unsusceptible to random noise. Importantly, the feature importance scores of  $I_P$  for the classification (Figs. 12c,d) exhibit a qualitative agreement with the scores obtained for the regression (Fig. 11), implying the robustness of  $I_P$  as compared with  $I_G$  as a feature importance measure.

#### e. Robustness

The sample size was effectively increased adopting the ensemble sampling approach (section 3), which enables the

TABLE 1. Labeling of the discrete value ranges of  $z_0$  and  $d$ ; the caret, or “hat,” denotes the ensemble average. The row with  $z_04$  and  $d3$  indicates classes that occupy the  $\pm 0.5\sigma$  or the 38% percentile.

	$z_0$	$d$
$z_01$	$z_0 \in [\hat{z}_0 - 2.0\sigma, \hat{z}_0 - 1.5\sigma]$	—
$z_02$	$z_0 \in [\hat{z}_0 - 1.5\sigma, \hat{z}_0 - 1.0\sigma]$	$d1 \quad d \in [\hat{d} - 1.5\sigma, \hat{d} - 1.0\sigma]$
$z_03$	$z_0 \in [\hat{z}_0 - 1.0\sigma, \hat{z}_0 - 0.5\sigma]$	$d2 \quad d \in [\hat{d} - 1.0\sigma, \hat{d} - 0.5\sigma]$
$z_04$	$z_0 \in [\hat{z}_0 - 0.5\sigma, \hat{z}_0 + 0.5\sigma]$	$d3 \quad d \in [\hat{d} - 0.5\sigma, \hat{d} + 0.5\sigma]$
$z_05$	$z_0 \in [\hat{z}_0 + 0.5\sigma, \hat{z}_0 + 1.0\sigma]$	$d4 \quad d \in [\hat{d} + 0.5\sigma, \hat{d} + 1.0\sigma]$
$z_06$	$z_0 \in [\hat{z}_0 + 1.0\sigma, \hat{z}_0 + 1.5\sigma]$	$d5 \quad d \in [\hat{d} + 1.0\sigma, \hat{d} + 1.5\sigma]$
$z_07$	$z_0 \in [\hat{z}_0 + 1.5\sigma, \hat{z}_0 + 2.0\sigma]$	$d6 \quad d \in [\hat{d} + 1.5\sigma, \hat{d} + 2.0\sigma]$
$z_08$	$z_0 \in [\hat{z}_0 + 2.0\sigma, \hat{z}_0 + 2.5\sigma]$	$d7 \quad d \in [\hat{d} + 2.0\sigma, \hat{d} + 2.5\sigma]$
$z_09$	$z_0 \in [\hat{z}_0 + 2.5\sigma, \hat{z}_0 + 3.0\sigma]$	$d8 \quad d \in [\hat{d} + 2.5\sigma, \hat{d} + 3.0\sigma]$
$z_{10}$	$z_0 \geq \hat{z}_0 + 3.0\sigma$	$d9 \quad d \geq \hat{d} + 3.0\sigma$

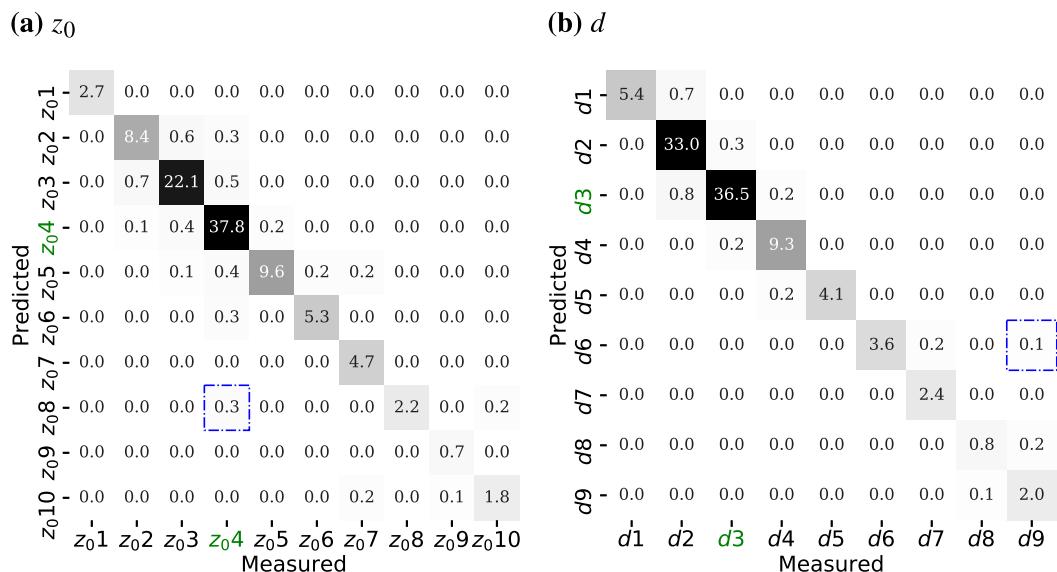


FIG. 10. Confusion matrices for the classification of (a)  $z_0$  and (b)  $d$ . The axial labels highlighted in green denote value ranges that lie within one-half standard deviation ( $\pm 0.5\sigma$ , or the 38% percentile) about the associated ensemble averages (see Table 1). The numbers in the cells denote the percentage of predictions, and the off-diagonal elements count the correctly predicted ones (%). The dash-dotted blue squares annotate the negligible amount of confusion: 0.3% of values in class  $z_04$  are incorrectly classified into class  $z_08$ ; similarly for the classification of  $d$ .

training of the ML model; however, the overlapping of neighboring units<sup>2</sup> may result in correlations in the calculated roughness aerodynamic parameters between adjacent analysis units. While this should be less of an issue for the RF classification since it is unlikely that values of the aerodynamics parameters of overlapped units would be separated into different classes [see section 5c(2)], it would possibly lead to overfitting of the RF regression. It is hard to acquire a large dataset with completely independent samples meanwhile maintaining a sufficient characterization of the surface topographic features; however, the performance scores in section 5c need to be interpreted with caution.

It is prudent to confirm the performance of the RF model for a dataset that is devoid of the neighboring overlap. This is now examined for  $d_x = l_x = 250$  m and  $d_y = l_y = 250$  m. As expected, the significant issue encountered in practice is the substantial decrease of the sample size (by more than an order of magnitude) and as a consequence, the surface morphometric features are also much less comprehensively represented. Unfortunately, there seems no economical means to resolve this other than considering more topographies. It is impractical and expensive to increase the sample size to an equivalent amount as before by doing so. Herein large-eddy simulations are performed for six extra subregions taken from Tokyo (see appendix D). The spatial dimensions and model setups remain the same as that for the Osaka topographies plotted in Fig. 1.

Relative to the prediction in section 5c, a significant drop is experienced ( $R^2 = 0.65$  for  $z_0$  and  $R^2 = 0.66$  for  $d$ ; figures not

shown), particularly for  $d$ , which is not as well predicted as the UCP scheme proposed in Kanda et al. (2013). The evidence to support the argument that an ML model could perform superiorly to the conventional UCP is inconclusive. Nevertheless, the slightly improved prediction of  $z_0$  is suggestive that ML can still be considered as a valuable complement to the conventional UCP once datasets are available for a sufficient training.

## 6. Summary and discussion

The current study explored the potential of ML as a complement to the conventional UCP schemes. Using LES and the traditional micrometeorological method, that is two-parameter nonlinear least squares regression of the logarithmic-layer wind profiles, in combination with the ensemble sampling approach (Duan and Takemi 2021), a dataset of approximately  $4.5 \times 10^3$  samples was established for the roughness aerodynamic parameters ( $z_0$  and  $d$ ) and the urban surface morphometric statistics. The latter included the building packing indices ( $\lambda_f$  and  $\lambda_p$ ), mean and maximum building heights ( $H_{ave}$  and  $H_{max}$ ) and associated standard deviations ( $\sigma_H$ ). A random forest algorithm (Breiman 2001), one of the most powerful ML algorithms that is capable of both regression and classification, was subsequently trained and implemented for the prediction of  $z_0$  and  $d$ , which are vital to the accuracy of the MOST.

