Multifidelity modeling and uncertainty quantification of heterogeneous roughness

By E. Y. Shin†, M. Chan, J. Wang, T. Zahtila, C. Gorle, G. Iaccarino AND M. F. Howland†

Atmosphere-surface interactions are complex and multiscale in nature. Therefore, numerical models of the high-Reynolds-number atmosphere do not resolve but instead model these interactions using surface parameterizations. To make predictions, the unknown parameters inherent to these surface models have been traditionally estimated deterministically, based on morphometric (geometry-based) or anemometric approaches. Yet these deterministic methods do not address the uncertainty in parameters that is inherent to parameterizing the turbulence processes, nor do they enable the assimilation of newly available data that capture the turbulence processes. In this study, we develop a data-driven approach to quantify uncertainty in the parameterization of surface roughness and we apply the methodology to a roughness transition problem. We learn the uncertainties in surface parameterizations within both mid-fidelity [wall-modeled large-eddy simulations (LES)] and low-fidelity [Reynolds-averaged Navier-Stokes (RANS)] models by leveraging geometry-resolving high-fidelity LES data. To enable tractable parameter estimation for computational fluid dynamics (CFD) models, we utilize a machine learning–accelerated inverse uncertainty quantification (UQ) approach that reduces the required model evaluations by three orders of magnitude. The UQ method reduces error in the modeled boundary-layer wind speed observed in the UQ by 71.7% and 75.0% compared to a standard morphometric model for LES and RANS, respectively. However, the reduction in mean predictive error for unseen statistics is not significant, suggesting opportunity to reduce model form error in this heterogeneous roughness transition environment. The choice of statistics assimilated in the inverse UQ method for this roughness transition problem affects both the mean and the uncertainty of the surface parameters.

1. Introduction

Accurately modeling the atmospheric boundary layer (ABL) is essential for applications, including numerical weather prediction (Cohen *et al.* 2015), climate modeling (Ma *et al.* 2022), and modeling urban environments (Bou-Zeid *et al.* 2009). The predictions in these applications serve as inputs for everyday decision-making. The surface boundary condition in the ABL dictates the character of the turbulence and the flow dynamics (Stull 2012). Therefore, accurate representation of the surface is critical to produce accurate ABL predictions (Bou-Zeid *et al.* 2020).

In high-Reynolds-number ABL flow over the multitude of natural and built surface geometries that exist in the ABL, multiscale turbulence processes dictate momentum and heat flux. Thus, the surface conditions are difficult to resolve (Bou-Zeid *et al.* 2020). Instead, numerical models parameterize the effects of these surface processes. Historically, the input parameters to these parameterizations, such as the aerodynamic roughness

† Department of Civil and Environmental Engineering, Massachusetts Institute of Technology

length z_0, have been estimated deterministically by morphometric (Macdonald *et al.* 1998; Grimmond & Oke 1999) or anemometric (Li *et al.* 2020) approaches. These methods utilize simplified geometric models and limited wind tunnel data or simulation data to tune the parameters. These morphometric, geometry-based methods produce deterministic parameters, leading to deterministic predictions of surface fluxes. Further, there is uncertainty about whether the equilibrium assumption applied in many of the commonly used surface parameterizations remains valid over heterogeneous surface conditions within the ABL (Bou-Zeid *et al.* 2020). Previous studies indicate sensitivity of parameters to sampling regions (Li *et al.* 2020), further emphasizing the uncertainty inherent in this deterministic estimation process of parameters. The existing, deterministic approaches to estimating surface model parameters do not recognize the uncertainty associated with parameterizing the turbulence processes nor are they flexible to assimilate newly available data that may be used to more broadly explore the parameter space.

The quantification of parameter uncertainty is computationally expensive, typically requiring $\mathcal{O}(10^5)$ forward evaluations of the model of interest for accurate sampling of the parameter posterior. This cost renders inverse UQ intractable for computational fluids applications. Therefore, the present study leverages a UQ methodology in which the number of forward model evaluations required is reduced 1000-fold (Howland *et al.* 2022; Oliver *et al.* 2024). This machine learning–accelerated methodology efficiently learns surface parameters and their uncertainties, based on high-fidelity geometry-resolving LES data. We apply this approach to a problem of high-Reynolds-number boundary-layer flow in the transition from relatively low surface roughness to an idealized urban canopy, considering multiple model fidelities. First, the inverse UQ method is verified for both mid-fidelity wall-modeled LES (WMLES) and low-fidelity RANS surface models in perfect-model numerical experiments. Then, UQ of surface model parameters in the lower-fidelity models are performed, leveraging the turbulence-capturing high-fidelity simulation data. The joint distributions over the surface parameters are quantified, and we evaluate the predictions from the two modeling fidelities with credible intervals generated by sampling from the parameter posteriors. Further, we evaluate the accuracy of the UQ approach against conventional morphometric parameter estimates and the degree to which the posteriors depend on the assimilated statistics.

