Received: 22 January 2022 | Revised: 28 February 2023 | Accepted: 4 March 2023 | Published on: 5 April 2023

DOI: 10.1002/qj.4445

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Predictability of passive scalar dispersion in atmospheric surface layers with urban-like roughness: A large-eddy simulations study

Yanle Lu[1] | **Luca Delle Monache**[2] | **Jeffrey Weil**[3] | **Keith Ngan**[4] | **Qi Li**[1]

[1]School of Civil and Environmental Engineering, Cornell University, Ithaca, New York USA

[2]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California San Diego, San Diego, California USA

[3]National Center for Atmospheric Research, Boulder, Colorado USA

[4]School of Energy and Environment, City University of Hong Kong, Hong Kong, China

Correspondence
Qi Li, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA.
Email: ql56@cornell.edu

Funding information
National Science Foundation, Grant/Award Numbers: NSF-AGS (2028644), NSF-CAREER (2143664), NSF-CBET (2028842)

Abstract

The predictability of passive scalar dispersion is of both theoretical interest and practical importance, for example for high-resolution numerical weather prediction and air quality modeling. However, the implications for the numerical modeling of urban areas remain relatively unexplored. Using obstacle-resolving large-eddy simulations (LES), we conducted twin experiments, with and without a velocity perturbation, to investigate how the presence of urban roughness affects error growth in streamwise velocity (u) and passive scalar (θ) fields, as well as the differences between error evolutions in u and θ fields. The predictability limit is characterized using the signal-to-noise ratio (SNR) as a continuous metric to indicate when error reaches saturation. The presence of urban roughness decreases T_p of the passive scalar by around 20% compared to cases without them. The error statistics of θ indicate that urban roughness-induced flow structures and different scalar source locations affect the scalar dispersion and relative fluctuations, which subsequently dictate the evolution of the SNR. Analysis of the passive scalar error energy ($\epsilon_\theta{}^2$) budget indicates that the contributions from advective transport by the velocity and velocity error dominate. The error energy spectra of both u and θ exhibit a $-5/3$ slope in flat-wall cases, but not in the presence of urban roughness, thereby highlighting the deviation from the assumption of locally isotropic turbulence. This study reveals that urban roughness can decrease the predictability of the passive scalar and destroy the similarity between the error statistics of the velocity and the passive scalar.

KEYWORDS

large-eddy simultion, urban canopy, passive scalar dispersion, predictability

1 | INTRODUCTION

Predictability refers to the capability to predict the future state of a system and the predictability horizon is the time period within which the prediction errors do not exceed some prechosen magnitude and a chaotic system can be appreciably predicted (Lorenz, 1969; Kalnay, 2003; Mukherjee *et al.*, 2016). Understanding the growth of errors, limitations to the predictability and extending the predictability horizon are subjects of great fundamental and practical importance in numerical weather prediction. The pioneering work in a series of remarkable papers

(Lorenz, 1963, 1965, 1969; Leith and Kraichnan, 1972) has investigated the predictability of fully developed homogeneous, isotropic turbulence using statistical turbulence theory. It has been shown that an initial uncertainty of arbitrarily small scale and amplitude spreads through the entire inertial range in a finite time. The results have obtained significant attention and were followed by a number of studies of chaotic dynamics, and of the predictability of atmospheric phenomena across multiple scales ranging from planetary to synoptic, with different atmospheric forecast models (e.g., Ruelle and Takens, 1971; Eckmann and Ruelle, 1985; Smith *et al.*, 1999; Laprise *et al.*, 2000; Tribbia and Baumhefner, 2004; Ngan and Eperon, 2012).

While the characteristic grid resolution of numerical weather prediction (NWP) models has increased, some of them start to resolve atmospheric motions including part of the boundary layer turbulence at unprecedented fine scales (i.e., 10–100 m) (e.g., Hacker, 2010; Schalkwijk *et al.*, 2015; Lee *et al.*, 2019; Wiersema *et al.*, 2020). The issues regarding how errors arising from initial and boundary conditions grow and affect the predicted quantities of interest are becoming more pressing with this development of high-resolution NWP models. For example, Mukherjee *et al.* (2016) observed that the characteristics of error growth in convective boundary layer exhibits the exponential error growth phase and saturation phase, and it is independent of perturbation wavenumber but is sensitive to the Reynolds number. It is expected that such advancement in high-resolution NWP modeling will bring about transformative changes to many meteorological applications in urban areas, such as prediction of extreme weather events and dispersion modeling of atmospheric pollutants, which require fine-scale, intra-city-level predictions. For example, a 100-m resolution NWP model was applied to study a clear convective boundary layer over London, UK (Lean *et al.*, 2019). As such, understanding how errors propagate in models is relevant to high-resolution NWP modeling in cities. Nevertheless, it is well known that the mean and turbulent flows in urban areas differ from the non-urban counterpart and influence other quantities of interest, for example, pollutant dispersion (e.g., Tomas *et al.*, 2015; 2016). Thus, the characteristics of error propagation, predictability and transient dynamics for flows in the urban context (e.g., Millward-Hopkins *et al.*, 2012; Ngan and Lo, 2017) are expected to deviate from that for canonical homogeneous, isotropic turbulence. For instance, Lo and Ngan (2015) found that the presence of urban street canyons induces a strong shear layer, which profoundly alters the error growth dynamics in the velocity field as errors are advected from the roof level into the canyon. Moreover, studies of these issues in the urban context are also relatively scarce compared to those for canonical turbulence. Therefore, in this study, we are first motivated to investigate the implications of the dynamical differences for behaviors of error propagation in turbulence-resolved atmospheric surface layer (ASL).

In addition, apart from the error propagation and dynamic predictability in velocity, the error growth in modeling passive scalar dispersion in the urban boundary layer has received relatively little attention, despite its theoretical and practical importance. This is perhaps due to invoking the Reynolds analogy (e.g., Van Driest, 1959; Brutsaert, 1965; Kays and Crawford, 1993), which assumes that momentum and passive scalars are transported similarly, especially at high Reynolds number when only turbulent velocity and length scales dominate their effective diffusivities (Yang and Abkar, 2018). However, this assumption may fail for high-roughness boundary layer flows, since there is no counterpart of the pressure gradient (PG) term (in Equation 1; PG here represents $-\frac{1}{\rho}\frac{\partial p}{\partial x_i}$) in the passive scalar equation (Equation 2) (van Ulden, 1978; Horst, 1979; Sullivan *et al.*, 2018). For incompressible flows, the Navier–Stokes, scalar conservation, and continuity equations are

$$\frac{Du_i}{Dt} = PG_i - \frac{\partial \tau_{ij}}{\partial x_j} \tag{1}$$

$$\frac{D\theta}{Dt} = S(\mathrm{x}_i, t)\,\delta(x_i - x_{0i}) - \frac{\partial q_i}{\partial x_i} \tag{2}$$

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{3}$$

where the subscript i represents the variables in streamwise, spanwise and vertical directions; u_i is the velocity vector; t refers to time; τ_{ij} is the stress tensor; PG_i is the pressure gradient; θ is the passive scalar; $S(x_i, t)$ is multiplied by the Dirac delta function, signifying that the source is zero except for the region whose locus is x_{0i} (Finnigan *et al.*, 2003).

Hence, extrapolating conclusions regarding dynamic predictability and error growth in velocity to the scalar counterpart may be questionable if the Reynolds analogy is violated. In fact, the error propagation and predictability for momentum and passive scalar may indeed differ. From the source term in Equation (2), $S(x_i, t)$, the specific details of both the locations (i.e., as a function of x_i) and the type of the scalar sources/sinks (i.e., the time-dependent nature of S) may be important in leading to different behaviors of error propagation in the scalar field compared to velocity. Therefore, the second motivation of this study is to further quantify and analyze such distinctions.

By conducting numerical simulations using large-eddy simulation (LES) of neutral, wall-bounded flows, we study the evolution of error fields diagnosed from identical-twin experiments in which small perturbations are added to a

well-developed velocity field. Our analyses of the numerical experiments are guided by the following questions: (1) how do the twin simulations of velocity and passive scalar fields diverge and what are the differences between velocity and passive scalar in a neutral ASL? (2) how is the error growth affected by the presence of urban street canyons and different scalar source locations?

This paper is organized as follows. LES and the details of simulations set up are introduced in Section 2, the determination of the predictability horizon of dispersion of the passive scalar and the difference between different cases are analyzed in Section 3, and Section 4 presents the summary and concluding remarks.

2 | METHODS

2.1 | LES

In the present study, we use an LES code which solves the Navier–Stokes equation using a pseudospectral discretization in the horizontal directions and a centered second-order finite difference in the vertical. The subgrid-scale LES model is the scale-dependent Lagrangian dynamic model developed in Bou-Zeid *et al.* (2005). For the simulations with building structures, the immersed boundary method (IBM) approach following Tseng *et al.* (2006), Chester *et al.* (2007), and Li *et al.* (2016a, 2016b) is adopted. It resolves the complex building structure explicitly when solving the pressure Poisson equation and enforces the velocity inside the buildings to be zero. For the simulations without building structures, we used the same numerical scheme and subgrid-scale model, but the code is based on the open-source LESGO (LESGO, 2020).

The governing-filtered incompressible Navier–Stokes equations with the continuity constraint and the scalar advection–diffusion equation are given below:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{4}$$

$$\frac{\partial u_i}{\partial t} + u_j\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right) = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + F_i + B_i \tag{5}$$

$$\frac{\partial \theta}{\partial t} + u_i\frac{\partial \theta}{\partial x_i} = -\frac{\partial q_i}{\partial x_i} \tag{6}$$

where the subscript i represents the variables in streamwise, spanwise, and vertical directions; u_i is the resolved velocity vector; t refers to time; τ_{ij} is the deviatoric part of the subgrid stress tensor; ρ is a constant of unity; p is the modified pressure that includes the resolved and subgrid-scale kinetic energy; F_i is a constant pressure gradient along the streamwise direction to drive the flow (here a homogeneous steady horizontal pressure gradient along the x direction with magnitude $\rho u_*^2/z_i$, for z_i being the boundary layer height in this study); B_i is the immersed boundary force representing the action of the obstacles on the fluid; θ is the resolved passive scalar; and q_i denotes the subgrid-scale scalar flux. The momentum equations are solved in the rotational form to ensure the conservation of kinetic energy (Orszag and Pao, 1975), and all the variables in the equations are filtered quantities.

2.2 | Simulation setup

For all the simulations, the computational domain dimensions are $(L_x, L_y, L_z) = (1{,}200,\ 600,\ 300)$ m with (192, 96, 48) grid points in the streamwise, spanwise, and vertical directions. A prescribed constant streamwise pressure gradient is imposed for all cases. The neutral condition is considered, and there is no inversion layer. The boundary layer height z_i is 200 m, which is constrained by the domain height and will be used for normalization. Periodic boundary conditions are applied at the lateral boundaries of the domain. The top boundary condition for momentum is zero stress. An equilibrium logarithmic wall model is applied for the ground and building surfaces. A roughness length of $z_0 = 0.005$ m, representing the microscale roughness of the surfaces is used to parameterize the momentum flux. The microscale is associated with its own local velocity, and the aggregate effects of the bluff-rough surfaces are corresponding to macroscale roughness (Li *et al.*, 2020). The same domain dimensions and boundary conditions are used in cases with the presence of urban roughness. Here the urban structures are considered to be idealized two-dimensional street canyons (Figure 1a), since it is a prototype used for urban climate studies and is therefore employed here as a necessary first step to investigate error propagation induced by resolving urban roughness (Cui *et al.*, 2004). The width and height of buildings are both equal to 50 m and represented by eight grid points, denoted as H, and the distance between each array of buildings is also H. The vertical profile of mean streamwise velocity in the simulation domain, with and without urban roughness, is shown in Figure 1c. The chosen resolution is based on previous sensitivity studies using the same code (Tseng *et al.*, 2006; Li and Bou-Zeid, 2019).

The boundary condition for the scalar is zero flux on the ground, the surface of buildings and top of the domain, except for the source region we defined. The passive scalar source, shown in Figure 1b, in the simulation domain is represented by an area surface flux located either on the ground or on the building roof. Similar to Caulton *et al.* (2018), a Gaussian smoothing is applied to

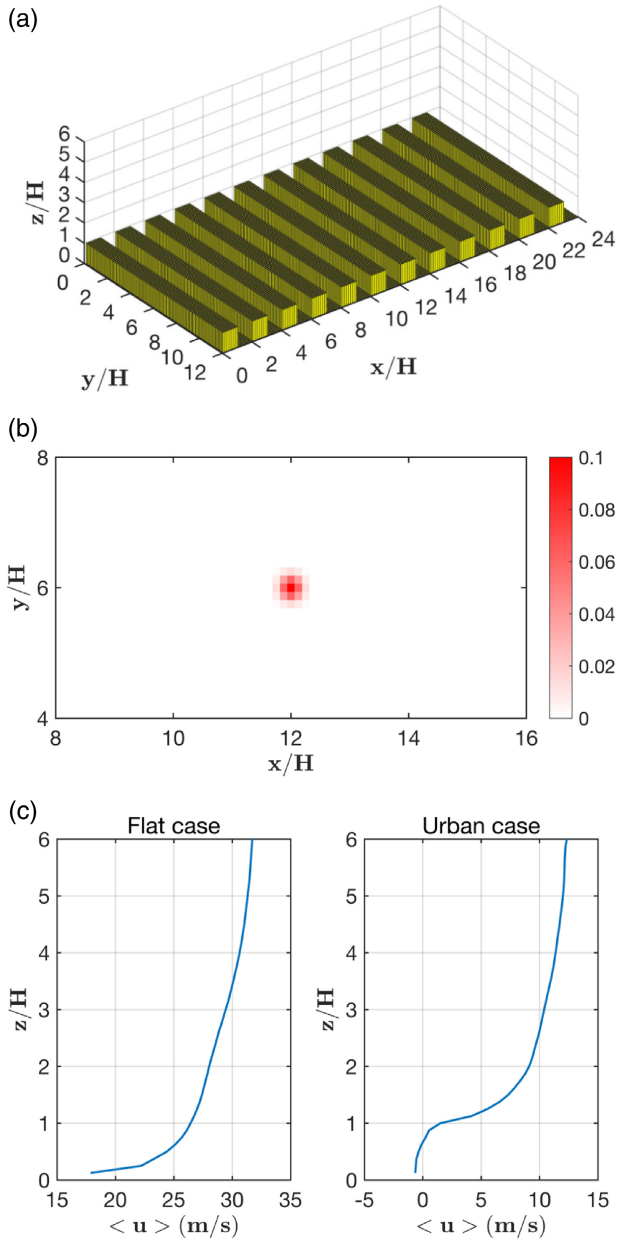

FIGURE 1 (a) Domain setup of urban topography in urban cases; H indicates building height; the aspect ratio of building is 1; (b) 2D projection of the source for passive tracer, which is located either on the ground or roof in different cases; (c) vertical profile for mean streamwise velocity over the simulation domain, with and without urban roughness [Colour figure can be viewed at wileyonlinelibrary.com]

the area surface flux to avoid numerical errors caused by grid discontinuities.