The RF algorithm predicted the categories of  $z_0$  and  $d$ , which were separated into discrete classes in terms of the departure (in units of the standard deviation) from the ensemble mean, with an exceptional performance score. The overall bell-shaped distributions were well predicted and the  $\pm 0.5\sigma$  (or 38%) percentile category competently captured: 38.0% for  $z_0$  and 36.5% for  $d$ . The classification is of particular relevance to

<sup>2</sup> For the sampling with lateral displacement distances ( $d_x$  and  $d_y$ ) smaller than the lateral dimensions ( $l_x$  and  $l_y$ ) of the analysis unit.

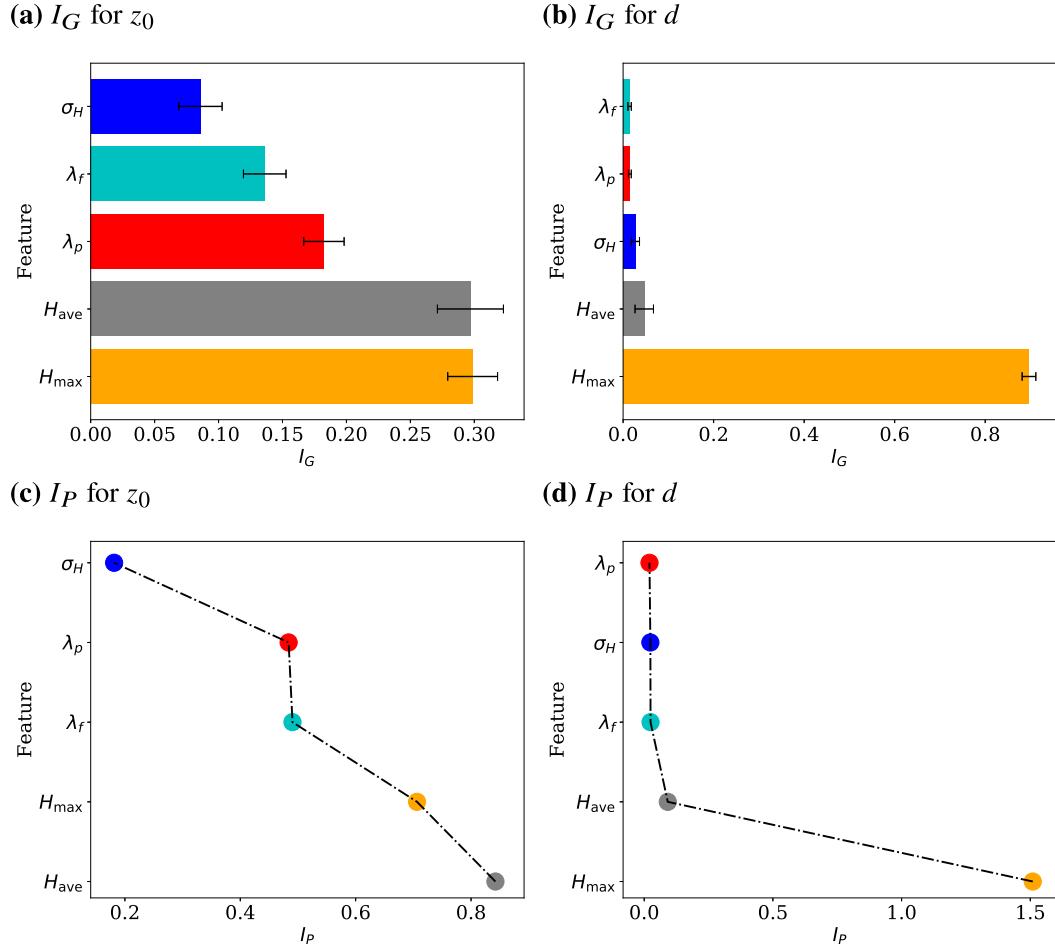


FIG. 11. The (a),(b)  $I_G$  and (c),(d)  $I_P$  scores (definitions see section 4b) for the RF regression for (left)  $z_0$  and (right)  $d$ . Because  $I_P$  is calculated as a relative change of  $R^2$  values before and after the permutation of a feature, it is not necessarily upper bounded by unity. Note the order change of feature labels on the ordinate axes between different panels, especially the top vs the bottom. For comparison, each feature is indicated in a different color.

the parameterization of surface characteristics in mesoscale models, wherein the roughness parameters are generally assigned in accord with land-use categories (Wieringa 1992; Davenport et al. 2000). While the RF regression experienced a significant drop in estimating  $d$  relative to the most recently proposed UCP scheme in Kanda et al. (2013), the performance for  $z_0$  appeared to be promising.

The most practical contribution is the identification of parameters that could be of predominant influence on the prediction of  $z_0$  and  $d$ . The surface morphometric features exhibited greatly differing influence on the roughness length than on the displacement height. For the former, all the morphometric parameters considered herein are found to be informative, doubly so for the building height metrics ( $H_{ave}$  and  $H_{max}$ ), which accounted for up to 60% of the influence. By contrast, the latter,  $d$ , appeared to be almost exclusively determined by  $H_{max}$ . The results coincide well with Kanda et al. (2013), wherein the importance of the maximum building height,  $H_{max}$ , was also emphasized. Direct supporting evidence is the inclusion of  $H_{max}$  in the proposed aerodynamic

parameterization schemes as compared with the original formulas in Macdonald et al. (1998).

A perhaps counterintuitive result is the considerably less striking importance of the building-height variability that reflected by the feature importance scores ( $I_G$  and  $I_P$ ). It is found that  $\sigma_H$  was of less influence as one would normally expect, particularly for the prediction of  $d$  (Fig. 11). While this is not entirely surprising because the values of  $\sigma_H$  for the urban topographies considered herein ( $\sigma_H < 25$  m; see Fig. 3) could be of secondary importance, a more likely explanation would be that  $\sigma_H$  is correlated with the other surface morphometric features, in particular with  $H_{max}$  and  $\lambda_f$  (Hagishima et al. 2009). The building-height-to-canyon-width aspect ratio (AR), which affects the flow regimes inside urban canyons (Ngan and Lo 2016), is another morphometric feature that could be correlated with  $\sigma_H$ , particularly considering the ubiquitous asymmetric canyon geometries in realistic urban regions. The influence of  $\sigma_H$  may be already embodied within those indices; however, disentangling the complexity of the correlation between features is difficult using the current ML algorithm.