2. Methods

In computational models of high-Reynolds-number flow, complex surface geometries can represented using the immersed boundary method (Iaccarino & Verzicco 2003; Bou-Zeid *et al.* 2009; Li *et al.* 2021) or using geometry-/terrain-following grids (Hwang & Gorlé 2023; Cooke *et al.* 2024). In this study, we leverage the flow solver charLES (developed by Cascade Technologies, Inc.), which uses an automated Voronoi diagram–based meshing approach that enables accurate simulations of natural roughness transitions in complex terrain (Cooke *et al.* 2024) and realistic urban settings (Hwang & Gorlé 2023). In this section, we discuss technical details of the models and UQ methodology used to leverage this technology in informing parameters.

2.1. *Multifidelity models and surface parameterizations*

The multifidelity modeling approach consists of three fidelity levels: (high) geometry-resolving LES, (mid) WMLES, and (low) RANS. The key difference that divides the numerical models into the different fidelities is the extent to which each resolves turbulence and the surface morphology. For each approach, we evaluate to what extent the model can accurately capture the effects of surface heterogeneity using only a limited number of model parameters. The parameters that govern the surface roughness parameterizations for WMLES and RANS are the aerodynamic roughness length z_0 and the displacement length d.

2.1.1. *High-fidelity: geometry-resolving LES*

The high-fidelity model utilizes the charLES solver. charLES uses a second-order-accurate, finite-volume method to solve the compressible, filtered Navier-Stokes equations using the low-Mach isentropic Helmholtz formulation (Ambo *et al.* 2020) with the Vreman subgrid-scale model (Vreman 2004). Time integration uses a third-order Runge-Kutta explicit method. The charLES code is widely used to investigate high-Reynolds-number wall turbulence (Hwang & Gorlé 2023; Cooke *et al.* 2024). In the computational domain, the mesh for the idealized urban canopy is generated automatically by clipped Voronoi diagrams. These geometry-resolving turbulence-capturing data are used as ground truth data in this study. Outside the urban canopy, an algebraic equilibrium wall model is used.

2.1.2. *Mid-fidelity: WMLES*

As the mid-fidelity model, the pseudo-spectral flow solver PadéOps is used to solve the filtered, incompressible Navier-Stokes equations (Ghate & Lele 2017; Howland *et al.* 2020). Fourier collocation in horizontal directions and a sixth-order staggered compact finite difference scheme is used in the vertical direction. Time integration uses a fourth-order strong stability-preserving Runge-Kutta method. In WMLES, surface processes are parameterized by the wall model. The standard equilibrium wall model is used here, which is equivalent to an algebraic wall model based on Monin–Obukhov similarity theory (Howland *et al.* 2020). The wall shear stress is calculated using a matching velocity at the first vertical location $z_m = \Delta z/2$ using the instantaneous, filtered velocity components

$$\tau_{i3}|_{\mathrm{wall}}/\rho = -\left[\frac{\kappa \hat{\tilde{U}}}{\log\left(\frac{z_m - d}{z_0}\right)}\right]^2 \frac{\hat{\tilde{u}}_i}{\hat{\tilde{U}}}, \tag{2.1}$$

where ρ is density, κ is the von Kármán constant, z_0 is the roughness length, d is the displacement length, and $\hat{\tilde{U}} = \sqrt{\hat{\tilde{u}}_m^2 + \hat{\tilde{v}}_m^2}$ is the filtered instantaneous wind speed. To alleviate log-layer mismatch in WMLES (Yang *et al.* 2017), we apply a first-order temporal filter, denoted with the overhanging $\hat{\cdot}$, with a filter timescale of $T_f = 6.25\Delta t$.