We set up identical-twin experiments to study the evolution of initial perturbations to the streamwise velocity component. First, the simulations are integrated over a spin-up period to obtain a statistical stationary state M_0 for the velocity field. Then the scalar starts to be continuously released as described in Figure 1b. One of the simulations continues from M_0 to yield the control case. A perturbation of small amplitude (to be defined later) is added to the streamwise component of velocity to form a perturbed state, which is integrated in time to yield the perturbation run. This setup is used to mimic the practical problem of predicting the dispersion given a known initial condition of the velocity field with some uncertainty, which is represented by the added perturbation. The difference between the control and the perturbation run, referred as the "error" hereafter, will propagate from velocity to the scalar and grow with time.

The initial perturbation is a two-dimensional sinusoidal field that is applied to the velocity only (Equation 7):

$$u_{\mathrm{p}}(x,y,z) = u_{\mathrm{c}}(x,y,z) + A\sin\left(\frac{4\pi kx}{L_x}\right)\sin\left(\frac{2\pi ky}{L_y}\right) \quad (7)$$

where subscript p represents the perturbation case, and c represents the control case. Its amplitude A is kept small (value shown in Table 1) and k is the dimensionless wavenumber in the x and y directions for all z, that is, the perturbation is independent of height. We here chose the same form of perturbation as what was applied by Mukherjee *et al.*, 2016 to their potential temperature fields. The added small-amplitude perturbation has negligible effects on the divergence of the velocity field as the current LES model applies a pseudospectral method, with which incompressibility can be enforced to a high degree of accuracy by applying the projection operator. Note that a height-dependent perturbation can be applied to study predictability within the urban canopy; however, Lo and Ngan (2015) found that error propagation is not sensitive to the location of the perturbation added to the streamwise velocity field. Thus, the perturbation for flat and urban cases follows Mukherjee *et al.* (2016). It is added to streamwise velocity only to facilitate analysis of the spatial error propagation and the interaction between velocity and scalar errors. Since the error growth rate is approximately independent of k (cf. Section 3.1.1), in agreement with Mukherjee *et al.* (2016), only a fixed value of k is considered.

The identical-twin simulations start to diverge after the perturbation (Equation 7) is added. Notice that with the same governing equation, boundary conditions and forcing, the twin simulations will reach the same statistical stationary state and the ensemble-averaged statistics of velocity and scalar will be independent of the initial conditions. Here we focus on T_{p}, the time elapsed between seeding and error saturation (the growth rate of error tends to zero, cf. Section 3.1.1). The error evolution and T_{p} may also be affected by the presence of the urban roughness and the location of scalar source. Thus, several different cases are examined (Table 1).

TABLE 1 Summary of cases: Flat Control (FC), Flat Amplitude (FA), Flat Wavenumber (FK), Urban Ground (UG), and Urban Roof (UR)

Case name	Morphology	Amplitude (A/u^*)	Wave number (k)	Source location
FC	Flat	0.001	4	Ground
FA	Flat	0.1	4	Ground
FK	Flat	0.001	2	Ground
UG	Urban	0.001	4	Ground
UR	Urban	0.001	4	Roof

Note: u^* is the friction velocity (here related to the background kinetic mean pressure gradient driving the flow F_i).

3 | RESULTS AND DISCUSSION

In this section, the evolution of errors in the velocity and scalar fields is analyzed to understand their similarities and differences. In addition, different metrics to recognize the stages of divergence and quantify the predictability horizon will be analyzed.

3.1 | Error evolution and predictability horizon analysis

3.1.1 | Error evolution

Here we first examine the volume-averaged errors, which are defined as:

$$e_\theta = < \left(\theta_p - \theta_b\right)^2 > \quad (8)$$

$$e_u = < \left(u_p - u_c\right)^2 > \quad (9)$$

where subscript p represents the data from the perturbation case, c represents the data from the control case, u and θ stand for streamwise velocity and passive scalar respectively, and <> refers to the spatial average over the whole 3D computational domain, D_0. Figure 2 shows e_u and e_θ, as a function of time t normalized by the characteristic time scale for wall-bounded flow, that is, eddy turnover time t^*, which is defined as the boundary layer height in the simulations, z_i, divided by the friction velocity u_*. The dotted red line in Figure 2 indicates $t/t^* = 1$. The perturbation is added to the u field only. Since the influence of the other two velocity components, that is, the spanwise velocity and vertical velocity, on the volume-averaged error is negligible (see Figure A1), the analysis of velocity errors will focus on the u field.

As shown in Figure 2, the growth of e_θ and e_u in all cases is similar and can be divided into three phases as in a convective atmospheric boundary layer (Mukherjee *et al.* (2016)). First, there is an onset period showing constant exponential-error growth, which is represented by the linear increase rate in the log–log graph, after seeding the perturbation; next, the error starts to saturate with a decreasing growth rate; finally, the growth rate tends to zero and the error approaches a plateau value, which can differ significantly among the cases for e_θ.

All the cases share similar exponential growth in the first phase for error fields (Figure 2a), which suggests the loss of predictability is governed by a linear instability. However, the growth rates differ: since flows over the urban canopy exhibit analogous characteristics to mixing-layer flows (Brunet *et al.*, 1994; Raupach *et al.*, 1996; Roth, 2000), differences in the vertical shear between flat and urban cases lead to different linear instabilities. The plateau value of e_θ in the Urban Ground (UG) case is two orders of magnitude larger than that of

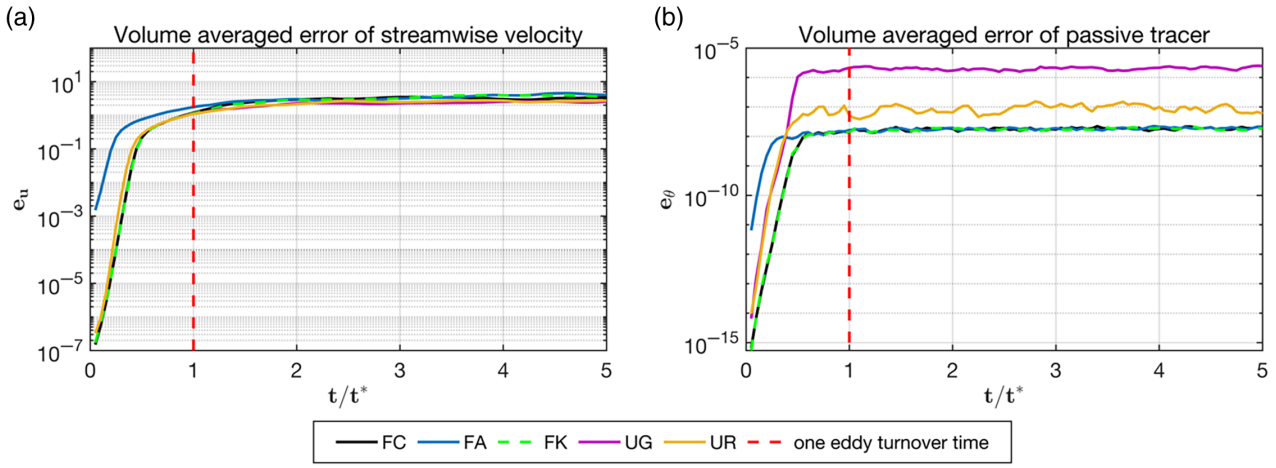

FIGURE 2 Volume-averaged errors for (a) streamwise velocity, u, and (b) the passive scalar, θ, at different times, where one eddy turnover time is indicated by $t^* = z_i/u_*$. The red dotted line indicates $t/t^* = 1$ [Colour figure can be viewed at wileyonlinelibrary.com]

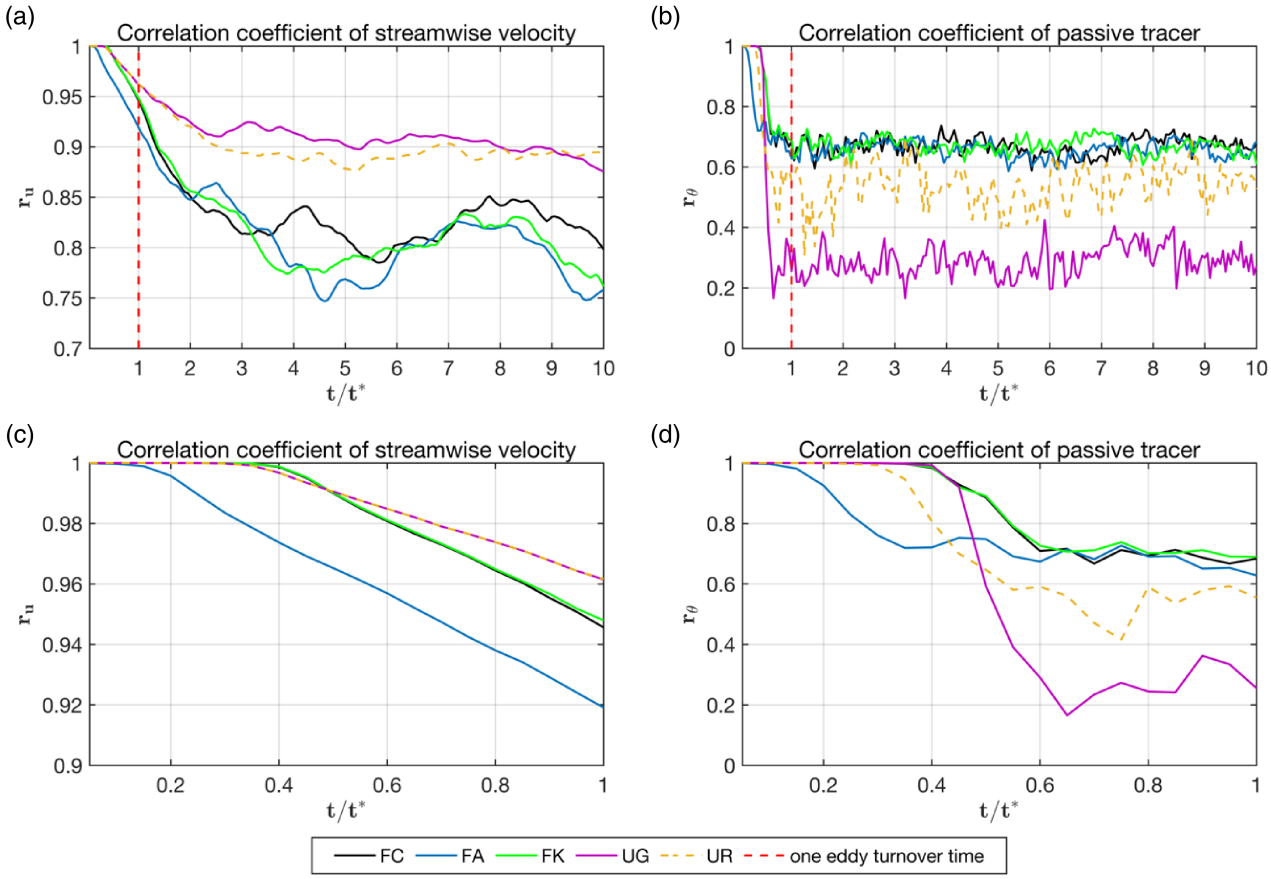

FIGURE 3 Correlation coefficient of (a,c) streamwise velocity, u, and (b,d) passive scalar, θ. The dotted line indicates $t/t^* = 1$ [Colour figure can be viewed at wileyonlinelibrary.com]

the cases without urban roughness (i.e., the flat cases), while the Urban Roof (UR) case falls between UG and the flat cases (Figure 2b). It suggests that even though the existence of urban roughness does not influence the saturation plateau of error in velocity, it does increase the magnitude of errors in the scalars (UG and UR), which are highly related to the influence of building structure in the canyon and roughness sublayers. In addition, these two cases indicate a clear dependence on the location of the scalar source, which is tied to the fact that error propagation near the shear-dominated building roof differs from that inside the canyon (Lo and Ngan, 2015). The natural evolution of tracer, that is, $\left\langle (\theta_c(t) - \theta_c(0))^2 \right\rangle^{1/2}$, decorrelates on a faster timescale than the perturbation, as there is almost no linear growth phase if in a log scale. The detailed physical meaning is discussed in Figure 4 (Section 3 below) as RMSC. More statistics of scalar error evolution from the physical perspective of scalar dispersion will be discussed in Section 3.2.

Compared to other cases, the Flat Amplitude (FA) case has a larger initial volume-averaged error in both u and θ. This is due to the larger magnitude of the initial perturbation. The larger initial error in θ also indicates that the initial error in the streamwise velocity field propagates quickly to the scalar field. The Flat Control (FC) and Flat Wavenumber (FK) cases always show the same trend and almost identical values, which is consistent with the insensitivity to the wavenumber k (Section 2.2; see also Figure A2).