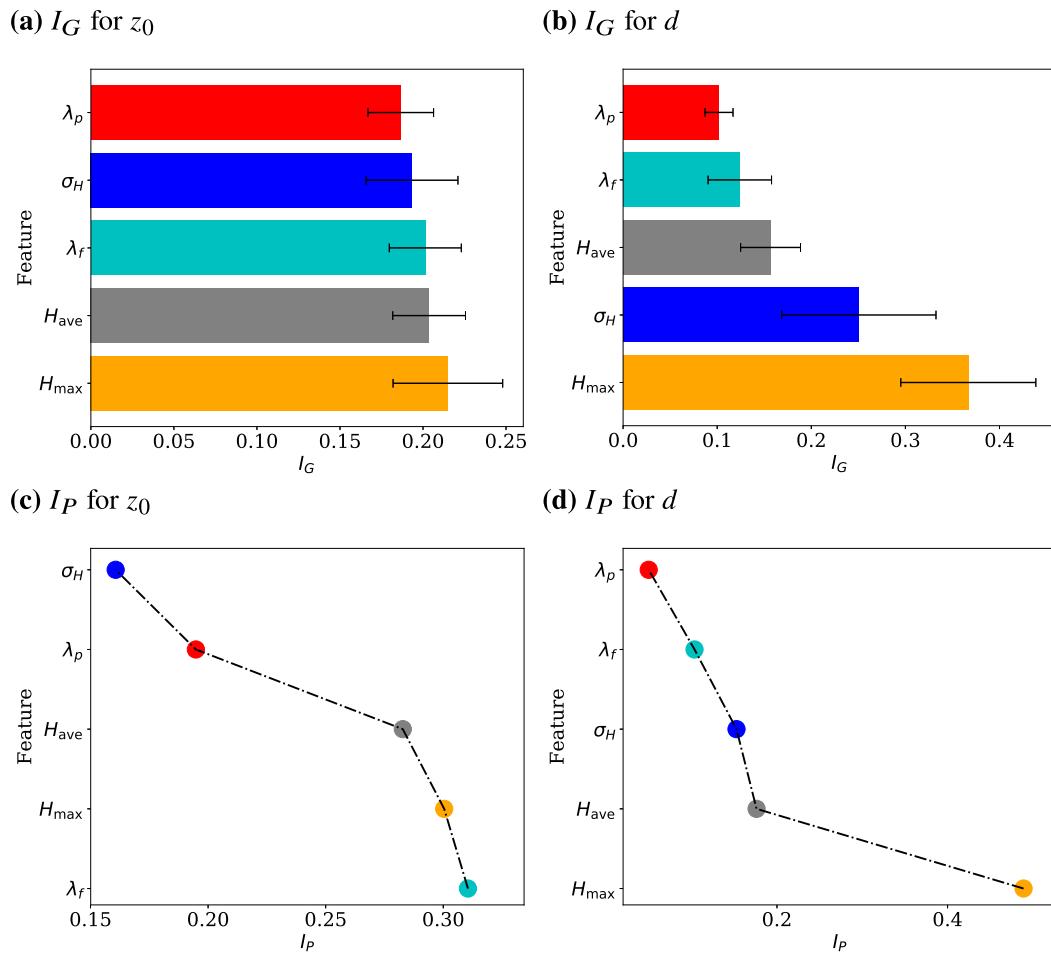


FIG. 12. As in Fig. 11, but for the RF classification.

Nevertheless, knowledge of the feature importance scores allows weights to be assigned accordingly to the different contributing factors, and those scores could be of usefulness guiding the selection of informative factors for the derivation of new UCP schemes.

The weaknesses of the current study lie in the following aspects. First, the urban topographies may not be representative of all urban surface types (Ratti et al. 2002), especially of areas occupied with sparsely distributed roughness elements, despite that the ensemble sampling approach (see section 3a) to a certain extent already accounts for the spatial variations of the morphometric statistics. Green canopies and vegetated surfaces might be another important consideration (Raupach et al. 1996; Kanda et al. 2007; Barbano et al. 2020). Second, the meteorological conditions were idealized, for example, the turbulent flow was driven by a prescribed pressure gradient with periodic boundary conditions and thermal effects were neglected. On the one hand, the estimation of the surface roughness aerodynamic parameters using the conventional micrometeorological method may not be straightforward for thermally stratified boundary layers, partially due to the growing of the internal boundary

layer (Rao 1975; Duan and Ngan 2020). On the other hand, the nonuniversal motions, for example, intermittency in stable conditions (Sun et al. 2012; Ansorge and Mellado 2016) or enhanced vertical transport of scalar covariance in unstable conditions (Cancelli et al. 2012), may result in possible violation against the similarity theory (Mahrt and Bou-Zeid 2020). Oftentimes “near neutral” assumption is imposed in studies of neutrally stratified boundary layer flows (Horiguchi et al. 2010; Klein and Galvez 2015). More sophisticated ML algorithms may be explored, for example, deep learning (e.g., Rasp et al. 2018; Cheng et al. 2019), to account for more complex perturbations that would arise in realistic meteorological scenarios. The above aspects may represent interesting directions for future investigation.

Establishing a universally accessible database to integrate  $z_0$ ,  $d$ , associated urban morphological indices, and possibly available meteorological conditions from discrete simulations, experiments, or field campaigns could bring efforts of different parties together and make well-trained ML algorithms possible. Given the proof-of-principle attempt of the current study, testing a simplified ML algorithm within a mesoscale model as a replacement of the internal UCP schemes could be an

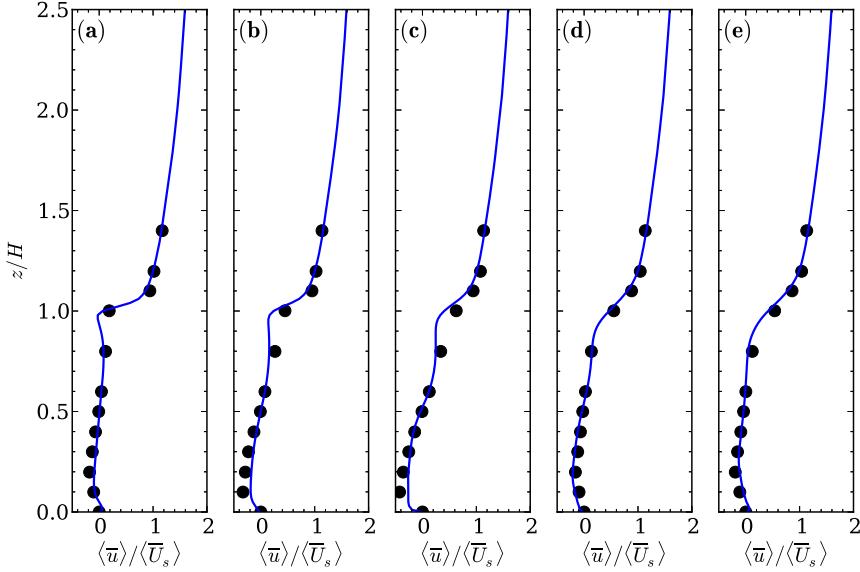


FIG. A1. Normalized mean streamwise velocity profiles,  $\langle \bar{u} \rangle / \langle \bar{U}_s \rangle$ , for the current LES (solid blue curve) and the wind-tunnel experiments of Brown et al. (2000) (black circles) for  $x/W =$  (a)  $-0.4$ , (b)  $-0.25$ , (c)  $0$ , (d)  $0.25$ , and (e)  $0.4$ ;  $\langle \bar{U}_s \rangle$  denotes the shear-layer average,  $1 \leq z/H \leq 1.5$ , of the streamwise velocity.

achievable short-term goal throughout the entire phase toward ML-based UCP.

**Acknowledgments.** This research was supported by the Environment Research and Technology Development Fund (ERTDF) JPMEERF20192005 of the Environmental Restoration and Conservation Agency (ERCA) of Japan and Japan Society for the Promotion of Science (JSPS) Kakenhi 18H01680. The computation was performed on the Large-Scale Computer Systems, Laurel 2, of Kyoto University. The authors thank the editor and the anonymous reviewers for the careful reading and valuable comments.

## APPENDIX A

### Validation

Figure A1 compares the vertical profiles of the temporal and spanwise averages of the streamwise velocity  $\langle \bar{u} \rangle$  at different streamwise locations  $x/W$  inside a unit-aspect-ratio canyon ( $H/W = 1$ , where  $H$  is the canyon height and  $W$  is the canyon width) with the wind-tunnel measurements of Brown et al. (2000). Statistical performance metrics (Eichhorn 2004) indicate that the model simulated the wind profiles very well: the normalized mean-square error (NMSE)  $\approx 0.01\text{--}0.03$  and the fractional bias (FB)  $\approx 0.01\text{--}0.03$  are consistent with the acceptance criteria for microscale flow models (VDI 2005); the hit rate,  $q \approx 68\%$ , also satisfies the criterion ( $q > 66\%$ ) for successful validation (Santiago et al. 2007). Comparable agreements have also been obtained in previous studies using the same LES model (e.g., Duan and Ngan 2018; Lo and Ngan 2020).