2.1.3. *Low-fidelity: RANS*

For the low-fidelity model, we use the solver OpenFOAM to integrate 2D steady, incompressible RANS equations. Turbulence is modeled using the standard $k-\epsilon$ model (Launder & Spalding 1974), which employs the closure coefficients $(C_\mu, C_1, C_2, \sigma_k, \sigma_\epsilon) = (0.09, 1.44, 1.92, 1.0, 1.3)$. In RANS simulations, surface parameterizations are termed wall functions or near-wall treatment. Based on the equilibrium assumption, wall functions are used to model the mean behavior of the velocity U, turbulence kinetic energy (TKE) k and TKE dissipation ϵ at the matching location z_m. The surface shear-stress boundary condition is given by

$$\tau_{i3}|_{\mathrm{wall}}/\rho = -\left[\frac{\kappa \overline{U}}{\log\left(\frac{z_m - d}{z_0}\right)}\right]^2 \frac{\overline{u}_i}{\overline{U}}, \tag{2.2}$$

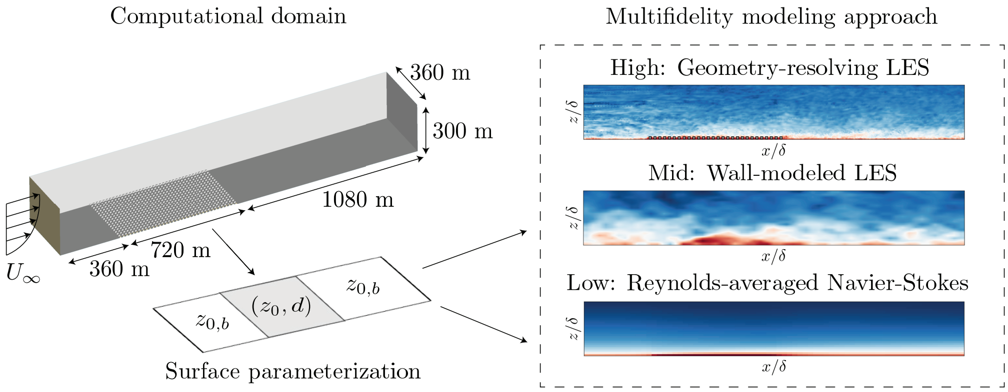

FIGURE 1. Computational domain and multifidelity modeling approach.

where $\overline{U} = \sqrt{\overline{u}_m^2 + \overline{v}_m^2}$ is the mean wind speed at $z_m = 7.5$ m. The geometry is parameterized using the two surface parameters, z_0 and d. The corresponding surface boundary conditions for the $k - \epsilon$ model are

$$k|_{z_m} = \frac{u_*^2}{\sqrt{C_\mu}}\ ,\ \epsilon|_{z_m} = \frac{u_*^3}{\kappa(z_m - d)}, \tag{2.3}$$

where friction velocity is $u_* = \sqrt{\tau|_{wall}/\rho}$.

2.2. *Computational domain and setup*

We use an idealized representation of an urban canopy with a square-array of cubes (Li *et al.* 2021). Coriolis forces are neglected and the flow is neutrally stratified. Our computational setup is shown on the left side of Figure 1. The surface transitions from a region of homogeneous roughness of $z_0 = 0.1$ m to an array of cubes and back to a region of homogeneous roughness equivalent to that upwind of the urban environment. In the lower-fidelity models, the surface roughness is represented by the surface parameterization. As in previous literature (Bou-Zeid *et al.* 2007; Li *et al.* 2021), the idealized urban environment is represented using an isolated patch with representative roughness and displacement lengths (z_0, d) that differ from those of the surrounding environment $(z_{0,b} = 0.1 \text{ m}, d = 0 \text{ m})$.

2.3. *Inverse problem*

We pose the learning of surface model parameters as a Bayesian inverse problem in which we infer parameters using high-fidelity reference data that are inherently noisy. This inverse approach enables the quantification of the uncertainty in model parameters for simplified surface parameterizations used in lower-fidelity models. The relationship between parameters and data is given by