Differences between the error growth in u and θ are revealed by considering the correlation coefficient r_X of X, with X being u or θ from the control (X_c) and perturbed (X_p) simulations. r_X is defined as:

$$r_X = \frac{\mathrm{cov}\left(X_c, X_p\right)}{\sigma_{X_c}\sigma_{X_p}} \tag{10}$$

where σ_{X_c}, σ_{X_p} are the standard deviation of X_c and X_p, respectively, and $\mathrm{cov}\left(X_c, X_p\right)$ is their covariance. σ_{X_c}, σ_{X_p} and $\mathrm{cov}\left(X_c, X_p\right)$ are computed from the three-dimensional snapshots at time t of $u(x, y, z, t)$ and $\theta(x, y, z, t)$.

Figure 3b shows that for both u and θ, the evolution of $r_u(t)$ and $r_\theta(t)$ can be divided into three stages, as with the volume-averaged errors of Figure 2. For $t/t^* < 0.5$, $r_u(t)$ and $r_\theta(t)$ almost remain constant at around one. Then there is a rapid decrease followed by oscillations around

some plateau value. In all cases, $r_\theta(t)$ is less than $r_u(t)$ for $t/t^* > 1$ and reaches the plateau levels much more rapidly. This indicates that passive scalar dispersion is much more sensitive to initial error compared to velocity. In general, the correlation coefficients of different cases indicate more distinctions compared to the volume-averaged errors of Figure 2. First, $r_u(t)$ and $r_\theta(t)$ in the FA case start to decrease earlier and more rapidly than in other cases, which is consistent with the trend in volume-averaged error and related to a non-linear effect. Secondly, the presence of street canyons significantly decreases $r_\theta(t)$ but slightly increases $r_u(t)$ during the final phase. The value at which oscillations around a constant level occurs is about 0.3 for the scalar in the UG case, and 0.5 in the UR case. Both are significantly lower than those in the flat cases (Figure 3b). The lower correlation coefficients in UR and UG are consistent with the large volume-averaged errors depicted in Figure 2b.

In summary, the volume-averaged error and correlation coefficient both indicate clear diverging behaviors of the twin simulations and some differences between u and θ across all cases. In the next section, we apply two methods to quantify the predictability horizon, T_p — that is the duration over which deterministic prediction of passive scalar dispersion can be made.

3.1.2 | Predictability horizon analysis

The noise-to-signal ratio (NSR), which is the inverse of the signal-to-noise ratio (SNR), is used to analyze outputs from the twin simulations and determine the predictability limit. NSR is defined as the ratio of the amplitude of noise power to the amplitude of the signal power. The signal refers to the root-mean-square difference of the scalar field between time t and the initial time in the control simulation (RMSC). The noise refers to the root-mean-square difference of the scalar field between the perturbed simulation and the control simulation (RMSP). This method has been applied in mesoscale modeling to indicate the divergence of twin simulations, such as in Zeng and Pielke (1993, 1995). The relevant formulas are listed below:

$$\mathrm{RMSC}(\theta) = \left[\frac{1}{N}\sum_{D_0}\left(\theta_C(t) - \theta_c(0)\right)^2\right]^{1/2} \tag{11}$$

$$\mathrm{RMSP}(\theta) = \left[\frac{1}{N}\sum_{D_0}\left(\theta_P(t) - \theta_c(t)\right)^2\right]^{1/2} \tag{12}$$

$$\mathrm{NSR} = \mathrm{SNR}^{-1} = \frac{\mathrm{RMSP}}{\mathrm{RMSC}} \tag{13}$$

where subscripts C and P denote the control and perturbed simulation respectively, and D_0 refers to the spatial domain in which these metrics are calculated. D_0 is taken to be the whole computational domain here, and N represents the number of grid points in D_0. These statistical metrics (i.e., RMSC, RMSP and NSR) quantify the divergence between the twin simulations. Moreover, they embody physical significance, which can be understood by connecting their analogous behaviors to the mean concentration ($\overline{\theta}$), root-mean-square concentration (σ_θ) and the concentration fluctuation intensity ($\sigma_\theta/\overline{\theta}$) respectively, for a surface point-source plume in a neutral atmospheric boundary layer (ABL). Further analysis regarding this later point will also be examined in the next section (cf. Section 3.2) by dividing D_0 into different subdomains to understand the reasons for the differences among cases.

First, we use the NSR to quantify how T_p varies across different cases. Based on the definitions, the noise power will exceed signal power when NSR > 1 (Zeng and Pielke, 1993), and we thus use NSR($t = T_p$) = 1 to indicate the threshold of predictability. Note that although different threshold values can be chosen, we choose NSR = 1 to denote the level above which the predicted scalar fields in the twin simulations diverge more than the variability of the control case.

From Figure 4a, the trend of NSR is almost the same for all the flat-wall cases, which increases first and then gradually decreases after $t = t^*$, and tends to a constant value (~0.55) at large times $t/t^* = 9$ or 10. It can be understood by the temporal variation of RMSC and RMSP in Figure 4b: RMSP and RMSC both increase rapidly right after the seeding of the perturbation, and the value of RMSC is larger than that of RMSP during this stage. At around one eddy turnover time, the difference between RMSC and RMSP is very small and as a result, NSR almost reaches one. However, as t approaches and goes beyond T_p, the volume error e_θ reaches a plateau (e.g., Figure 2b); thus RMSP, which is related to e_θ, also tends to a constant level. RMSP slightly increases with time afterwards while RMSC monotonically increases since the scalar continuously released from a constant source is mixed within the domain D_0. Therefore, NSR gradually decreases.

However, for cases with street canyons, the evolution of these quantities is different. First, the NSR becomes larger than one for urban cases at around $t/t^* = 0.6$, and then levels off with more oscillations. This is more obvious for the UG case while the UR case falls between UG and the flat-wall cases, whose trend is similar to those of flat cases but with more oscillations and higher NSR value. More subdomain analysis to reveal the physical reasons for the different performance is in next section (cf. Section 3.2). The trend of RMSC is non-monotonic in the UG case (Figure 4c), while being almost monotonic for the case

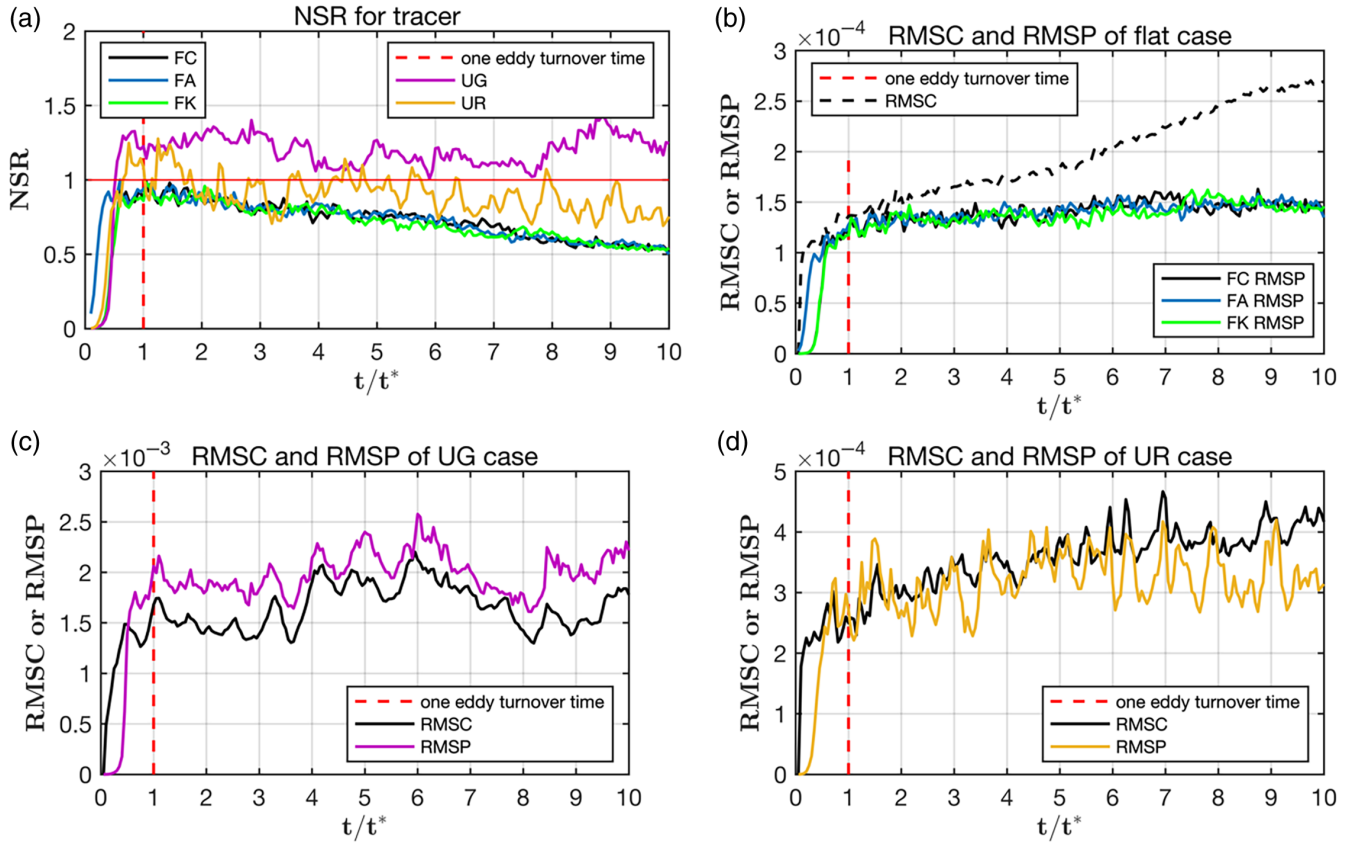

FIGURE 4 (a) NSR for scalar field of all cases, and RMSC and RMSP of (b) flat cases, (c) case UG, (d) case UR [Colour figure can be viewed at wileyonlinelibrary.com]

UR (Figure 4d) although increasing more slowly than for flat cases (Figure 4b). Notice that the magnitude of RMSC and RMSP in the UR case is of the same order as in the flat cases, that is, about twice the flat-case value, whereas those in case UG are about one order of magnitude larger. This can be explained by a significant fraction of the scalar being trapped within the canopy and the reduced scalar dispersion above the canopy sublayer in UG as shown in Figure 5 that the scalar error, $\theta_p - \theta_c$, is one order of magnitude larger than UR. Thus, higher concentration inside the canyon and near-source high scalar fluctuations give rise to large RMSC and RMSP. In addition, oscillations of larger amplitude and lower frequencies are observed for cases UG and UR after one eddy turnover time, For example, the deviation of UG RMSC is up to 20% of the average value of approximately 1.7×10^{-3}. The larger amplitude of these oscillations than for the flat cases may be due to the turbulence-organized motions leading to ejections (Inagaki *et al.*, 2012; Li and Bou-Zeid, 2019). These motions cause the mean concentration ($\overline{\theta}$) and root-mean-square concentration (σ_θ) to fluctuate more strongly than in the flat cases. The UR case shares features of both the flat and UG cases, and the characteristic values of all properties lie between them.

In summary, initial velocity errors propagate very quickly to the scalar field and subsequently influence the dispersion of the passive scalar. The continuous metric NSR conveys the message about the predictability horizon, T_p, differences among cases. The urban roughness tends to decrease T_p, and if the source of the passive scalar is located within the canyon (i.e., case UG) rather than on the roof (i.e., case UR), the characteristics of predictability will be more significantly modified compared to the flat-wall cases. On the other hand, the exact magnitudes of T_p can depend on the Reynolds number (Re) of the LES results as discussed in Mukherjee *et al.* (2016). They concluded that the error growth and predictability horizon, T_p, is expected to be resolution-dependent using LES since the effective Re of LES is inherently resolution-dependent, and the time of the initial-error growth phase is proportional to $\mathrm{Re}^{-1/2}$. Although a neutral flow over the urban canopy is considered here, which is different from their convective boundary layer setup, a similar conclusion of reduction in T_p by about 50% is found in double-resolution cases (see Figure A3). This is as expected since Re of LES is physically related to the scale separation between the smallest scale resolved in LES (i.e., scales modeled by the subgrid-scale model) and the energy-containing

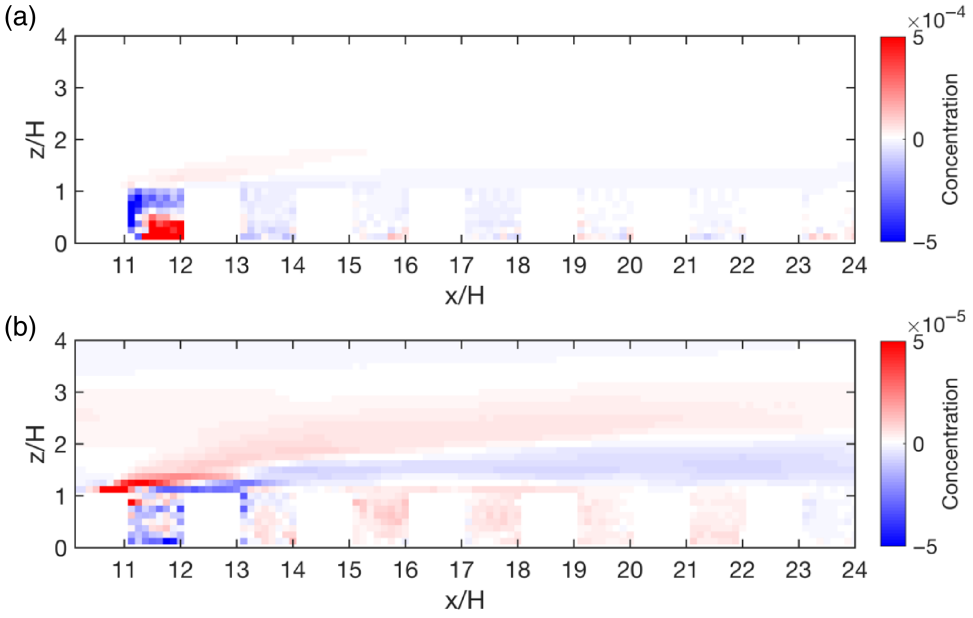

FIGURE 5 The spatial distribution of the scalar error for cases (a) UG and (b) UR. The scalar error shown here is the average of the first three eddy turnover times [Colour figure can be viewed at wileyonlinelibrary.com]

scale (Sullivan and Patton, 2011; Li *et al.*, 2018), where $\mathrm{Re} = \left(\frac{z_i}{\Delta^f}\right)^{\frac{4}{3}}$, z_i is the outer scale and Δ^f is the filter size commensurate with the LES grid resolution. Thus, for the same z_i, T_p is proportional to $\Delta^{f-2/3}$ and doubling the resolution leads to $(1/2)^{-2/3}T_\mathrm{p} \sim 0.6T_\mathrm{p}$. While the predictability of velocity and scalar are dependent on Re and resolution as described above, we found that the differences between the predictability of velocity and scalar remain independent of Re and the resolution (see Figures A4 and A5). Furthermore, a test run of a small ensemble of eight realizations is conducted for ground-level source release with different initial conditions by adding noise to the velocity field. The probability density function of the spatially averaged scalar for each realization shows similar performance with weak variabilities (see Figure A6), which indicates that a scalar continuously released in a statistical stationary flow is not sensitive to the initial conditions. Conclusions drawn based on the performance of RMSC, RMSP, NSR in one pair of twin simulations are representative.