The degree of pattern similarity between the simulated and measured wind profiles is quantified using Taylor diagram (Taylor 2001; Ghorbani et al. 2020) in Fig. A2. The markers on the diagram denote the statistical measures of the comparison for each  $x/W$ . A perfect validation would be the one that overlaps the measurement (indicated as “REF” on the  $x$  axis). It can be seen that the simulated patterns of the wind profiles

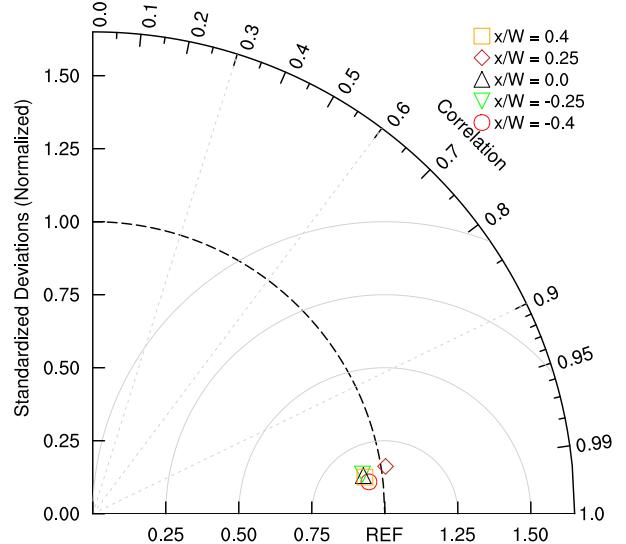


FIG. A2. Taylor diagram for the comparison of temporally and spanwise-averaged streamwise velocity profiles obtained from the current LES model with that from the wind-tunnel experiments of Brown et al. (2000). REF denotes the measured values.

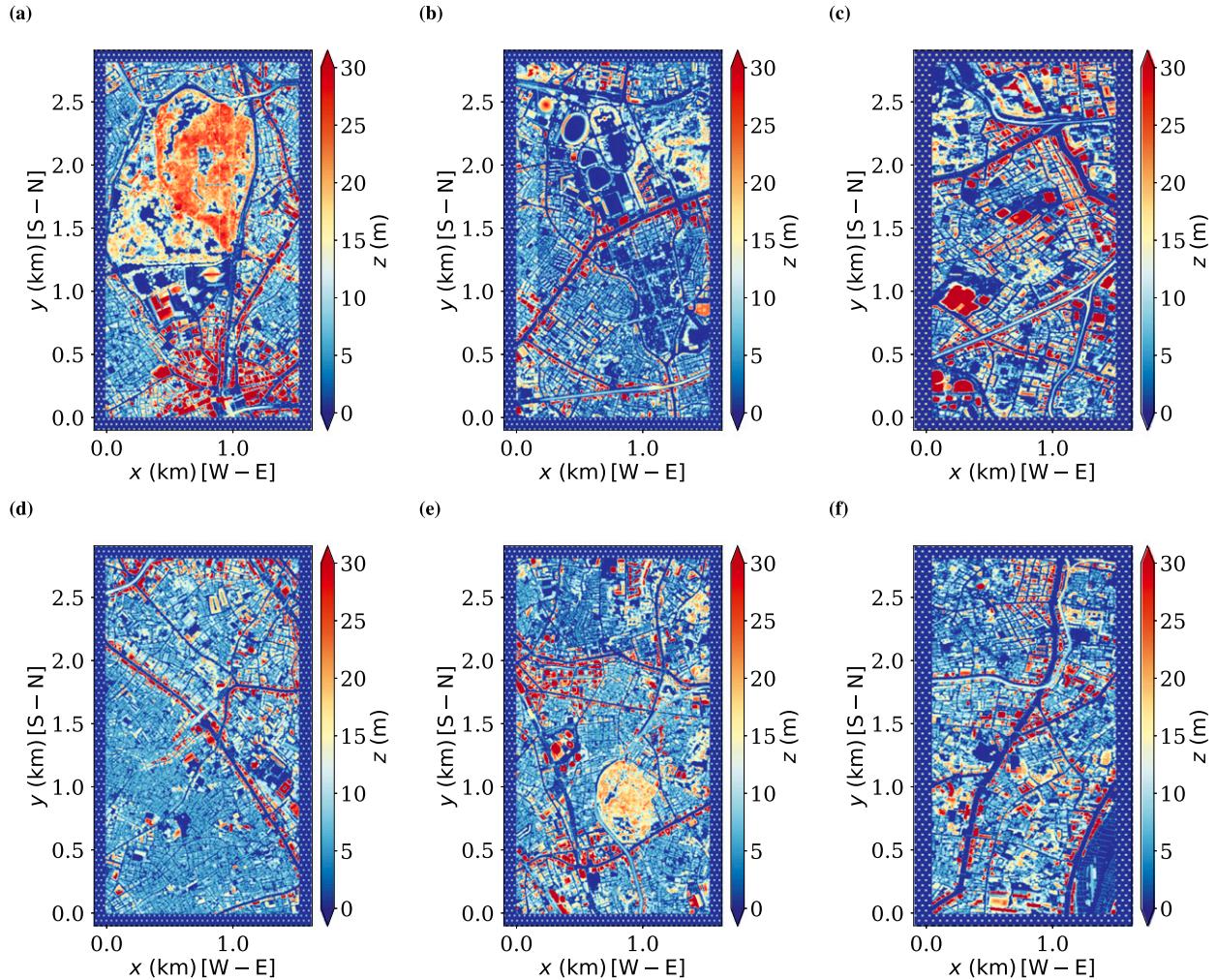


FIG. D1. Urban areas within Tokyo considered for the extra LES in section 5e.

match well with the measurements: the Pearson correlation coefficients (the azimuthal angle) are close to unity; the RMS errors (gray contours) are small, and the standard deviations (the dashed arc) are comparable to the measured values.

Successful validations of second-order turbulence statistics can be found in Duan and Ngan (2019) and Nazarian et al. (2020) for idealized building arrays and in Gronemeier et al. (2021) and Resler et al. (2021) for real-case building setups. Since the dataset in the current study is primarily derived from least squares regression of the time-averaged wind profiles (see sections 3b and 5b), validations of second-order statistics are not repeated here.

## APPENDIX B

### A Simplified Mathematical Illustration of RF

RF has often been presented in a schematic style in its practical applications, while the associated mathematical properties and statistical mechanisms have been largely neglected. This is because internally an ML algorithm can be hardly represented

using a series of formulas—it is an interdisciplinary field and involves linear algebra, multivariate calculus, probability theory and statistics, algorithms, and complex optimization (Mitchell 1997). For completeness and clarity, we show below a simplified mathematical illustration of the procedures that RF uses.

Given a dataset,  $\mathcal{D}_n = \{(X^i, Y^i)\}_{i=1}^n$  [Eq. (12)], where  $(X^i, Y^i)$  denotes an element pair [Eqs. (3) and (4)] of the feature set  $X$  and the response set  $Y$ , that is,  $(X^i, Y^i) \in (X, Y)$ , the objective is to predict  $Y^i$  for given  $X^i$ , namely,

$$\hat{f}(X^i) = \mathbb{E}(Y^i|X^i), \quad (\text{B1})$$

using an RF algorithm that trained on bootstrap subsamples of  $\mathcal{D}_n$ , herein  $\mathcal{D}_n \equiv (X, Y)$ . A collection of the base trees that trained in parallel on  $B$  subsets  $\mathcal{D}_B$  (resulting from a  $B$  times bootstrapping),<sup>B1</sup>

<sup>B1</sup>The process of training the decision trees of a random forest in parallel on random subsets of the training dataset  $\mathcal{D}_n$ .

$$\mathcal{D}_B = \{(X_b, Y_b)\}_{b=1}^B, \quad (\text{B2})$$

wherein  $\mathcal{D}_b = (X_b, Y_b) \subset \mathcal{D}_B$ , constitutes a random forest.