$$\boldsymbol{y} = \mathcal{G}_\infty(\boldsymbol{\theta}) + \eta, \tag{2.4}$$

in which $\boldsymbol{\theta}$ is the set of parameters, $\boldsymbol{y}$ is the reference data, $\eta \sim N(0, \Sigma)$ is a realization of the normally distributed noise, and $\mathcal{G}_\infty$ is the infinite time average of the forward model, in this case, numerical models of different fidelities (low and mid). In this study, the data $\boldsymbol{y}$ is a finite time average generated using the high-fidelity, geometry-resolving LES. The noise associated with the data $\boldsymbol{y}$ is assumed to be normally distributed following the central limit theorem, characterized by the covariance matrix of the internal variability of the system Σ. We note that in this setting there is no measurement error, and we have neglected model error in the inverse problem. The infinite time average is used in Eq. (2.4) because posing the inverse problem using instantaneous data or finite time averages from the model $\mathcal{G}(\boldsymbol{\theta})$ would require the joint estimation of the initial conditions (i.e., state estimation) in the chaotic, turbulent boundary layer. While the infinite time average is available by definition from the solution of a RANS model (low-fidelity), it is not directly available as an output of LES. As described in Section 2.4, a key benefit of the proposed UQ methodology is that the emulator provides a surrogate of the infinite time average of the statistics from model $\mathcal{G}(\boldsymbol{\theta})$, which eliminates the requirement to estimate the initial conditions jointly with the model parameters that are fixed in time. Further discussion is provided in Dunbar *et al.* (2021) and Howland *et al.* (2022).

2.4. *Calibrate-emulate-sample*

We approximate the posterior distribution $\mathbb{P}(\boldsymbol{\theta} \mid \boldsymbol{y})$ using calibrate-emulate-sample (CES), a UQ methodology that combines stochastic optimization and machine learning surrogate modeling with Markov chain Monte Carlo (MCMC) sampling (2021; Oliver *et al.* 2024). The CES methodology significantly accelerates the expensive UQ process by sampling from a surrogate model that is trained using data generated in the calibration stage.

The method is composed of three modular steps. The calibrate stage involves the use of ensemble Kalman inversion (EKI) to perform the calibration of parameters that are fixed in time in the form of an inverse problem. In the emulate stage, the EKI-generated parameter-model output pairs are used to train a Gaussian process emulator to replace the expensive forward model. Because the emulator is trained with parameter-model output pairs that are subject to independent realizations of noise in the chaotic system, the emulator approximates the infinite time (ensemble) average (Howland *et al.* 2022). In the final step of the CES methodology, we sample from the emulator to approximate parameter posteriors. In sampling from the computationally cheap emulator, we effectively estimate posteriors using only the model runs in the EKI process, which is $O(10^2)$ evaluations of the model $\mathcal{G}(\boldsymbol{\theta})$. For a more detailed description of the CES methodology and its previous applications, we refer to (2021), Howland *et al.* (2022), and Oliver *et al.* (2024), with the last referencing the open-source software available on GitHub.

In this study, the CES methodology is used with the mid-fidelity model (WMLES) and low-fidelity model (RANS) separately. In the calibration phase, we use ensemble sizes of 40 members for the WMLES and 50 members for the RANS, with five iteration steps. A fundamental aspect in the Bayesian inversion is the embedding of prior knowledge about the parameters through the prior distribution. We provide physically feasible bounds that minimally inform the parameter space to allow for exploration. An exponential distribution is used for the roughness length, $z_0 \sim \exp(N(0,1))$, and a uniform distribution is used for the displacement length, $d \sim U(0, 7.5)$. The exponential distribution for z_0 is used because values of the aerodynamic roughness length span several decades (Stull 2012). The LES data used here are time averaged for 3 minutes. This averaging timescale was selected as a balance between removing dependence on initial conditions and maintaining a sufficient number of independent samples in the high-fidelity LES.

3. Results

We present the results of the CES methodology in two parts. The first set of results is perfect-model numerical experiments, which are used to verify the UQ method for the two models of different fidelities. The second set of results is imperfect-model experiments in

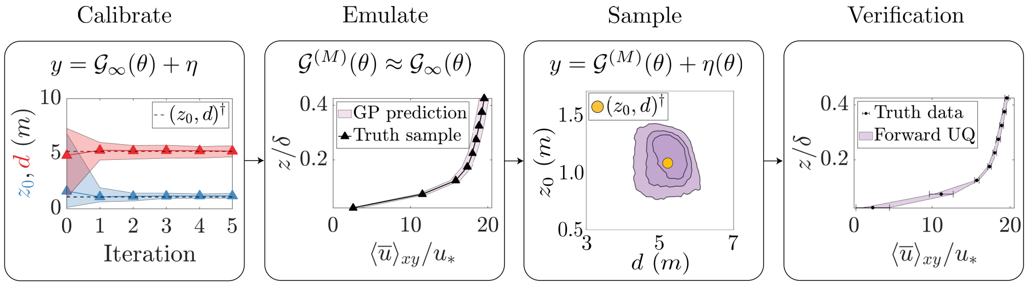

FIGURE 2. Perfect-model experiment with RANS demonstrating the CES pipeline with verification of the UQ method.