3.2 | Factors influencing the divergence of θ and T_p

Next, we examine the effects of canopy and roughness sublayers on error evolution for different height intervals, that is, different subdomains D_0 (Figure 6). The height of the buildings is represented by H and we consider three subdomains: inside canyon, D_{01}, (i.e., the canopy sublayer), between building height H and $2H$, D_{02}, (i.e., the roughness sublayer) and above $2H$, D_{03} (i.e., the upper half of the boundary layer).

For the upper half of the boundary layer, that is, $D_0 \in z > 2H$, the source in all cases appears more like a scalar released from the surface. Thus, the UG and UR cases are expected to look more like the results in the flat cases, and as shown in Figure 6c, NSRs do share a similar trend across all cases. Large differences seen between the urban and flat cases are for $D_0 \in z < 2H$. To explain the differences, we make connections between RMSC, RMSP, NSR and their analogous behaviors to the mean concentration ($\overline{\theta}$), root-mean-square concentration (σ_θ) and the concentration fluctuation intensity ($\sigma_\theta/\overline{\theta}$) respectively, for a surface point-source plume (Fackrell and Robins, 1982; Sykes, 1988; Weil *et al.*, 1992) in a neutral ABL.

First, for flat cases (Figure 6d–f), these metrics generally resemble those at larger z, except with earlier attainment of a plateau with higher magnitudes, for RMSC and RMSP, due to the proximity (in z) to scalar source locations. The length of time for attainment of the constant for RMSC is due to the plume evolution with distance and time. The region nearer the source tends to reach a statistically steady state earlier. With a continuously releasing scalar source, the more distant region will continue to grow or evolve even though further concentration changes may be small for statistics averaged over the whole domain (i.e., Figure 5). The attainment of an approximately constant NSR far downstream (Figure 6d–f) is analogous to the finding of a constant fluctuation intensity ($\sigma_\theta/\overline{\theta}$) for a continuous surface point source in neutral boundary layer experiments (Fackrell and Robins, 1982). In the case of $\sigma_\theta/\overline{\theta}$, Sykes (1988) argues that the plume is always about the same size as the local turbulent eddies, which leads to self-similar behavior and a constant $\sigma_\theta/\overline{\theta}$ with distance ($x \leq 7z_i$). A key difference between the fluctuation

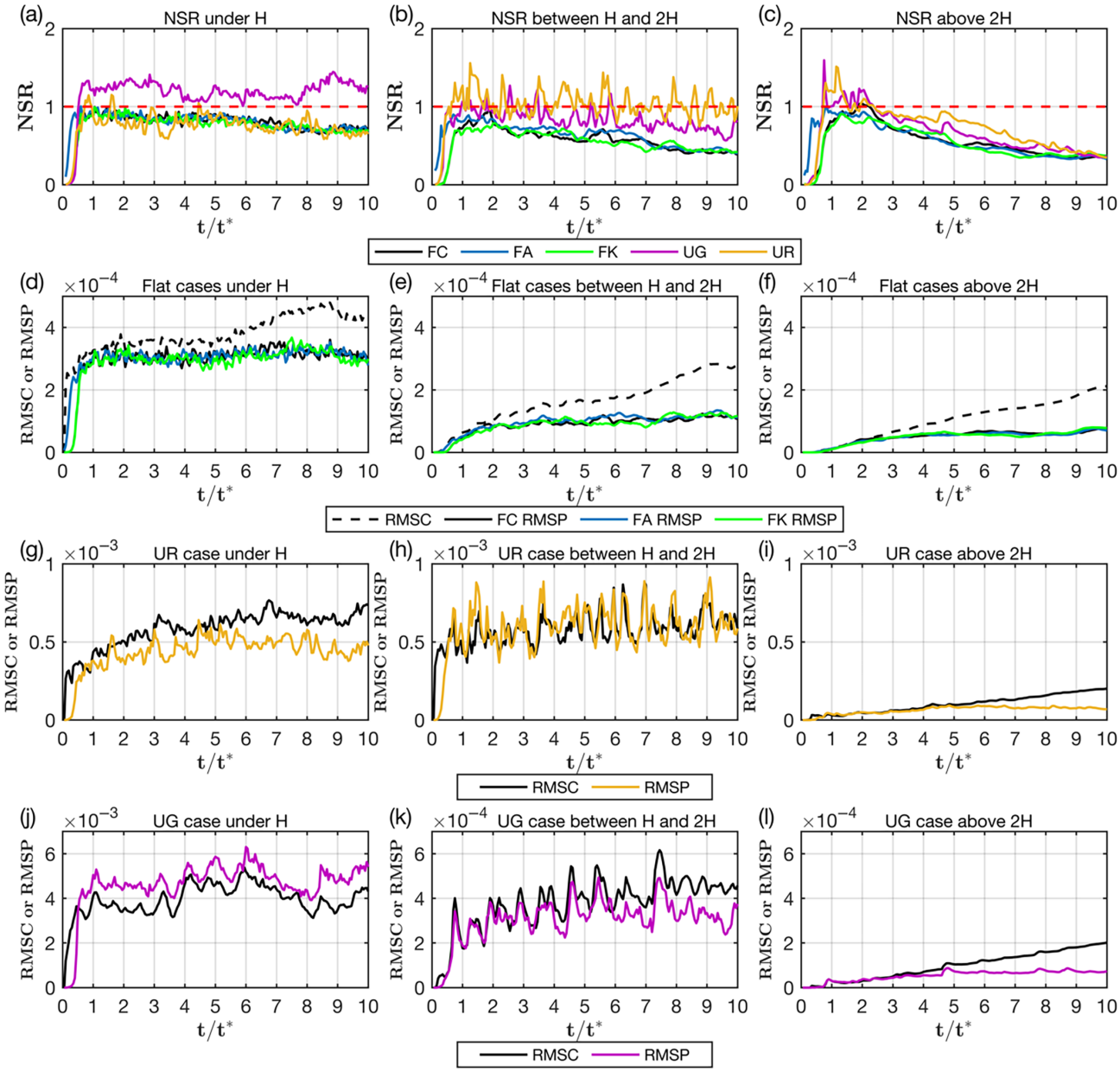

FIGURE 6 Noise-to-signal ratio (NSR) of the scalar field for all cases in the subdomains (a) $z < H$ (building height), (b) $H < z < 2H$ and (c) $z > 2H$; RMSC and RMSP for all flat cases in the subdomain; (d–f) same as (a–c) except for case UR; (j–l) same as (a–c) except for case UG [Colour figure can be viewed at wileyonlinelibrary.com]

intensity and NSR is that the former is based on "point" concentration values in a steady-state plume, whereas the NSR is determined from volume-averaged concentrations of a time-evolving plume. Nevertheless, the NSR ($\cong 0.55$) at large times, $\frac{t}{t^*} = 9$ or 10, is close to the fluctuation intensity (0.5) in the plume experiments.

In urban cases, RMSC and RMSP have larger fluctuations and exhibit different temporal trends. The RMSC does not increase monotonically with time in all height intervals, rather RMSP and RMSC are almost synchronized in their fluctuations (e.g., Figure 6h,k). For Figure 6j, RMSP has a greater value than RMSC, which is especially different from the flat-case counterparts and is tied to the trapping of the scalar in the canopy and the low transport in and out of the canopy. Also, the magnitudes of RMSC and RMSP for $D_0 \in H < z < 2H$ is one order of magnitude smaller than in the region $D_0 \in z < H$ for the UG case (Figure 6j compared to k). This is physically consistent with the fact that most of the production of scalar fluctuations occurs near the source (i.e., hence large RMSP for $z < H$). The scalar from the near-surface source is effectively mixed within the canopy but little escapes above given the current geometry of street canyons, thus a large RMSC for $z < H$. On the other hand, in the UR case with

roof-level source location, the magnitudes of RMSC are comparable for $D_0 \in 0 < z < H$ and for $D_0 \in H < z < 2H$ in Figure 6g,h. This is because the scalar from the roof level can be entrained into the street canyon and meanwhile can be transported downwind above the canopy. Similarly, scalar fluctuations in both $D_0 \in 0 < z < H$ and $D_0 \in H < z < 2H$ are comparable, leading to RMSP of the same order of magnitude in both subdomains.

Next, to better understand how errors propagate after the initial seeding, the error budget is now investigated. The starting point for the error budget analysis is the filtered equation for the passive scalar that is solved:

$$\frac{\partial \widetilde{\theta}}{\partial t} + \widetilde{u}_i \frac{\partial \widetilde{\theta}}{\partial x_i} = s - \frac{\partial q_i}{\partial x_i} = s + \frac{\partial}{\partial x_i} D \frac{\partial \widetilde{\theta}}{\partial x_i} \tag{14}$$

where $\widetilde{\theta}$ refers to the resolved passive scalar, subscript i represents the variables in streamwise, spanwise and vertical directions, $\widetilde{u}_i$ refers to the resolved velocity field, and s is the source of the scalar and will be zero for the locations where there is no release. $\frac{\partial q_i}{\partial x_i}$ is the subgrid-scale scalar flux and it is calculated by $\frac{\partial}{\partial x_i} D \frac{\partial \widetilde{\theta}}{\partial x_i}$, in which D is the subgrid-scale scalar diffusivity. Superscripts c and p are used to represent the control case and perturbed case respectively, and therefore we have $\widetilde{\theta}^{\mathrm{c}}$, $\widetilde{\theta}^{\mathrm{p}}$, $\widetilde{u}_i^{\mathrm{c}}$ and $\widetilde{u}_i^{\mathrm{p}}$. Then the differences between $\widetilde{u}_i^{\mathrm{c}}$ and $\widetilde{u}_i^{\mathrm{p}}$ ($\widetilde{\theta}^{\mathrm{c}}$ and $\widetilde{\theta}^{\mathrm{p}}$) in the twin simulations (here also referred to as errors) are represented by $\varepsilon_i = \widetilde{u}_i^{\mathrm{p}} - \widetilde{u}_i^{\mathrm{c}}$ ($\varepsilon_\theta = \widetilde{\theta}^{\mathrm{p}} - \widetilde{\theta}^{\mathrm{c}}$). The equation of ε_θ can then be obtained by subtracting equation Equation (14) for $\widetilde{\theta}^{\mathrm{p}}$ *and* $\widetilde{\theta}^{\mathrm{c}}$, while substituting $\varepsilon_i = \widetilde{u}_i^{\mathrm{p}} - \widetilde{u}_i^{\mathrm{c}}$ whenever appropriate:

$$\begin{aligned} &\frac{\partial \varepsilon_\theta}{\partial t} + \widetilde{u}_i^{\mathrm{c}} \frac{\partial \varepsilon_\theta}{\partial x_i} + \varepsilon_i \frac{\partial \varepsilon_\theta}{\partial x_i} + \varepsilon_i \frac{\partial \widetilde{\theta}^{\mathrm{c}}}{\partial x_i} \\ &\quad = \frac{\partial}{\partial x_i} \left(D^{\mathrm{c}} \frac{\partial \widetilde{\theta}^{\mathrm{c}}}{\partial x_i} - D^{\mathrm{p}} \frac{\partial \widetilde{\theta}^{\mathrm{p}}}{\partial x_i} \right) \end{aligned} \tag{15}$$

Multiplying Equation (15) by ε_θ and defining the scalar error energy as $\frac{1}{2}\varepsilon_\theta^2$, then gives:

$$\begin{aligned} &\underbrace{\frac{\partial \left(\frac{1}{2}\varepsilon_\theta^2\right)}{\partial t}}_{\text{(I)}} = \underbrace{-\widetilde{u}_i^{c} \frac{\partial \left(\frac{1}{2}\varepsilon_\theta^2\right)}{\partial x_i}}_{\text{(II)}} \underbrace{- \varepsilon_i \frac{\partial \left(\frac{1}{2}\varepsilon_\theta^2\right)}{\partial x_i}}_{\text{(III)}} \\ &\underbrace{- \varepsilon_i \, \varepsilon_\theta \frac{\partial \left(\widetilde{\theta^c}\right)}{\partial x_i}}_{\text{(IV)}} + \underbrace{\varepsilon_\theta \frac{\partial}{\partial x_i} \left[\mathrm{D}^{c} \frac{\partial \widetilde{\theta^c}}{\partial x_i} - \mathrm{D}^{p} \frac{\partial \widetilde{\theta^p}}{\partial x_i} \right]}_{\text{(V)}} \end{aligned} \tag{16}$$

On the left-hand side, term (I) is the local storage of scalar error "energy". This quantity is also related to RMSP examined above, which is important to understand the transition of NSR and T_{p}. On the right-hand side, term (II) denotes the advection by the velocity of the control case, term (III) is the transport by error of velocity, term (IV) represents the interaction between errors of the scalar and the velocity acting on the scalar gradient, and term (V) is the subgrid-scale scalar flux between twin simulations.