Algorithms for growing the trees generally work from the top down. The best split of each node aims at minimizing the impurity  $G$ , which is a measure conditioned on the feature entries for the optimal partitioning of each node into two such that similar values of the target variable end up in the same set. Exact forms of the impurity function depend on the specific RF tasks. For regression, common measures include minimizing the L1 and L2 errors; for classification, there is Gini impurity (MacKay 2003; D'Ambrosio and Tutore 2011). The optimum split option  $\theta^*$  of a sample set  $Q$  (where  $Q \subseteq \mathcal{D}_b$ ) at node  $m$  is determined such that  $G$  is minimized, namely,

$$\theta^* = \operatorname{argmin}_{\theta} G(Q, \theta). \quad (\text{B3})$$

Recursively doing so for subsets  $Q_{\text{left}}$  and  $Q_{\text{right}}$  (where  $Q = Q_{\text{left}} \cup Q_{\text{right}}$ ) until specified a criterion (e.g., the maximum depth for a tree or the minimum allowable leaf number for a node) is reached. The final prediction of RF for input  $X^i$  is given by averaging over the ensemble of the trees for regression,

$$\hat{f}_{\text{final}}(X^i) = \frac{1}{B} \sum_{b=1}^B \hat{f}_b(X^i), \quad (\text{B4})$$

or voting the majority for classification. This final process is also referred to as aggregating, and jointly with bootstrapping it is known as bagging.

## APPENDIX C

### A Brief Description of the Feature Importance Measures, $I_G$ and $I_P$

For  $I_G$ , a typical choice would be the Gini impurity decrease for RF classification and the variance reduction for RF regression. The former is a function of the probability  $1 - p_i$  that a randomly chosen class is incorrectly categorized; the latter is simply the mean-square error. Formally,

$$I_G = \Delta \begin{cases} \sum_{i=1}^J p_i(1-p_i) & \text{for RF classification} \\ \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 & \text{for RF regression} \end{cases}, \quad (\text{C1})$$

where  $\Delta$  indicates the deduction of associated impurity measure,  $J$  denotes the number of the classification,  $N$  denotes the number of predicted values  $y_i$ , and  $\bar{y}$  denotes the associated mean. For  $I_P$ , there is

$$I_P = \sum_{b=1}^B \frac{P(X_b, Y_b) - P[\mathbf{S}(X_b)_j, Y_b]}{P(X_b, Y_b)}, \quad (\text{C2})$$

where  $P$  denotes the model performance score,  $\mathbf{S}(X_b)_j$  is the  $j$ th feature with its element values randomly shuffled, and  $B$  is the total number of training subsets drawn out of a random split of the original database (see appendix B for details).

For simplicity, herein  $P$  is chosen to be the  $R$  squared for the RF regression and the proportion of correctly classified counts for the RF classification.

## APPENDIX D

### Subregions within Tokyo

[Figure D1](#) plots the topographies of six urban areas within Tokyo considered for the extra LES in [section 5e](#). The spatial dimensions and model setups remain the same as that for the Osaka subregions plotted in [Fig. 1](#).

## REFERENCES

- Akinlade, O. G., D. J. Bergstrom, M. F. Tachie, and L. Castillo, 2004: Outer flow scaling of smooth and rough wall turbulent boundary layers. *Exp. Fluids*, **37**, 604–612, <https://doi.org/10.1007/s00348-004-0856-5>.
- Ansorge, C., and J. P. Mellado, 2016: Analyses of external and global intermittency in the logarithmic layer of Ekman flow. *J. Fluid Mech.*, **805**, 611–635, <https://doi.org/10.1017/jfm.2016.534>.
- Barbano, F., S. Di Sabatino, R. Stoll, and E. R. Pardyjak, 2020: A numerical study of the impact of vegetation on mean and turbulence fields in a European-city neighbourhood. *Build. Environ.*, **186**, 107293, <https://doi.org/10.1016/j.buildenv.2020.107293>.
- Bardina, J., J. Ferziger, and W. Reynolds, 1980: Improved subgrid-scale models for large-eddy simulation. *13th Fluid and Plasma Dynamics Conf.*, Snowmass, CO, AIAA, <https://doi.org/10.2514/6.1980-1357>.
- Biesbroek, R., S. Badloe, and I. N. Athanasiadis, 2020: Machine learning for research on climate change adaptation policy integration: An exploratory UK case study. *Reg. Environ. Change*, **20**, 85, <https://doi.org/10.1007/s10113-020-01677-8>.
- Blocken, B., T. Stathopoulos, and J. Carmeliet, 2007: CFD simulation of the atmospheric boundary layer: Wall function problems. *Atmos. Environ.*, **41**, 238–252, <https://doi.org/10.1016/j.atmosenv.2006.08.019>.
- Bodini, N., J. K. Lundquist, and M. Optis, 2020: Can machine learning improve the model representation of turbulent kinetic energy dissipation rate in the boundary layer for complex terrain? *Geosci. Model Dev.*, **13**, 4271–4285, <https://doi.org/10.5194/gmd-13-4271-2020>.
- Bou-Zeid, E., W. Anderson, G. G. Katul, and L. Mahrt, 2020: The persistent challenge of surface heterogeneity in boundary-layer meteorology: A review. *Bound.-Layer Meteor.*, **177**, 227–245, <https://doi.org/10.1007/s10546-020-00551-8>.
- Breiman, L., 1996: Bagging predictors. *Mach. Learn.*, **24**, 123–140, <https://doi.org/10.1007/BF00058655>.
- , 2001: Random forests. *Mach. Learn.*, **45**, 5–32, <https://doi.org/10.1023/A:1010933404324>.
- Britter, R. E., and S. R. Hanna, 2003: Flow and dispersion in urban areas. *Annu. Rev. Fluid Mech.*, **35**, 469–496, <https://doi.org/10.1146/annurev.fluid.35.101101.161147>.
- Brown, M. J., 2000: Urban parameterizations for mesoscale meteorological models. *Mesoscale Atmospheric Dispersion*, WIT Press, 193–255.
- , R. Lawson, D. DeCroix, and R. Lee, 2000: Mean flow and turbulence measurements around a 2-D array of buildings in a