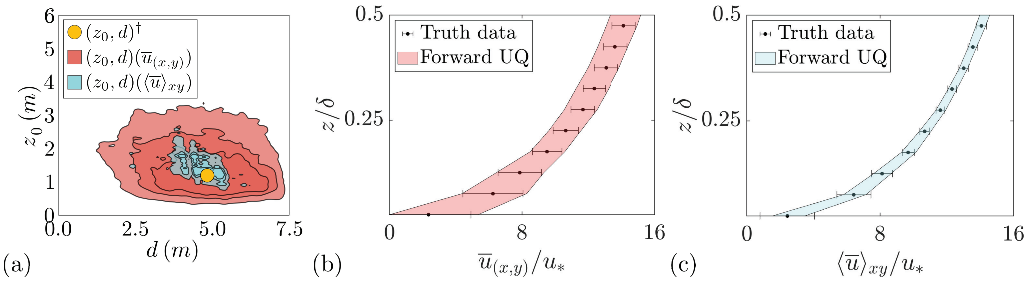

FIGURE 3. Perfect-model experiment with WMLES. (a) Joint posterior distributions inferred from assimilated statistics of different spatial averaging. (b) Forward UQ runs for single-point statistics. (c) Forward UQ runs for horizontally averaged statistics.

which we apply the methodology to learn surface parameters in the low- and mid-fidelity models using data generated by the geometry-resolved high-fidelity LES.

3.1. *Perfect-model setting*

In the perfect-model setting, the ground truth data $\boldsymbol{y}$ are constructed from the same (fidelity) forward model with a prescribed set of parameters. This eliminates any model form error that may arise from the use of different models, and leaves only sample bias as the mode of deviation in parameter estimation. In this perfect-model context, we prescribe truth parameters $\boldsymbol{\theta}^\dagger$ that are used to generate data, $\boldsymbol{y} = \mathcal{G}(\boldsymbol{\theta}^\dagger)$. Because the CES pipeline observes only data $\boldsymbol{y}$ and not the prescribed parameters, the successful recovery of these prescribed parameters $\boldsymbol{\theta}^\dagger$ serves as a verification of the UQ process.

The flowchart in Figure 2 presents results of a perfect-model experiment using RANS. In the calibrate phase, the EKI performs gradient-free optimization that iteratively concentrates the ensemble of parameters around the true parameters. These are used to train an emulator of the parameter-to-forward model mapping. The MCMC sampling of the emulator shows a joint posterior distribution of the 50%, 75%, and 99% highest posterior density regions (HPDR) with the prescribed truth parameters (in yellow) contained. In the verification step, the posterior is sampled, here using Monte Carlo sampling, and propagated through the forward model. We observe the 95% credible intervals from the forward model evaluations of the horizontally averaged nondimensional mean wind speed, $\langle\overline{u}\rangle_{xy}/u_*$, to fairly accurately capture the spread of the synthetic truth data. The overbar $\bar{\cdot}$ represents a time average and $\langle\cdot\rangle$ represents a spatial average. The horizontal average is performed over the urban canopy, shown in Figure 1.

We also perform a perfect-model experiment using WMLES as the model $\mathcal{G}(\boldsymbol{\theta})$. In this second test, we also examine the impact of spatial averaging in assimilated statistics on parameter uncertainty. In Figure 3(a), the red posterior is estimated on the basis of the assimilation of nonspatially averaged mean wind speed, denoted $\overline{u}_{(x,y)}$, whereas the blue posterior is estimated using horizontally averaged mean wind speed, denoted $\langle\overline{u}\rangle_{xy}$. The prescribed truth parameter is contained in both posteriors, indicating the robustness of the UQ methodology to the assimilated statistics. Taking the 75% HPDR, we observe the area of the posterior from nonspatially averaged statistics to be approximately eight times larger than the posterior from horizontally averaged statistics due to the spatial averaging operation, which reduces the noise. The key result of this section is that we are able to achieve tractable inverse UQ for CFD models, as demonstrated using both RANS and WMLES.