The contributions of different terms in Equation (16) are calculated for each grid point and averaged over the whole computational domain, D_0. Their changes with time are shown in Figure 7. First, the change of scalar energy has a constant exponential growth rate (Figure 7b) for $t/t^* < 0.5$ in most cases, which is consistent with Figure 2, even though it looks almost identical to zero because of the linear scale of axes. However, the oscillations and different term contributions for the error energy become clear after error starts to saturate. It takes a finite time for the seeding perturbation in the velocity field to propagate to θ and become saturated, and this timeframe is significantly reduced in FA with large amplitude of initial error. For the flat cases, term (IV) and term (V) dominate, which means that the scalar error energy is produced by errors of both u and θ acting on the local scalar gradient, which is dissipated by the subgrid-scale motions. The relative magnitudes of these terms change significantly from flat cases to urban cases. For the urban cases, all terms are of similar orders of magnitude. In other words, term (II) and term (III), representing the advective transport by velocity and error of velocity, become more important than in the flat cases, which is a key difference of the scalar error energy budget between the flat-wall and urban cases. This difference is directly attributed to the horizontal spatial inhomogeneity introduced by the street canyons. In the flat cases, with horizontally homogeneous advective transport, all terms with the $\frac{\partial}{\partial x_i}$ component will be small with $i = 1, 2$, except the $\frac{\partial \widetilde{\theta}^{\mathrm{c}}}{\partial x_i}$ because of the local scalar source in the domain. At the same time, the magnitude of the vertical velocity is small, which means that $\widetilde{u}_3^{\mathrm{c}}$ and ε_3 are close to zero. Therefore, only the horizontal components of $\frac{\partial \widetilde{\theta}^{\mathrm{c}}}{\partial x_i}$ matter in the flat cases. However, in urban cases, the urban street canyon introduces horizontal inhomogeneity for $z < 2\mathrm{H}$ and strong vertical motions within the canopy sublayer are expected because of continuity, which will increase the amplitude of the vertical velocity. Therefore, term (II) and term (III) become non-negligible in urban cases.

Apart from the differences, they do share similarities. Term (IV) is always positive, while term (V) is always negative. Term (II) is positive for most of the time, even though it is small in the non-urban cases. Term (III) does not show a consistent sign but oscillate around zero. Therefore, the interactions of scalar errors and the transport of scalar by

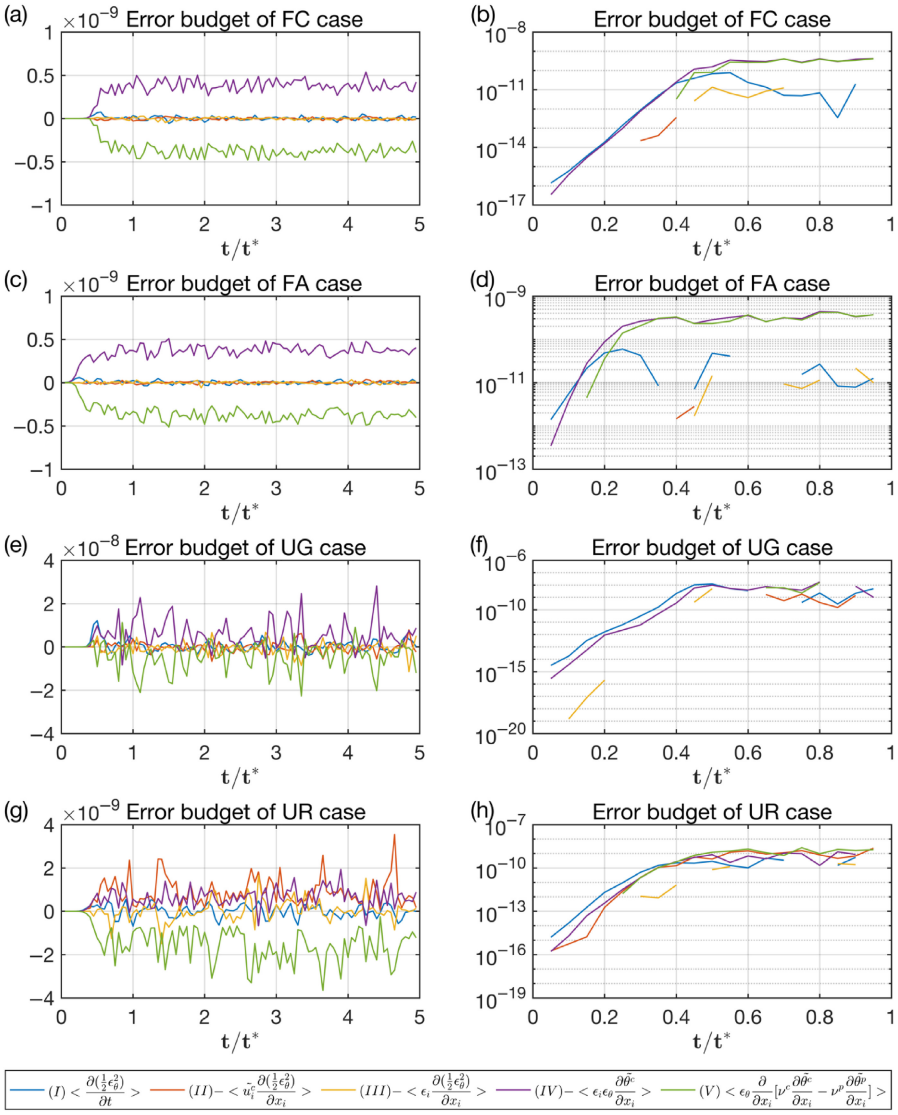

FIGURE 7 All terms in the error budget change with time for case (a,b) FC, (c,d) FA, (e,f) UG, and (g,h) UR. <> represents averaging over the whole domain, D_0. Panels b, d, f and h are shown with a log-linear axis with zoomed-in view of terms in the error budget [Colour figure can be viewed at wileyonlinelibrary.com]

the velocity error (i.e., term IV) and advective transport by the velocity error (i.e., term II in urban cases) provide dominant sources to the scalar error magnitude, while the subgrid-scale transport (i.e., term V) is the main sink.

The vertical profiles of these terms are shown in Figure 8, which are averaged over both the horizontal plane and the time of 0∼10t^* after perturbation. They further confirm the above observations. In addition, all the terms only have a large value near the source level and almost zero for other height.

3.3 | Spectral analysis

All above analyses in Sections 3.1 and 3.2 are conducted for quantities in physical space. To further understand how errors diverge and propagate among different scales in velocity and scalar fields and find out if the building structures influence the error spectra, the spectral correlation coefficient and power spectra of error are investigated.

As the twin simulations begin to diverge, the correlation of cases for both the velocity and scalar fields decreases in three similar stages as described (cf. Section 3.1) but has a different rate of decrease and plateau value. Here we calculate the spectral correlation coefficient between the twin simulations, which is defined as Equation (17), to further understand the divergence in spectral space.

$$\rho(k,z,t) = \frac{E_{\mathrm{CP}}(k,z,t)}{\sqrt{E_{\mathrm{c}}(k,z,\mathrm{t})E_{\mathrm{p}}(k,z,t)}}, \tag{17}$$

where $\rho(k,z,t)$ is the spectral correlation coefficient at wavenumber k, height z and time t; E_{cp} is the cospectrum for the velocity or scalar field of twin simulations; and E_{c} and E_{p} are the power spectra of the control and perturbed simulation respectively. Considering the discontinuities

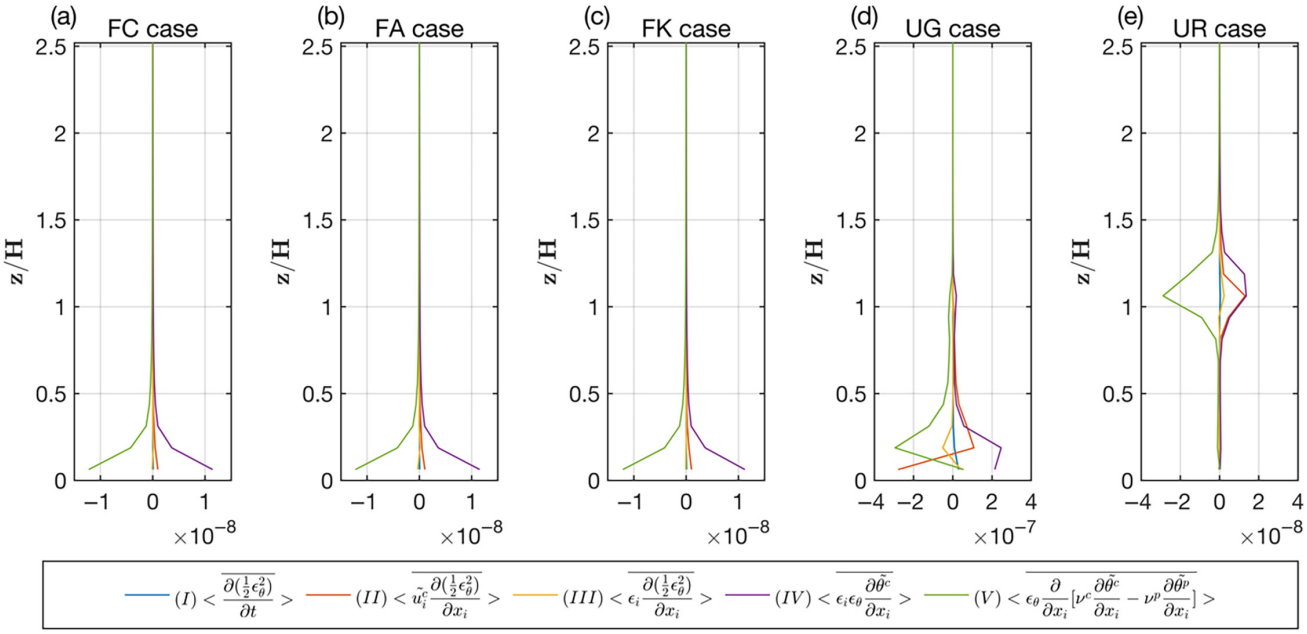

FIGURE 8 Vertical profile of all terms in the error budget for case (a) FC, (b) FA, (c) FK, (d) UG and (e) UR. <> represents averaging in the horizontal plane, represents averaging over $0{\sim}10t^*$ [Colour figure can be viewed at wileyonlinelibrary.com]

induced by the building structures, the $\rho(k,z,t)$ is only calculated for a height z above the roof level.

The vertically averaged spectral correlation coefficient for above roof level, that is, $\rho(k,t) =< \rho(k,z,t) >$, for the velocity and scalar in all cases shares an identical trend. The correlation decreases for all wavenumbers but with different rates among scales. Higher wavenumbers drop to $\rho = 0$ more quickly, which means that the decorrelation time is shorter for small scales (see Figures A7 and A8). This feature was consistent with the paper of Mukherjee *et al.* (2016), in which they proposed that there is a Reynolds-dependent regime for small scales and a Reynolds-independent regime for larger scales when considering decorrelation time.

The power spectra-of-error field, that is, the difference in the twin simulations, is Fourier-transformed in the streamwise direction, and then averaged over the spanwise direction. To analyze the vertical variability, the results of three different heights ($z = 1.125H$, $2H$ and $4H$) are shown below. The definition is given by:

$$\Delta E_u(k,z,t) =< \frac{1}{2}\boldsymbol{\varepsilon_u}\cdot\boldsymbol{\varepsilon_u^*} > \tag{18}$$

$$\Delta E_\theta(k,z,t) =< \frac{1}{2}\boldsymbol{\varepsilon_\theta}\cdot\boldsymbol{\varepsilon_\theta^*} > \tag{19}$$

where ΔE_u and ΔE_θ are error energy spectra, $\boldsymbol{\varepsilon_u}$ and $\boldsymbol{\varepsilon_\theta}$ are the Fourier transform of errors in u and θ, respectively, and <> represents averaging in the spanwise direction.

The existence of buildings can induce high-order harmonics for spatial spectra analysis; $z = 1.125H$, that is, just above the roof level, is therefore chosen to understand the influence of urban structures. The evolution of ΔE_u in the first two eddy turnover times after seeding of perturbation for case FC and case UG is shown in Figure 9, and the color map represents the normalized time after the seeding. The growth of the error spectra in both cases among different scales is similar, but the existence of building affects the error propagation to large scales. The perturbation added to a certain wavenumber appears as the peak for ΔE_u for $t = 0$ (i.e., the lowest line in Figure 9), and the perturbation in both cases spreads to all scales within $t = 0.1t^*$, that is, cascades both upscale and downscale simultaneously. After the initial adjustment, ΔE_u increases with time and the rate of increase is similar across a wide range of scales prior to saturation. Such self-similar spectral shape is also reported by Mukherjee *et al.* (2016) in the potential temperature error spectrum during the constant error growth phase. Then, the error will first saturate at smaller scales and at larger scales saturate at a slower rate. Another trend that can be noticed is that the error growth rate decreases with increasing height, thus a longer time is needed to reach saturation for higher regions. A $k^{-5/3}$ scaling can be observed in the inertial subrange when the error spectrum becomes saturated. This range is almost the same for the different selected heights in the flat cases but changes in the urban cases as height changes. Buildings affect the error propagation to large scales inside the canyon, and further influence the roughness sublayer. Therefore, a shorter subrange follows $k^{-5/3}$ after saturation and tends to be flat for large scales.