- wind tunnel. *11th Joint Conf. on the Applications of Air Pollution Meteorology*, Long Beach, CA, Amer. Meteor. Soc., 4A.2, [https://ams.confex.com/ams/annual2000/techprogram/paper\\_314.htm](https://ams.confex.com/ams/annual2000/techprogram/paper_314.htm).
- Buccolieri, R., M. Sandberg, H. Wigö, and S. Di Sabatino, 2019: The drag force distribution within regular arrays of cubes and its relation to cross ventilation—Theoretical and experimental analyses. *J. Wind Eng. Ind. Aerodyn.*, **189**, 91–103, <https://doi.org/10.1016/j.jweia.2019.03.022>.
- Cancelli, D. M., N. L. Dias, and M. Chamecki, 2012: Dimensionless criteria for the production-dissipation equilibrium of scalar fluctuations and their implications for scalar similarity. *Water Resour. Res.*, **48**, W10522, <https://doi.org/10.1029/2012WR012127>.
- Cheng, H., and I. Castro, 2002: Near wall flow over urban-like roughness. *Bound.-Layer Meteor.*, **104**, 229–259, <https://doi.org/10.1023/A:1016060103448>.
- , P. Hayden, A. Robins, and I. Castro, 2007: Flow over cube arrays of different packing densities. *J. Wind Eng. Ind. Aerodyn.*, **95**, 715–740, <https://doi.org/10.1016/j.jweia.2007.01.004>.
- Cheng, W., and C.-H. Liu, 2011: Large-eddy simulation of turbulent transports in urban street canyons in different thermal stabilities. *J. Wind Eng. Ind. Aerodyn.*, **99**, 434–442, <https://doi.org/10.1016/j.jweia.2010.12.009>.
- Cheng, Y., and Coauthors, 2019: Deep learning for subgrid-scale turbulence modeling in large-eddy simulations of the atmospheric boundary layer. arXiv, 33 pp., <https://arxiv.org/ftp/arxiv/papers/1910/1910.12125.pdf>.
- , Q. Li, A. Grachev, S. Argentini, H. J. S. Fernando, and P. Gentine, 2020: Power-law scaling of turbulence spectra for the stably stratified atmospheric boundary layer. *Bound.-Layer Meteor.*, **177**, 1–18, <https://doi.org/10.1007/s10546-020-00545-6>.
- Chew, L. W., X. Liu, X.-X. Li, and L. K. Norford, 2020: Interaction between heat wave and urban heat island: A case study in a tropical coastal city, Singapore. *Atmos. Res.*, **247**, 105134, <https://doi.org/10.1016/j.atmosres.2020.105134>.
- Cui, Z., X. Cai, and C. J. Baker, 2004: Large-eddy simulation of turbulent flow in a street canyon. *Quart. J. Roy. Meteor. Soc.*, **130**, 1373–1394, <https://doi.org/10.1256/qj.02.150>.
- D'Ambrosio, A., and V. A. Tutore, 2011: Conditional classification trees by weighting the Gini impurity measure. *New Perspectives in Statistical Modeling and Data Analysis*, S. Ingrassia, R. Rocci, and M. Vichi, Eds., Springer-Verlag, 273–280.
- Davenport, A. G., C. S. B. Grimmond, T. R. Oke, and J. Wieringa, 2000: Estimating the roughness of cities and sheltered country. *12th Conf. on Applied Climatology*, Asheville, NC, Amer. Meteor. Soc., 4B.2, [https://ams.confex.com/ams/May2000/techprogram/paper\\_13744.htm](https://ams.confex.com/ams/May2000/techprogram/paper_13744.htm).
- Deardorff, J., 1980: Stratocumulus-capped mixed layers derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495–527, <https://doi.org/10.1007/BF00119502>.
- Duan, G., and K. Ngan, 2018: Effects of time-dependent inflow perturbations on turbulent flow in a street canyon. *Bound.-Layer Meteor.*, **167**, 257–284, <https://doi.org/10.1007/s10546-017-0327-1>.
- , and —, 2019: Sensitivity of turbulent flow around a 3-D building array to urban boundary-layer stability. *J. Wind Eng. Ind. Aerodyn.*, **193**, 103958, <https://doi.org/10.1016/j.jweia.2019.103958>.
- , and —, 2020: Influence of thermal stability on the ventilation of a 3-D building array. *Build. Environ.*, **183**, 106969, <https://doi.org/10.1016/j.buildenv.2020.106969>.
- , and T. Takemi, 2021: Gustiness in thermally-stratified urban turbulent boundary-layer flows and the influence of surface roughness. *J. Wind Eng. Ind. Aerodyn.*, **208**, 10442, <https://doi.org/10.1016/j.jweia.2020.104442>.
- Edwards, J. M., A. C. M. Beljaars, A. A. M. Holtlag, and A. P. Lock, 2020: Representation of boundary-layer processes in numerical weather prediction and climate models. *Bound.-Layer Meteor.*, **177**, 511–539, <https://doi.org/10.1007/s10546-020-00530-z>.
- Eichhorn, J., 2004: Application of a new evaluation guideline for microscale flow models. *Ninth Int. Conf. on Harmonisation within Atmospheric Dispersion Modeling for Regulatory Purposes*, Garmisch-Partenkirchen, Germany, HARMO, [http://www.harmo.org/Conferences/Proceedings/\\_Garmisch/publishedSections/PPT/1.10-Eichhorn.pdf](http://www.harmo.org/Conferences/Proceedings/_Garmisch/publishedSections/PPT/1.10-Eichhorn.pdf).
- Fang, J., and F. Porté-Agel, 2015: Large-eddy simulation of very-large-scale motions in the neutrally stratified atmospheric boundary layer. *Bound.-Layer Meteor.*, **155**, 397–416, <https://doi.org/10.1007/s10546-015-0006-z>.
- Foken, T., 2006: 50 years of the Monin–Obukhov similarity theory. *Bound.-Layer Meteor.*, **119**, 431–447, <https://doi.org/10.1007/s10546-006-9048-6>.
- Gagne, D. J., II, H. M. Christensen, A. C. Subramanian, and A. H. Monahan, 2020: Machine learning for stochastic parameterization: Generative adversarial networks in the Lorenz '96 model. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001896, <https://doi.org/10.1029/2019MS001896>.
- Ghorbani, M. A., R. C. Deo, S. Kim, M. Hasanzadeh Kashani, V. Karimi, and M. Izadkhah, 2020: Development and evaluation of the cascade correlation neural network and the random forest models for river stage and river flow prediction in Australia. *Soft Comput.*, **24**, 12 079–12 090, <https://doi.org/10.1007/s00500-019-04648-2>.
- Grimmond, C. S. B., and T. R. Oke, 1999: Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Meteor.*, **38**, 1262–1292, [https://doi.org/10.1175/1520-0450\(1999\)038<1262:APOUAD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2).
- Gronemeier, T., K. Surm, F. Harms, B. Leitl, B. Maronga, and S. Raasch, 2021: Evaluation of the dynamic core of the PALM model system 6.0 in a neutrally stratified urban environment: Comparison between LES and wind-tunnel experiments. *Geosci. Model Dev.*, **14**, 3317–3333, <https://doi.org/10.5194/gmd-14-3317-2021>.
- Hagishima, A., J. Tanimoto, K. Nagayama, and S. Meno, 2009: Aerodynamic parameters of regular arrays of rectangular blocks with various geometries. *Bound.-Layer Meteor.*, **132**, 315–337, <https://doi.org/10.1007/s10546-009-9403-5>.
- Hanafusa, T., T. Fujitani, Y. Kobori, and Y. Mitsuta, 1982: A new type sonic anemometer-thermometer for field operation. *Pap. Meteor. Geophys.*, **33**, 1–19, <https://doi.org/10.2467/mripapers.33.1>.
- Hanna, S. R., and Coauthors, 2006: Detailed simulations of atmospheric flow and dispersion in downtown Manhattan: An application of five computational fluid dynamics models. *Bull. Amer. Meteor. Soc.*, **87**, 1713–1726, <https://doi.org/10.1175/BAMS-87-12-1713>.
- Haugen, D., J. Kaimal, and E. Bradley, 1971: An experimental study of Reynolds stress and heat flux in the atmospheric surface layer. *Quart. J. Roy. Meteor. Soc.*, **97**, 168–180, <https://doi.org/10.1002/qj.49709741204>.
- Haupt, S. E., and Coauthors, 2020: On bridging a modeling scale gap: Mesoscale to microscale coupling for wind energy. *Bull. Amer. Meteor. Soc.*, **100**, 2533–2550, <https://doi.org/10.1175/BAMS-D-18-0033.1>.
- Horiguchi, M., T. Hayashi, H. Hashiguchi, Y. Ito, and H. Ueda, 2010: Observations of coherent turbulence structures in the