3.2. *Imperfect-model setting*

In the imperfect-model setting, the lower-fidelity model surface parameters are learned from geometry-resolved high-fidelity data $\boldsymbol{y}$ for which there are no known truth parameters. To evaluate the data-driven UQ approach to estimating the parameters, we compare results with those of a standard morphometric approach to estimate the necessary input surface parameters (z_0, d) for the RANS and WMLES. We use the morphometric model presented in Macdonald *et al.* (1998) as the reference method. The model predicts the input surface parameters (z_0, d) on the basis of key nondimensional parameters in the urban morphology: plan area density λ_p and frontal area density λ_f, which are determined by the building height H and distance between buildings D in an array of cubes.

A comparison is done for the case of building height $H = 10$ m and distance between buildings $D = 20$ m. The UQ is performed for the RANS model on the basis of horizontally averaged statistics from the urban canopy. In Figure 4(a), the joint posterior distribution resulting from the CES UQ methodology yields a significantly different displacement length from that estimated using the morphometric approach. In Figure 4(b), we show that the 95% credible intervals of an ensemble of samples from the posterior propagated through the model accurately capture the spread of high-fidelity data observed during the UQ process at the first several grid points. We observe 75.0% reduction in root-mean-square error (RMSE), compared to the morphometric approach, in evaluations against the calibration data. This finding demonstrates the effectiveness of the CES methodology in inferring surface parameters from high-fidelity simulation data and reducing error relative to the data seen during the UQ process. The small discrepancy in the profile beyond these points is likely due to model form error between the RANS and the geometry-resolving LES. In Figure 4(c), we show out-of-sample predictions in a comparison of the forward model prediction of CES posteriors with those of reference morphometric parameters. In predicting unobserved wind speeds downwind of the urban canopy, errors are more variable, demonstrating that the model form error is nonnegligible.

Next, in Figure 5, we evaluate the impact of spatial averaging of the statistics $\boldsymbol{y}$ provided to the UQ method on the posterior distributions for both the RANS and the WMLES. In the WMLES, only the cases for assimilated statistics of no spatial averaging and horizontal averaging are plotted due to computational limitations. Here, we discuss results of the RANS posterior distributions from Figure 5(b); however, we note that WMLES shows a similar qualitative trend for the cases investigated here. We observe that the uncertainty associated with inferred parameters is highest when using single-point assimilated statistics(red shading) because the noise is highest without spatial averaging. When examining horizontally averaged statistics, we evaluate two streamwise mea-

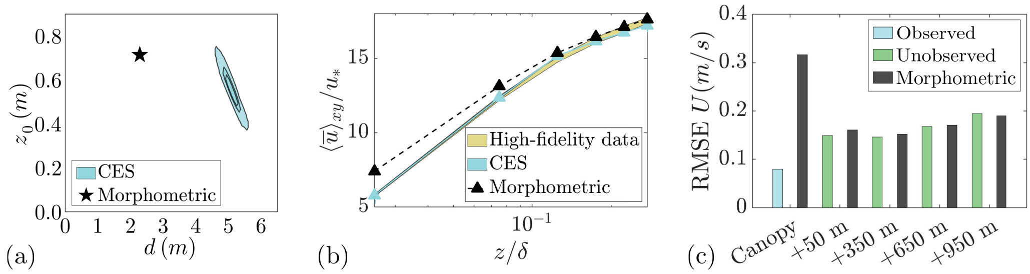

FIGURE 4. Comparison in the RANS model setting. (a) Joint posterior distributions estimated by CES with parameters estimated by the morphometric model. (b) Spatially averaged wind speed from RANS using parameters estimated by CES and the morphometric model. (c) RMSE for observed and unobserved wind speed using parameters estimated by CES and the morphometric model.