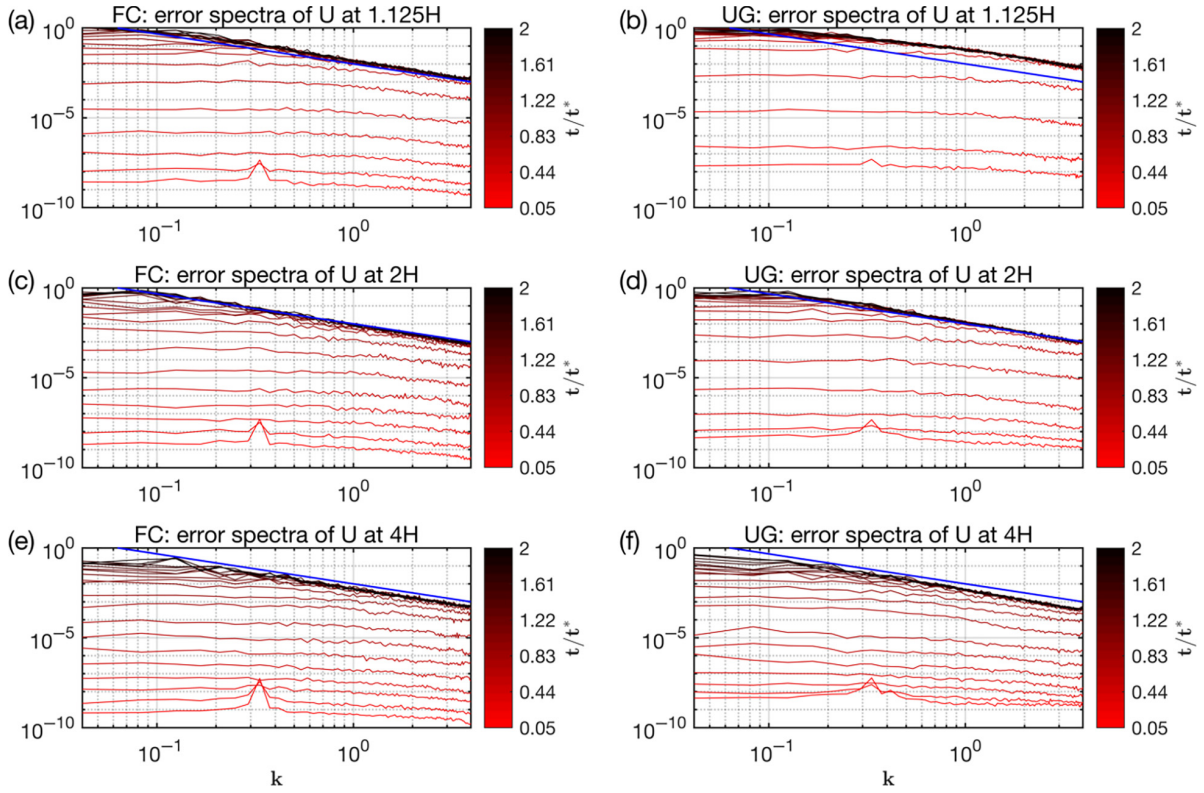

FIGURE 9 Error spectra of streamwise velocity at different heights and time for case (a,c,e) FC and (b,d,f) UG; eddy turnover time: $t^* = z_i/u^*$. The solid line indicates $k^{-5/3}$ [Colour figure can be viewed at wileyonlinelibrary.com]

As height increases, this subrange for urban cases increases and becomes similar to the flat cases. This performance is also consistent with the previous results that the influence of buildings cannot be perceived above $2H$.

Similarly, the ΔE_θ is analyzed and shown in Figures 10 and 11. The scalar error energy propagated from the velocity field is evenly distributed over all scales in the beginning with almost identical magnitude. Similar to the evolution of ΔE_u, the error growth rate for the scalar field is also identical for a wide range of scales prior to saturation, and the smaller scales saturate first. However, the spectral slope of ΔE_θ is different between urban and flat cases, which indicates that building structures can influence the propagation of error among scales. For the flat cases (Figure 10, FK is not shown here because of the similarity with FC), $\Delta E_\theta(k, z)$ at different heights also follows the $k^{-5/3}$ scaling after saturation, like the velocity field. For the urban cases (Figure 11), though, the spectral shape in terms of the power exponent changes with height. It has a smaller absolute value at lower height but becomes $k^{-5/3}$ above $2H$.

From the definition of ΔE_u and ΔE_θ, it is expected that after the twin simulations become essentially uncorrelated for $t \gg T_p$, the error energy spectra exhibit a return-to-turbulence feature (Mukherjee *et al.*, 2016), which attains the $k^{-5/3}$ scaling in the inertial subrange. This could be understood by considering:

$$\int_0^\infty \Delta E_X(k, z, t)dk = \mathrm{Var}\left[X_p - X_c\right] = \mathrm{Var}\left[X_p\right] + \mathrm{Var}\left[X_c\right] \tag{20}$$

where Var represents the variance for X being u or θ. Note that the covariance term is not included in Equation (20) because it is zero as the twin simulations eventually become uncorrelated. Since $\mathrm{Var}\left[X_p\right]$ and $\mathrm{Var}\left[X_c\right]$ can be related to their respective energy spectra, it is expected that the error energy spectra eventually will follow the behavior of the energy spectra $E_{X_c}(k)$. For flat cases, the error energy spectra of u and θ both show a $-5/3$ spectral slope, which is consistent with the $-5/3$ spectral slope expected in the inertial and inertial-convective subrange of the energy spectra of the streamwise velocity and passive scalar (see Figures 12 and A9). However, when an urban canopy is present, E_u still exhibits a $-5/3$ slope, but a spectral slope of -1, which is shallower than $-5/3$ is seen in both ΔE_θ (Figure 11) and E_θ (Figure 12) below $z = 2H$. This is consistent with Batchelor's prediction that the slope is -1 in the canyon (Batchelor, 1959; Warhaft, 2000). In general, passive scalars mixed by complicated turbulent flows within the canopy sublayer can violate the locally isotropic condition, and the spectral slopes of E_θ can also be affected by Reynolds number and turbulent intermittency,

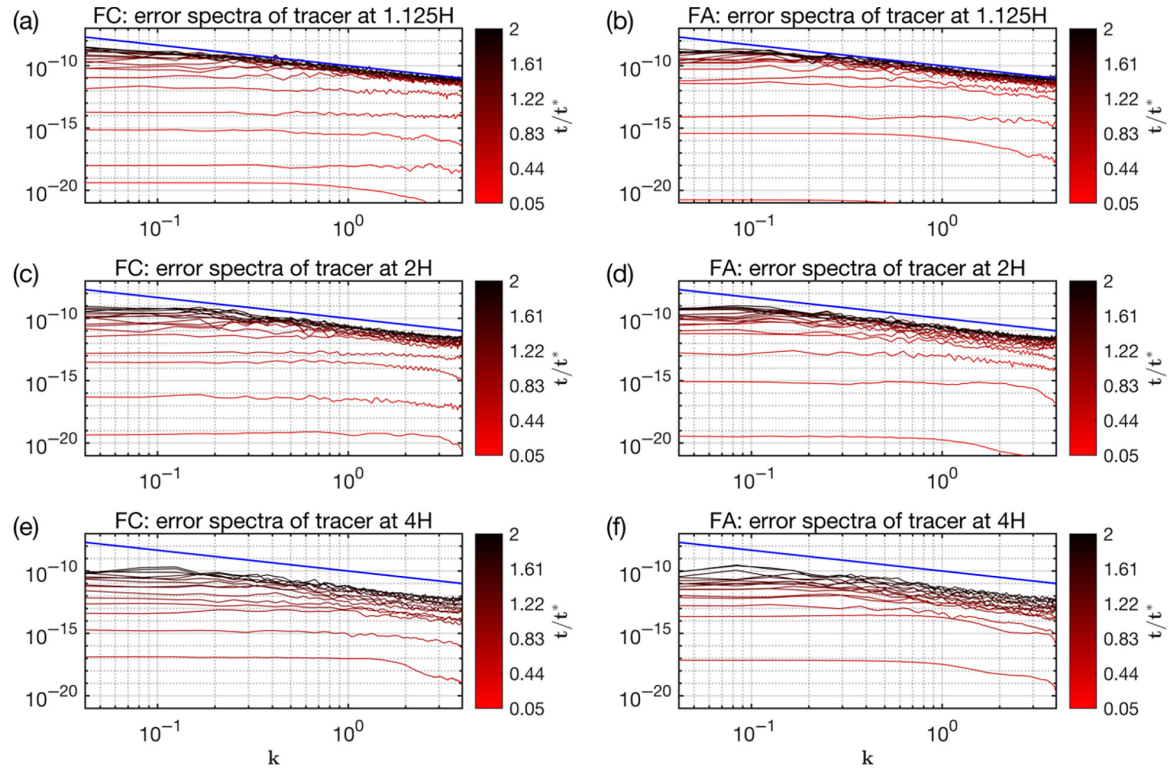

FIGURE 10 Error spectra of the scalar at different heights at different times for case (a,c,e) FC and (b,d,f) FA; eddy turnover time: $t^* = z_i/u^*$. The solid line indicates $k^{-5/3}$ [Colour figure can be viewed at wileyonlinelibrary.com]

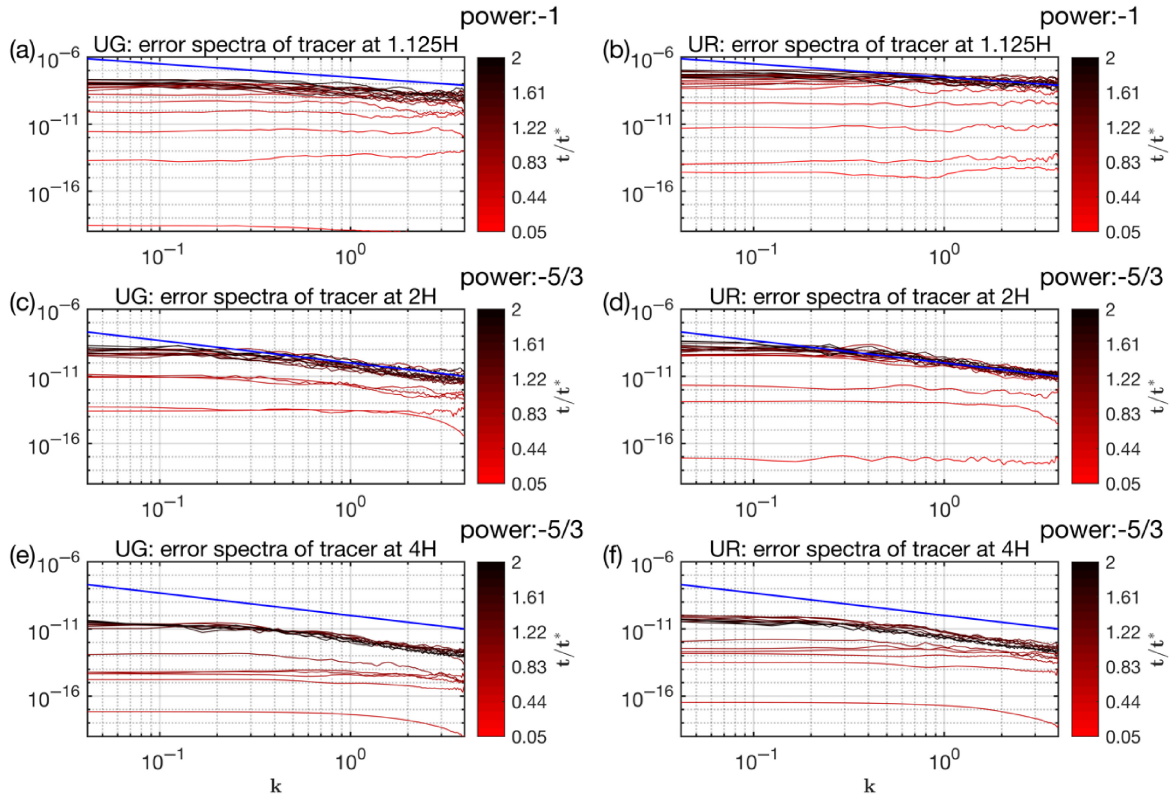

FIGURE 11 Error spectra of the scalar at different heights at different times for case (a,c,e) UG and (b,d,f) UR; eddy turnover time: $t^* = z_i/u^*$. The solid line indicates k^{power} [Colour figure can be viewed at wileyonlinelibrary.com]

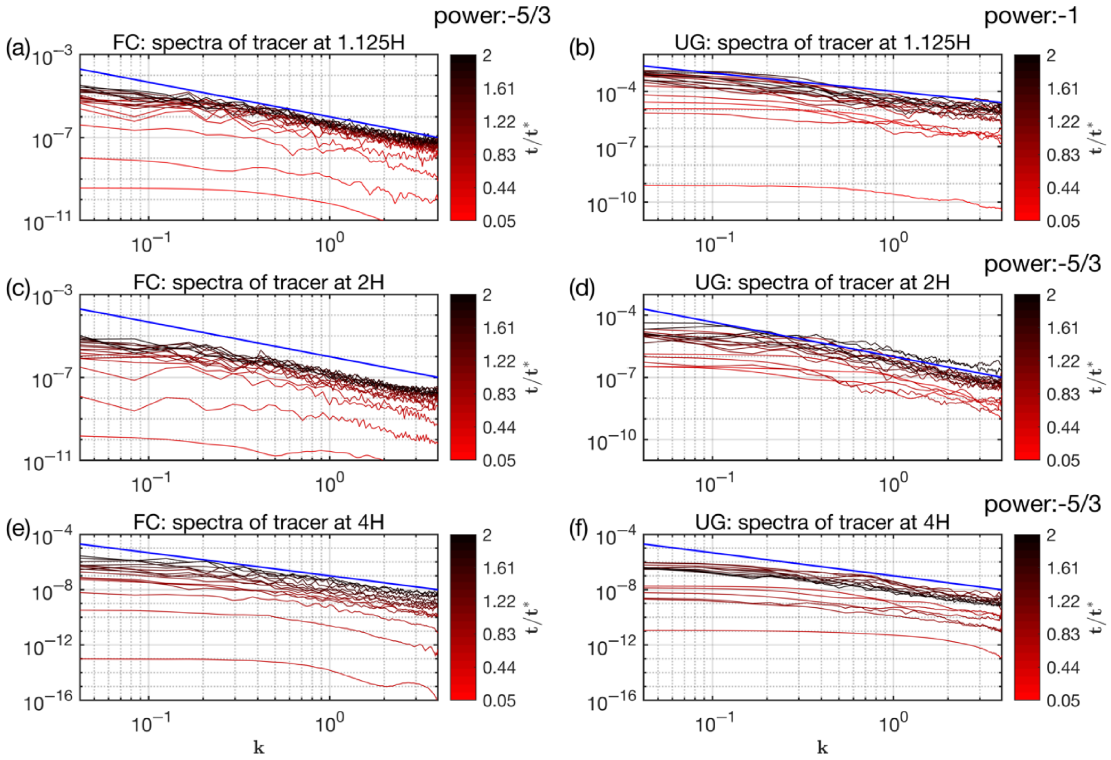

FIGURE 12 Power spectra of the scalar at different heights and times for case (a,c,e) FC and (b,d,f) UG; eddy turnover time: $t^* = z_i/u^*$. The solid line indicates k^{power} [Colour figure can be viewed at wileyonlinelibrary.com]

especially in shear turbulence (Sreenivasan, 1996). Also, a $-5/3$ spectral slope present in E_u (and hence ΔE_u) is not a sufficient condition for the passive scalar to follow similar behavior, since the non-universal large-scale fluctuations can affect the passive scalar in the inertial range. Therefore, the error energy spectrum ΔE_θ at different scales could be highly impacted by these different turbulent structures appearing in the urban cases, leading to distinct errors of u and θ in spectral space.