- near-neutral atmospheric boundary layer. *Bound.-Layer Meteor.*, **136**, 25–44, <https://doi.org/10.1007/s10546-010-9500-5>.
- Hutchins, N., and I. Marusic, 2007: Evidence of very long meandering features in the logarithmic region of turbulent boundary layers. *J. Fluid Mech.*, **579**, 1–28, <https://doi.org/10.1017/S0022112006003946>.
- Inagaki, A., and M. Kanda, 2008: Turbulent flow similarity over an array of cubes in near-neutrally stratified atmospheric flow. *J. Fluid Mech.*, **615**, 101–120, <https://doi.org/10.1017/S0022112008003765>.
- Kaimal, J. C., and J. A. Businger, 1963: A continuous wave sonic anemometer-thermometer. *J. Appl. Meteor.*, **2**, 156–164, [https://doi.org/10.1175/1520-0450\(1963\)002<0156:ACWSAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1963)002<0156:ACWSAT>2.0.CO;2).
- , —, and J. C. Wyngaard, 1990: The Kansas and Minnesota experiments. *Bound.-Layer Meteor.*, **50**, 31–47, <https://doi.org/10.1007/BF00120517>.
- , —, D. Haugen, O. Coté, Y. Izumi, S. Caughey, and C. Readings, 1976: Turbulence structure in the convective boundary layer. *J. Atmos. Sci.*, **33**, 2152–2169, [https://doi.org/10.1175/1520-0469\(1976\)033<2152:TSITCB>2.0.CO;2](https://doi.org/10.1175/1520-0469(1976)033<2152:TSITCB>2.0.CO;2).
- Kanda, M., M. Kanega, T. Kawai, R. Moriwaki, and H. Sugawara, 2007: Roughness lengths for momentum and heat derived from outdoor urban scale models. *J. Appl. Meteor. Climatol.*, **46**, 1067–1079, <https://doi.org/10.1175/JAM2500.1>.
- , A. Inagaki, T. Miyamoto, M. Gryschka, and S. Raasch, 2013: A new aerodynamic parametrization for real urban surfaces. *Bound.-Layer Meteor.*, **148**, 357–377, <https://doi.org/10.1007/s10546-013-9818-x>.
- Kastner-Klein, P., and M. W. Rotach, 2004: Mean flow and turbulence characteristics in an urban roughness sublayer. *Bound.-Layer Meteor.*, **111**, 55–84, <https://doi.org/10.1023/B:BOUN.0000010994.32240.b1>.
- Klein, P. M., and J. M. Galvez, 2015: Flow and turbulence characteristics in a suburban street canyon. *Environ. Fluid Mech.*, **15**, 419–438, <https://doi.org/10.1007/s10652-014-9352-5>.
- Lellep, M., J. Prexl, M. Linkmann, and B. Eckhardt, 2020: Using machine learning to predict extreme events in the Hénon map. *Chaos*, **30**, 013113, <https://doi.org/10.1063/1.5121844>.
- Li, B., X. Zhang, and X. Zhang, 2020: Classifying wakes produced by self-propelled fish-like swimmers using neural networks. *Theor. Appl. Mech. Lett.*, **10**, 149–154, <https://doi.org/10.1016/j.taml.2020.01.010>.
- Lo, K. W., and K. Ngan, 2017: Characterizing ventilation and exposure in street canyons using Lagrangian particles. *J. Appl. Meteor. Climatol.*, **56**, 1177–1194, <https://doi.org/10.1175/JAMC-D-16-0168.1>.
- , —, and —, 2020: Multiscale parameterisation of passive scalars via wavelet-based numerical homogenisation. *Appl. Math. Model.*, **82**, 217–234, <https://doi.org/10.1016/j.apm.2020.01.018>.
- Macdonald, R., R. Griffiths, and D. Hall, 1998: An improved method for the estimation of surface roughness of obstacle arrays. *Atmos. Environ.*, **32**, 1857–1864, [https://doi.org/10.1016/S1352-2310\(97\)00403-2](https://doi.org/10.1016/S1352-2310(97)00403-2).
- MacKay, D. J. C., 2003: *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 628 pp.
- Mahrt, L., and E. Bou-Zeid, 2020: Non-stationary boundary layers. *Bound.-Layer Meteor.*, **177**, 189–204, <https://doi.org/10.1007/s10546-020-00533-w>.
- Maronga, B., and Coauthors, 2020a: Overview of the PALM model system 6.0. *Geosci. Model Dev.*, **13**, 1335–1372, <https://doi.org/10.5194/gmd-13-1335-2020>.
- , C. Knigge, and S. Raasch, 2020b: An improved surface boundary condition for large-eddy simulations based on Monin–Obukhov similarity theory: Evaluation and consequences for grid convergence in neutral and stable conditions. *Bound.-Layer Meteor.*, **174**, 297–325, <https://doi.org/10.1007/s10546-019-00485-w>.
- Martilli, A., A. Clappier, and M. W. Rotach, 2002: An urban surface exchange parameterisation for mesoscale models. *Bound.-Layer Meteor.*, **104**, 261–304, <https://doi.org/10.1023/A:1016099921195>.
- Menze, B. H., B. M. Kelm, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich, and F. A. Hamprecht, 2009: A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinf.*, **10**, 213, <https://doi.org/10.1186/1471-2105-10-213>.
- Michioka, T., and A. Sato, 2012: Effect of incoming turbulent structure on pollutant removal from two-dimensional street canyon. *Bound.-Layer Meteor.*, **145**, 469–484, <https://doi.org/10.1007/s10546-012-9733-6>.
- , H. Takimoto, H. Ono, and A. Sato, 2019: Large-eddy simulation of the effects of wind-direction fluctuations on turbulent flow and gas dispersion within a cubical canopy. *Bound.-Layer Meteor.*, **173**, 243–262, <https://doi.org/10.1007/s10546-019-00467-y>.
- Mirocha, J., B. Kosović, and G. Kirkil, 2014: Resolved turbulence characteristics in large-eddy simulations nested within mesoscale simulations using the Weather Research and Forecasting Model. *Mon. Wea. Rev.*, **142**, 806–831, <https://doi.org/10.1175/MWR-D-13-00064.1>.
- Mitchell, T. M., 1997: *Machine Learning*. McGraw-Hill, 414 pp.
- Monin, A. S., and A. M. Obukhov, 1954: Basic laws of turbulent mixing in the atmosphere near the ground. *Tr. Geofiz. Inst. Akad. Nauk SSSR*, **24**, 163–187.
- Muñoz-Esparza, D., B. Kosović, J. Mirocha, and J. van Beeck, 2014: Bridging the transition from mesoscale to microscale turbulence in numerical weather prediction models. *Bound.-Layer Meteor.*, **153**, 409–440, <https://doi.org/10.1007/s10546-014-9956-9>.
- Munters, W., C. Meneveau, and J. Meyers, 2016: Shifted periodic boundary conditions for simulations of wall-bounded turbulent flows. *Phys. Fluids*, **28**, 025112, <https://doi.org/10.1063/1.4941912>.
- Nakayama, H., T. Takemi, and H. Nagai, 2011: LES analysis of the aerodynamic surface properties for turbulent flows over building arrays with various geometries. *J. Appl. Meteor. Climatol.*, **50**, 1692–1712, <https://doi.org/10.1175/2011JAMC2567.1>.
- Nazarian, N., E. S. Krayenhoff, and A. Martilli, 2020: A one-dimensional model of turbulent flow through “urban” canopies (MLUCM v2.0): Updates based on large-eddy simulation. *Geosci. Model Dev.*, **13**, 937–953, <https://doi.org/10.5194/gmd-13-937-2020>.
- Ngan, K., and K. W. Lo, 2016: Revisiting the flow regimes for urban street canyons using the numerical Green’s function. *Environ. Fluid Mech.*, **16**, 313–334, <https://doi.org/10.1007/s10652-015-9422-3>.
- Oikawa, S., and Y. Meng, 1995: Turbulence characteristics and organized motion in a suburban roughness sublayer. *Bound.-Layer Meteor.*, **74**, 289–312, <https://doi.org/10.1007/BF00712122>.
- Olesen, H. R., S. E. Larsen, and J. Højstrup, 1984: Modelling velocity spectra in the lower part of the planetary boundary layer. *Bound.-Layer Meteor.*, **29**, 285–312, <https://doi.org/10.1007/BF00119794>.