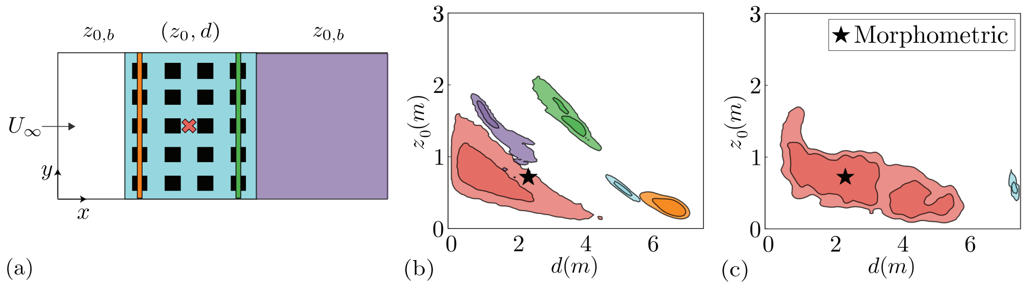

FIGURE 5. (a) Schematic of spatial averaging of the statistics assimilated by inverse UQ: no spatial averaging(red shading), lateral averaging near start(orange shading) and end(green shading) of urban canopy, and horizontal averaging within(cyan shading) and downwind(purple shading) of urban canopy. Joint posterior distributions of 75% and 99% HPDR inferred from averaged statistics for (b) RANS and (c) WMLES, respectively.

surement locations: within(cyan shading) and downwind(purple shading) of the urban canopy. The parameters from assimilating statistics within the urban canopy are closest to equilibrium values over the entire urban environment. However, in results not shown, horizontal averaging of the urban canopy over different streamwise distances shows that the velocity does not reach equilibrium by the end of the patch. Assimilating statistics from the region downwind of the canopy yields a low displacement height and a high roughness length. This may be attributed to enhanced turbulent mixing over the urban canopy, which results in a more uniform velocity profile. Further, the information related to displacement length diminishes due to the disappearance of canopy wind shear near the surface in the region downwind of the urban environment.

In the case of laterally averaged statistics, we examine two streamwise positions: one near the start(orange shading) and the other near the end(green shading) of the urban canopy. At the end of the canopy, the inferred roughness length is observed to be higher, while the displacement height is lower than the parameter values inferred from horizontally averaged statistics. On the other hand, for parameters inferred from laterally averaged statistics near the start of the canopy, we observe a lower roughness length and a higher displacement height than those from horizontally averaged statistics.

4. Conclusions

In this study, we investigate surface model parameter uncertainty in an idealized urban canopy by coupling geometry-resolving LES using charLES with an efficient machine learning–accelerated inverse UQ method. As a key result, we demonstrate inverse UQ of CFD models with a relatively small number $\mathcal{O}(10^2)$ of forward model evaluations. We note that if the parameters are universal, as typically assumed in standard morphometric approaches, then this cost of $\mathcal{O}(10^2)$ CFD simulations would need to be performed only once. The UQ method is verified in a perfect-model setting in which we prescribe and recover ground truth parameters. In evaluating the proposed data-driven UQ approach, we show a reduction in modeling error within the training data compared to the standard morphometric approach, but a negligible impact on unobserved data, suggesting an opportunity to reduce model form error. We demonstrate that the parameter calibration and quantified uncertainty significantly vary as a result of spatial averaging in assimilated statistics, and that the data-driven parameter estimates differ from those of the morphometric model in a statistically significant manner. This data-driven, multifidelity approach to learning surface parameterizations for the urban environment lays out a flexible framework to improve computational models for turbulent geophysical flows. Future work will investigate a range of urban environments and atmospheric conditions and will validate the efficient inverse UQ methodology using field observations.

Acknowledgements

This research was performed with the support of the Center for Turbulence Research (CTR) Summer Program 2024 at Stanford University. The authors thank Themistoklis Vargiemezis from Stanford for assistance with OpenFOAM. E.Y.S. and M.F.H. acknowledge funding from MIT UMRP. E.Y.S. acknowledges gracious support from the MIT MathWorks fellowship. Simulations were performed on the Stampede3 supercomputers under the NSF ACCESS project ATM170028 and on Stanford's Yellowstone supercomputer.

REFERENCES

Ambo, K., Nagaoka, H., Philips, D. A., Ivey, C., Brès, G. A. & Bose, S. T. 2020 Aerodynamic force prediction of the laminar to turbulent flow transition around the front bumper of the vehicle using dynamic-slip wall model LES. In *AIAA Scitech 2020 Forum*, p. 0036.

Bou-Zeid, E., Anderson, W., Katul, G. G. & Mahrt, L. 2020 The persistent challenge of surface heterogeneity in boundary layer meteorology: a review. *Bound.-Layer Meteorol.* **177**, 227–245.