4 | CONCLUSION

The divergence of both the velocity and scalar fields after a small-amplitude perturbation and the corresponding predictability limit of the passive scalar in the atmospheric boundary layer is studied by simulations conducted at high spatiotemporal resolution and a turbulence-resolving model, the LES. With the identical-twin simulation method, an initial perturbation is added to one of the simulations and the evolution of error is analyzed. Overall, this study highlights that the presence of large roughness structures can lead to diverging behaviors of error propagation, saturation and hence predictability between the velocity and passive scalar in the ASL. The following specific remarks can be made regarding the questions raised in the introduction.

1. The evolution of volume errors and correlation coefficients with time indicates three phases of error growth in both the flat-wall and urban cases for both u and θ. The volume errors of velocity are similar in all cases, but scalar source locations (e.g., rooftop vs near the ground) affect their magnitudes. A significant decrease in correlation coefficients of the passive scalar is seen compared to the high value for velocity. To quantify the predictability limit and analyze the differences among cases for a passive scalar, the NSR is used to resolve when error saturation is reached and how the urban roughness influences it.
2. T_{P} determined from the proposed methods indicates that presence of an urban canopy decreases T_{p} for θ compared to the flat-wall case by around 20%. In addition, a larger amplitude of the initial perturbation further decreases T_{p}, while the predictability time is not sensitive to the frequency of the initial perturbation. Different source locations influence the evolution of the error field locally but do not impact T_{p} quantified with the methods above. The magnitude of T_{p} is affected by the Reynolds number of the LES and thus can be sensitive to resolution.
3. The presence of urban canyons does not significantly change the magnitude of the volume error of velocity, but the scalar error field is affected within the roughness sublayer height. Analysis of the height

dependency of statistics of errors (i.e., the RMSC, RMSP, and NSR) indicates that below $z = 2H$, urban and flat-wall cases show different evolution. This can be attributed to the effects of the same large-scale flow structures in the twin simulations transporting the scalar plume, resulting in an almost synchronized mean concentration and fluctuations or analogously synchronized RMSC and RMSP values. Analysis of the scalar-error energy budget also indicates that advective transports by velocity and the velocity error significantly increase in urban cases compared to the flat-wall cases.

4. Spectral space analysis indicates that beyond $z = 2H$, the urban canopy effect does not affect the error spectra, and thus results are similar to those in the flat-wall cases. However, below $z = 2H$, the velocity and scalar in flat cases both show a $-5/3$ spectral slope but not in urban cases. A shallower than $-5/3$ spectral slope in the scalar error spectra for urban cases implies that a $-5/3$ range in velocity error spectra is not a sufficient condition for similarity of the passive scalar error spectra due to the influence of the non-universal large-scale motions when urban roughnesses are present.

This study is limited by conducting experiments under an idealized domain setup, where a horizontal periodic boundary condition is imposed for both the momentum and passive scalar. Dispersion from a point source, which yields as single plume, can be considered in future studies. The simulations only consider neutrally stratified flows, and the effect of buoyancy is deliberately neglected. Under unstably stratified conditions, for instance, we can conjecture that the coupling between potential temperature and momentum may further change how errors transmit between the velocity and active scalar fields, as backscatters of turbulent kinetic energy from the small-scale eddies to large-scale coherent ones (Zilitinkevich *et al.*, 2021) may facilitate error propagating from velocity to scalar fields. This consequently modifies errors in passive scalar dispersion. Thus, future studies can further explore the predictability limit of both active and passive scalar dispersion under unstable conditions. In addition, the idealized two-dimensional urban street canyons cannot account for the irregularities at different scales (e.g., variations in building heights, street canyon widths, and surface roughness) and how these irregularities impact the predictability remains an open question. However, it can be deduced from results of the current study that error propagation in tracer dispersion is significantly tied to the spatial variations of the mean and variance of scalars, which is dependent on both the street canyon configurations and source locations. Thus, future research may benefit from a systematic sweep of the parameter space of urban form for better understanding.

As a final remark, despite its somewhat idealized setting and theoretical motivation, this study is intended to address an important practical problem, namely the response of the system to a perturbation. Our results may not be directly applicable to the issue of predictability in current NWP models, but will shed light on future NWP models, in which buildings and turbulence down to certain scales will be resolved. Since the analysis increments are not guaranteed to be turbulence-free, at least with existing assimilation systems, results for random perturbations, such as those of this study, are relevant to data assimilation in practice when higher temporal and spatial resolution measurements become available.

AUTHOR CONTRIBUTIONS

Yanle Lu: Conceptualization; data curation; formal analysis; visualization; writing – original draft; writing – review and editing. **Luca Delle Monache:** Conceptualization. **Jeffrey Weil:** Conceptualization; supervision; writing – original draft. **Keith Ngan:** Conceptualization; supervision; writing – review and editing. **Qi Li:** Conceptualization; data curation; formal analysis; funding acquisition; supervision; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

The authors would like to thank three anonymous reviewers for their suggestions and comments. The authors acknowledge Dr. Gao Shang for his help during the revision process. Qi Li acknowledges support from the US National Science Foundation (NSF-CAREER-2143664, NSF-AGS-2028644, NSF-CBET-2028842) and computational resources from the National Center for Atmospheric Research (UCOR-0049).

CONFLICT OF INTEREST STATEMENT

All authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

ORCID

Qi Li https://orcid.org/0000-0003-4435-6220

REFERENCES

Batchelor, G.K. (1959) Small-scale variation of convected quantities like temperature in turbulent fluid Part 1. General discussion and the case of small conductivity. *Journal of Fluid Mechanics*, 5, 113–133. https://doi.org/10.1017/S002211205900009X.

Bou-Zeid, E., Meneveau, C. and Parlange, M. (2005) A scale-dependent Lagrangian dynamic model for large eddy simulation of complex turbulent flows. *Physics of Fluids*, 17, 25105. https://doi.org/10.1063/1.1839152.

Brunet, Y., Finnigan, J. and Raupach, M. (1994) A wind tunnel study of air flow in waving wheat: Single-point velocity statistics. *Boundary-Layer Meteorology*, 70, 95–132. https://doi.org/10.1007/BF00712525.

Brutsaert, W. (1965) Equations for Vapor Flux as a Fully Turbulent Diffusion Process Under Diabatic Conditions. *International Association of Scientific Hydrology Bulletin*, 10, 11–21. https://doi.org/10.1080/02626666509493386.

Caulton, D.R., Li, Q., Bou-Zeid, E., Fitts, J.P., Golston, L.M., Pan, D., Lu, J., Lane, H.M., Buchholz, B., Guo, X., McSpiritt, J., Wendt, L. and Zondlo, M.A. (2018) Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods. *Atmospheric Chemistry and Physics*, 18, 15145–15168. https://doi.org/10.5194/acp-18-15145-2018.

Chester, S., Meneveau, C. and Parlange, M.B. (2007) Modeling turbulent flow over fractal trees with renormalized numerical simulation. *Journal of Computational Physics*, 225, 427–448. https://doi.org/10.1016/j.jcp.2006.12.009.

Cui, Z., Cai, X. and Baker, C.J. (2004) Large-eddy simulation of turbulent flow in a street canyon. *Quarterly Journal of the Royal Meteorological Society*, 130, 1373–1394. https://doi.org/10.1256/qj.02.150.

Van Driest, E.R. (1959) Convective heat transfer in gases. *High Speed Aerodynamics and Jet Propulsion*, 5, 339–427.

Eckmann, J.-P. and Ruelle, D. (1985) Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57, 617–656. https://doi.org/10.1103/RevModPhys.57.617.

Fackrell, J.E. and Robins, A.G. (1982) Concentration fluctuations and fluxes in plumes from point sources in a turbulent boundary layer. *Journal of Fluid Mechanics*, 117, 1–26. https://doi.org/10.1017/S0022112082001499.

Finnigan, J.J., Clement, R., Malhi, Y., Leuning, R. and Cleugh, H.A. (2003) A re-evaluation of long-term flux measurement techniques Part I: averaging and coordinate rotation. *Boundary-Layer Meteorology*, 107, 1–48. https://doi.org/10.1023/A:1021554900225.

Hacker, J.P. (2010) Spatial and temporal scales of boundary layer wind predictability in response to small-amplitude land surface uncertainty. *Journal of the Atmospheric Sciences*, 67, 217–233. https://doi.org/10.1175/2009JAS3162.1.

Horst, T.W. (1979) Lagrangian similarity modeling of vertical diffusion from a ground-level source. *Journal of Applied Meteorology and Climatology*, 18, 733–740. https://doi.org/10.1175/1520-0450(1979)018<0733:LSMOVD>2.0.CO;2.

Inagaki, A., Castillo, M.C.L., Yamashita, Y., Kanda, M. and Takimoto, H. (2012) Large-Eddy Simulation of Coherent Flow Structures within a Cubical Canopy. *Boundary-Layer Meteorology*, 142, 207–222. https://doi.org/10.1007/s10546-011-9671-8.

Kalnay, E. and Eugenia, K. (2003) *Atmospheric Modeling, Data Assimilation and Predictability*. United Kingdom: Cambridge University Press.

Kays, W.M. and Crawford, M.E. (1993) *Convective Heat and Mass Transfer*. New York: McGraw-Hill, pp. 255–282.

Laprise, R., Varma, M.R., Denis, B., Caya, D. and Zawadzki, I. (2000) Predictability of a Nested Limited-Area Model. *Monthly Weather Review*, 128, 4149–4154. https://doi.org/10.1175/1520-0493(2000)129<4149:POANLA>2.0.CO;2.

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

Lee, G.-J., Muñoz-Esparza, D., Yi, C. and Choe, H.J. (2019) Application of the Cell Perturbation Method to Large-Eddy Simulations of a Real Urban Area. *Journal of Applied Meteorology and Climatology*, 58, 1125–1139. https://doi.org/10.1175/JAMC-D-18-0185.1.

Leith, C.E. and Kraichnan, R.H. (1972) Predictability of Turbulent Flows. *Journal of the Atmospheric Sciences*, 29, 1041–1058. https://doi.org/10.1175/1520-0469(1972)029<1041:potf>2.0.co;2.

LESGO. (2020) LESGO: a parallel pseudo-spectral large-eddy simulation code. https://lesgo.me.jhu.edu.

Li, Q. and Bou-Zeid, E. (2019) Contrasts between momentum and scalar transport over very rough surfaces. *Journal of Fluid Mechanics*, 880, 32–58. https://doi.org/10.1017/jfm.2019.687.

Li, Q., Bou-Zeid, E. and Anderson, W. (2016a) The impact and treatment of the Gibbs phenomenon in immersed boundary method simulations of momentum and scalar transport. *Journal of Computational Physics*, 310, 237–251. https://doi.org/10.1016/j.jcp.2016.01.013.

Li, Q., Bou-Zeid, E., Anderson, W., Grimmond, S. and Hultmark, M. (2016b) Quality and reliability of LES of convective scalar transfer at high Reynolds numbers. *International Journal of Heat and Mass Transfer*, 102, 959–970. https://doi.org/10.1016/j.ijheatmasstransfer.2016.06.093.

Li, Q., Gentine, P., Mellado, J.P. and McColl, K.A. (2018) Implications of nonlocal transport and conditionally averaged statistics on Monin–Obukhov similarity theory and townsend's attached eddy hypothesis. *Journal of the Atmospheric Sciences*, 75, 3403–3431. https://doi.org/10.1175/JAS-D-17-0301.1.

Li, Q., Bou-Zeid, E., Grimmond, C., Zilitinkevich, S. and Katul, G. (2020) Revisiting the relation between momentum and scalar roughness lengths of urban surfaces. *Quarterly Journal of the Royal Meteorological Society*, 146, 3144–3164. https://doi.org/10.1002/qj.3839.

Lo, K.W. and Ngan, K. (2015) Predictability of turbulent flow in street canyons. *Boundary-Layer Meteorology*, 156, 191–210. https://doi.org/10.1007/s10546-015-0014-z.

Lorenz, E.N. (1963) Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20, 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Lorenz, E.N. (1965) A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17, 321–333. https://doi.org/10.3402/tellusa.v17i3.9076.