- Park, S.-B., J.-J. Baik, and B.-S. Han, 2015a: Large-eddy simulation of turbulent flow in a densely built-up urban area. *Environ. Fluid Mech.*, **15**, 235–250, <https://doi.org/10.1007/s10652-013-9306-3>.
- , —, and S.-H. Lee, 2015b: Impacts of mesoscale wind on turbulent flow and ventilation in a densely built-up urban area. *J. Appl. Meteor. Climatol.*, **54**, 811–824, <https://doi.org/10.1175/JAMC-D-14-0044.1>.
- Rao, K. S., 1975: Effect of thermal stratification on the growth of the internal boundary layer. *Bound.-Layer Meteor.*, **8**, 227–234, <https://doi.org/10.1007/BF00241339>.
- Rasp, S., M. S. Pritchard, and P. Gentine, 2018: Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci. USA*, **115**, 9684–9689, <https://doi.org/10.1073/pnas.1810286115>.
- Ratti, C., S. Di Sabatino, R. Britter, M. Brown, F. Caton, and S. Burian, 2002: Analysis of 3-D urban databases with respect to pollution dispersion for a number of European and American cities. *Water Air Soil Pollut. Focus*, **2**, 459–469, <https://doi.org/10.1023/A:1021380611553>.
- Raupach, M. R., J. J. Finnigan, and Y. Brunei, 1996: Coherent eddies and turbulence in vegetation canopies: The mixing-layer analogy. *Bound.-Layer Meteor.*, **78**, 351–382, <https://doi.org/10.1007/BF00120941>.
- Resler, J., and Coauthors, 2021: Validation of the PALM model system 6.0 in real urban environment; case study of Prague-Dejvice, Czech Republic. *Geosci. Model Dev.*, <https://doi.org/10.5194/gmd-2020-175>, in press.
- Rodriguez-Galiano, V. F., M. Sanchez-Castillo, J. Dash, P. M. Atkinson, and J. Ojeda-Zujar, 2016: Modelling interannual variation in the spring and autumn land surface phenology of the European forest. *Biogeosciences*, **13**, 3305–3317, <https://doi.org/10.5194/bg-13-3305-2016>.
- Rotach, M. W., 1993a: Turbulence close to a rough urban surface Part I: Reynolds stress. *Bound.-Layer Meteor.*, **65**, 1–28, <https://doi.org/10.1007/BF00708816>.
- , 1993b: Turbulence close to a rough urban surface Part II: Variances and gradients. *Bound.-Layer Meteor.*, **66**, 75–92, <https://doi.org/10.1007/BF00705460>.
- , 1995: Profiles of turbulence statistics in and above an urban street canyon. *Atmos. Environ.*, **29**, 1473–1486, [https://doi.org/10.1016/1352-2310\(95\)00084-C](https://doi.org/10.1016/1352-2310(95)00084-C).
- , 1999: On the influence of the urban roughness sublayer on turbulence and dispersion. *Atmos. Environ.*, **33**, 4001–4008, [https://doi.org/10.1016/S1352-2310\(99\)00141-7](https://doi.org/10.1016/S1352-2310(99)00141-7).
- Roth, M., 2000: Review of atmospheric turbulence over cities. *Quart. J. Roy. Meteor. Soc.*, **126**, 941–990, <https://doi.org/10.1002/qj.49712656409>.
- Santiago, J., A. Martilli, and F. Martín, 2007: CFD simulation of airflow over a regular array of cubes. Part I: Three-dimensional simulation of the flow and validation with wind-tunnel measurements. *Bound.-Layer Meteor.*, **122**, 609–634, <https://doi.org/10.1007/s10546-006-9123-z>.
- Shen, C., and Coauthors, 2019: Impacts of high-resolution urban canopy parameters within the WRF Model on dynamical and thermal fields over Guangzhou, China. *J. Appl. Meteor. Climatol.*, **58**, 1155–1176, <https://doi.org/10.1175/JAMC-D-18-0114.1>.
- Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, **91**, 99–164, [https://doi.org/10.1175/1520-0493\(1963\)091<0099:GCEWTP>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2).
- Sun, J., L. Mahrt, R. M. Banta, and Y. L. Pichugina, 2012: Turbulence regimes and turbulence intermittency in the stable boundary layer during CASES-99. *J. Atmos. Sci.*, **69**, 338–351, <https://doi.org/10.1175/JAS-D-11-082.1>.
- Takemi, T., T. Yoshida, M. Horiguchi, and W. Vanderbauwhede, 2020: Large-eddy-simulation analysis of airflows and strong wind hazards in urban areas. *Urban Climate*, **32**, 100625, <https://doi.org/10.1016/j.ulclim.2020.100625>.
- Tang, Y., J. Kurths, W. Lin, E. Ott, and L. Kocarev, 2020: Introduction to focus issue: When machine learning meets complex systems: Networks, chaos, and nonlinear dynamics. *Chaos*, **30**, 063151, <https://doi.org/10.1063/5.0016505>.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192, <https://doi.org/10.1029/2000JD900719>.
- VDI, 2005: Environmental meteorology-prognostic microscale windfield models—Evaluation for flow around buildings and obstacles. VDI Standard 3783, 68 pp.
- Wang, C., and W. Anderson, 2019: Turbulence coherence within canonical and realistic aeolian dune-field roughness sublayers. *Bound.-Layer Meteor.*, **173**, 409–434, <https://doi.org/10.1007/s10546-019-00477-w>.
- Wang, W., and E. Ng, 2018: Air ventilation assessment under unstable atmospheric stratification—A comparative study for Hong Kong. *Build. Environ.*, **130**, 1–13, <https://doi.org/10.1016/j.buildenv.2017.12.018>.
- , Y. Xu, E. Ng, and S. Raasch, 2018: Evaluation of satellite-derived building height extraction by CFD simulations: A case study of neighborhood-scale ventilation in Hong Kong. *Landscape Urban Plann.*, **170**, 90–102, <https://doi.org/10.1016/j.landurbplan.2017.11.008>.
- Were, K., D. T. Bui, Ø. B. Dick, and B. R. Singh, 2015: A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecol. Indic.*, **52**, 394–403, <https://doi.org/10.1016/j.ecolind.2014.12.028>.
- Wicker, L. J., and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097, [https://doi.org/10.1175/1520-0493\(2002\)130<2088:TSMFEM>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2).
- Wieringa, J., 1992: Updating the davenport roughness classification. *J. Wind Eng. Ind. Aerodyn.*, **41**, 357–368, [https://doi.org/10.1016/0167-6105\(92\)90434-C](https://doi.org/10.1016/0167-6105(92)90434-C).
- Williamson, J., 1980: Low-storage Runge-Kutta schemes. *J. Comput. Phys.*, **35**, 48–56, [https://doi.org/10.1016/0021-9991\(80\)90033-9](https://doi.org/10.1016/0021-9991(80)90033-9).
- Yoshida, T., T. Takemi, and M. Horiguchi, 2018: Large-eddy-simulation study of the effects of building-height variability on turbulent flows over an actual urban area. *Bound.-Layer Meteor.*, **168**, 127–153, <https://doi.org/10.1007/s10546-018-0344-8>.
- Zitnik, M., F. Nguyen, B. Wang, J. Leskovec, A. Goldenberg, and M. M. Hoffman, 2019: Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. *Inf. Fusion*, **50**, 71–91, <https://doi.org/10.1016/j.inffus.2018.09.012>.