Bou-Zeid, E., Overney, J., Rogers, B. D. & Parlange, M. B. 2009 The effects of building representation and clustering in large-eddy simulations of flows in urban canopies. *Bound.-Layer Meteorol.* **132**, 415–436.

Bou-Zeid, E., Parlange, M. B. & Meneveau, C. 2007 On the parameterization of surface roughness at regional scales. *J. Atmos. Sci.* **64**, 216–227.

Cleary, E., Garbuno-Inigo, A., Lan, S., Schneider, T. & Stuart, A.M. 2021 Calibrate, emulate, sample. *J. Comput. Phys.* **424**, 109716.

Cohen, A. E., Cavallo, S. M., Coniglio, M. C. & Brooks, H. E. 2015 A review of planetary boundary layer parameterization schemes and their sensitivity in simulating southeastern U.S. cold season severe weather environments. *Weather Forecast.* **30**, 591–612.

Cooke, J., Jerolmack, D. & Park, G. I. 2024 Mesoscale structure of the atmospheric boundary layer across a natural roughness transition. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2320216121.

Dunbar, O. R., Garbuno-Inigo, A., Schneider, T. & Stuart, A. M. 2021 Calibration and uncertainty quantification of convective parameters in an idealized GCM. *J. Adv. Model. Earth Sys* **13**, e2020MS002454.

Ghate, A. S. & Lele, S. K. 2017 Subfilter-scale enrichment of planetary boundary layer large eddy simulation using discrete Fourier–Gabor modes. *J. Fluid Mech.* **819**, 494–539.

Grimmond, C. S. B. & Oke, T. R. 1999 Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Meteorol. Climatol.* **38**, 1262–1292.

Howland, M. F., Dunbar, O. R. & Schneider, T. 2022 Parameter uncertainty quantification in an idealized GCM with a seasonal cycle. *J. Adv. Model. Earth Sys* **14**, e2021MS002735.

Howland, M. F., Ghate, A. S. & Lele, S. K. 2020 Influence of the geostrophic wind direction on the atmospheric boundary layer flow. *J. Fluid Mech.* **883**, A39.

Hwang, Y. & Gorlé, C. 2023 Large-eddy simulations to define building-specific similarity relationships for natural ventilation flow rates. *Flow* **3**, E10.

Iaccarino, G. & Verzicco, R. 2003 Immersed boundary technique for turbulent flow simulations. *Appl. Mech. Rev.* **56**, 331–347.

Launder, B. & Spalding, D. 1974 The numerical computation of turbulent flows. *Comput. Methods Appl. Mech. Eng.* **3**, 269–289.

Li, Q., Bou-Zeid, E., Grimmond, S., Zilitinkevich, S. & Katul, G. 2020 Revisiting the relation between momentum and scalar roughness lengths of urban surfaces. *Q. J. R. Meteorol. Soc.* **146**, 3144–3164.

Li, Q., Yang, J. & Yang, L. 2021 Impact of urban roughness representation on regional hydrometeorology: an idealized study. *J. Geophys. Res. Atmos.* **126**, e2020JD033812.

Ma, P.-L. et al. 2022 Better calibration of cloud parameterizations and subgrid effects increases the fidelity of the E3SM Atmosphere Model version 1. *Geosci. Model Dev.* **15**, 2881–2916.

Macdonald, R., Griffiths, R. & Hall, D. 1998 An improved method for the estimation of surface roughness of obstacle arrays. *Atmos. Environ.* **32**, 1857–1864.

Oliver, R., Bieli, M., Garbuno-Iñigo, A., Howland, M., de Souza, A.N., Mansfield, L.A., Wagner, G.L. and Efrat-Henrici, N. 2024 Calibrateemulatesample.jl: accelerated parametric uncertainty quantification. *J. Open Source Softw.* **9**, 6372.

Stull, R. B. 2012 *An Introduction to Boundary Layer Meteorology*. Springer Science & Business Media.

Vreman, A. 2004 An eddy-viscosity subgrid-scale model for turbulent shear flow: algebraic theory and applications. *Phys. Fluids* **16**, 3670–3681.

Yang, X. I., Park, G. I. & Moin, P. 2017 Log-layer mismatch and modeling of the fluctuating wall stress in wall-modeled large-eddy simulations. *Phys. Rev. Fluids* **2**, 104601.