Lorenz, E.N. (1969) The predictability of a flow which possesses many scales of motion. *Tellus*, 21, 289–307. https://doi.org/10.3402/tellusa.v21i3.10086.

Millward-Hopkins, J.T., Tomlin, A.S., Ma, L., Ingham, D. and Pourkashanian, M. (2012) The predictability of above roof wind resource in the urban roughness sublayer. *Wind Energy*, 15, 225–243. https://doi.org/10.1002/we.463.

Mukherjee, S., Schalkwijk, J.M. and Jonker, H.J.J. (2016) Predictability of dry convective boundary layers: An LES study. *Journal of*

the Atmospheric Sciences, 73, 2715–2727. https://doi.org/10.1175/JAS-D-15-0206.1.

Ngan, K. and Eperon, G.E. (2012) Middle atmosphere predictability in a numerical weather prediction model: Revisiting the inverse error cascade. *Quarterly Journal of the Royal Meteorological Society*, 138, 1366–1378. https://doi.org/10.1002/qj.984.

Ngan, K. and Lo, K.W. (2017) Linear error dynamics for turbulent flow in urban street canyons. *Journal of Applied Meteorology and Climatology*, 56, 1195–1208. https://doi.org/10.1175/JAMC-D-16-0173.1.

Orszag, S.A. and Pao, Y.-H. (1975) Numerical Computation of Turbulent Shear Flows11This work is supported by Fluid Dynamics Branch, Office of Naval Research under Navy Contract No. NOO0I4-72-C-O355, ONR Task No. NR 062-464 and the Climatic Impact Assessment Program, Department of Transport. In: Frenkiel, F.N. and R.E.B.T.-A. in Munn, G. (Eds.) *Turbulent Diffusion in Environmental Pollution*, Vol. 18 of Elsevier, pp. 225–236.

Raupach, M.R., Finnigan, J.J. and Brunet, Y. (1996) In: Garratt, J.R. and Taylor, P.A. (Eds.) *Coherent Eddies and Turbulence in Vegetation Canopies: The Mixing-Layer Analogy BT – Boundary-Layer Meteorology 25th Anniversary Volume, 1970–1995: Invited Reviews and Selected Contributions to Recognise Ted Munn's Contribution as Editor over the Past 25*. Netherlands: Springer, pp. 351–382.

Roth, M. (2000) Review of atmospheric turbulence over cities. *Quarterly Journal of the Royal Meteorological Society*, 126, 941–990. https://doi.org/10.1256/smsqj.56408.

Ruelle, D. and Takens, F. (1971) On the nature of turbulence. *Communications in Mathematical Physics*, 20, 167–192. https://doi.org/10.1007/BF01646553.

Schalkwijk, J., Jonker, H.J.J., Siebesma, A.P. and Van Meijgaard, E. (2015) Weather forecasting using GPU-based large-eddy simulations. *Bulletin of the American Meteorological Society*, 96, 715–723. https://doi.org/10.1175/BAMS-D-14-00114.1.

Smith, L.A., Ziehmann, C. and Fraedrich, K. (1999) Uncertainty dynamics and predictability in chaotic systems. *Quarterly Journal of the Royal Meteorological Society*, 125, 2855–2886. https://doi.org/10.1002/qj.49712556005.

Sreenivasan, K.R. (1996) The passive scalar spectrum and the Obukhov-Corrsin constant. *Physics of Fluids*, 8, 189–196. https://doi.org/10.1063/1.868826.

Sullivan, P.P. and Patton, E.G. (2011) The effect of mesh resolution on convective boundary layer statistics and structures generated by large-eddy simulation. *Journal of the Atmospheric Sciences*, 68, 2395–2415. https://doi.org/10.1175/JAS-D-10-05010.1.

Sullivan, P.P., Banner, M.L., Morison, R.P. and Peirson, W.L. (2018) Turbulent flow over steep steady and unsteady waves under strong wind forcing. *Journal of Physical Oceanography*, 48, 3–27. https://doi.org/10.1175/JPO-D-17-0118.1.

Sykes, R.I. (1988) Concentration Fluctuations in Dispersing Plumes. In: Venkatram, A. and Wyngaard, J.C. (Eds.) *Lectures on Air Pollution Modeling*. Boston, MA: American Meteorological Society. https://doi.org/10.1007/978-1-935704-16-4_8

Tomas, J.M., Pourquie, M.J.B.M. and Jonker, H.J.J. (2015) The influence of an obstacle on flow and pollutant dispersion in neutral and stable boundary layers. *Atmospheric Environment*, 113, 236–246. https://doi.org/10.1016/j.atmosenv.2015.05.016.

Tomas, J.M., Pourquie, M.J.B.M. and Jonker, H.J.J. (2016) Stable stratification effects on flow and pollutant dispersion in boundary layers entering a generic urban environment. *Boundary-Layer Meteorology*, 159, 221–239. https://doi.org/10.1007/s10546-015-0124-7.

Tribbia, J.J. and Baumhefner, D.P. (2004) Scale interactions and atmospheric predictability: an updated perspective. *Monthly Weather Review*, 132, 703–713. https://doi.org/10.1175/1520-0493(2004)132<0703:SIAAPA>2.0.CO;2.

Tseng, Y.-H., Meneveau, C. and Parlange, M.B. (2006) Modeling flow around bluff bodies and predicting urban dispersion using large eddy simulation. *Environmental Science & Technology*, 40, 2653–2662. https://doi.org/10.1021/es051708m.

van Ulden, A.P. (1978) Simple estimates for vertical diffusion from sources near the ground. *Atmospheric Environment*, 12, 2125–2129. https://doi.org/10.1016/0004-6981(78)90167-1.

Warhaft, Z. (2000) Passive scalars in turbulent flows. *Annual Review of Fluid Mechanics*, 32, 203–240. https://doi.org/10.1146/annurev.fluid.32.1.203.

Weil, J.C., Sykes, R.I. and Venkatram, A. (1992) Evaluating air-quality models: review and outlook. *Journal of Applied Meteorology and Climatology*, 31, 1121–1145. https://doi.org/10.1175/1520-0450(1992)031<1121:EAQMRA>2.0.CO;2.

Wiersema, D.J., Lundquist, K.A. and Chow, F.K. (2020) Mesoscale to microscale simulations over complex terrain with the immersed boundary method in the weather research and forecasting model. *Monthly Weather Review*, 148, 577–595. https://doi.org/10.1175/MWR-D-19-0071.1.

Yang, X.I.A. and Abkar, M. (2018) A hierarchical random additive model for passive scalars in wall-bounded flows at high Reynolds numbers. *Journal of Fluid Mechanics*, 842, 354–380. https://doi.org/10.1017/jfm.2018.139.

Zeng, X. and Pielke, R.A. (1993) Error-growth dynamics and predictability of surface thermally induced atmospheric flow. *Journal of the Atmospheric Sciences*, 50, 2817–2844. https://doi.org/10.1175/1520-0469(1993)050<2817:EGDAPO>2.0.CO;2.

Zeng, X. and Pielke, R.A. (1995) Further study on the predictability of landscape-induced atmospheric flow. *Journal of the Atmospheric Sciences*, 52, 1680–1698. https://doi.org/10.1175/1520-0469(1995)052<1680:fsotpo>2.0.co;2.

Zilitinkevich, S., Kadantsev, E., Repina, I., Mortikov, E. and Glazunov, A. (2021) Order out of chaos: Shifting paradigm of convective turbulence. *Journal of the Atmospheric Sciences*, 78, 3925–3932. https://doi.org/10.1175/jas-d-21-0013.1.

How to cite this article: Lu, Y., Monache, L.D., Weil, J., Ngan, K. & Li, Q. (2023) Predictability of passive scalar dispersion in atmospheric surface layers with urban-like roughness: A large-eddy simulations study. *Quarterly Journal of the Royal Meteorological Society*, 149(752), 994–1017. Available from: https://doi.org/10.1002/qj.4445

APPENDIX A

The almost identical trends and plateau values show the insensitivity of error fields growth of perturbation wavenumber k.

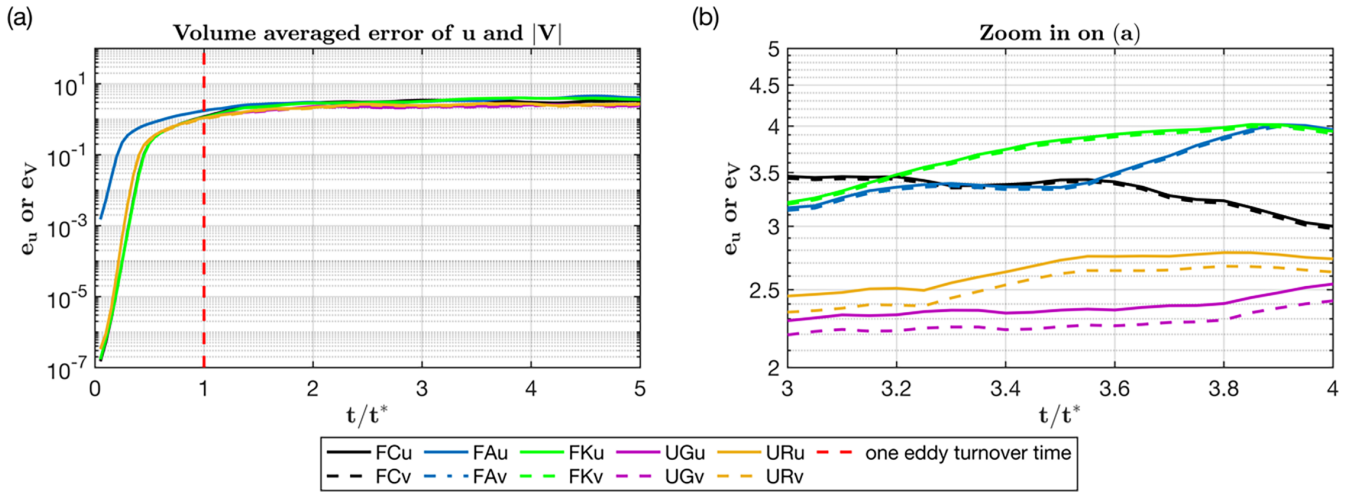

FIGURE A1 Volume- averaged error of twin simulations of the streamwise component of velocity, u, and velocity magnitude, V, at different times, where one eddy turnover time is indicated by $t^* = z_i/u_*$. The red dotted line indicates t/t^*=1. Panel b is the zoomed-in version for (a), indicating that some trivial differences do exist [Colour figure can be viewed at wileyonlinelibrary.com]

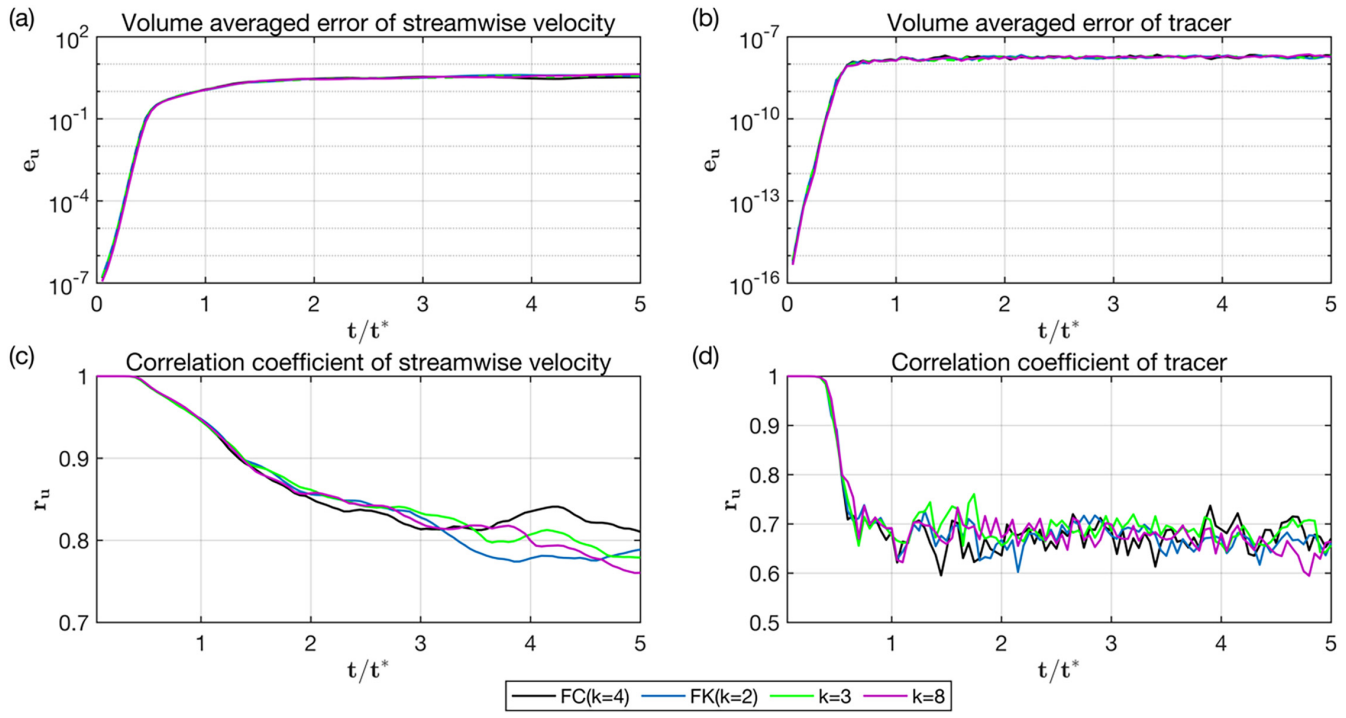

FIGURE A2 Volume-averaged error of twin simulations of (a) the streamwise velocity, u, and (b) the passive scalar, θ; (c) correlation coefficient of twin simulations of the streamwise velocity, u, and (d) of changes with time of the passive scalar, θ. $t^* = z_i/u_*$ [Colour figure can be viewed at wileyonlinelibrary.com]

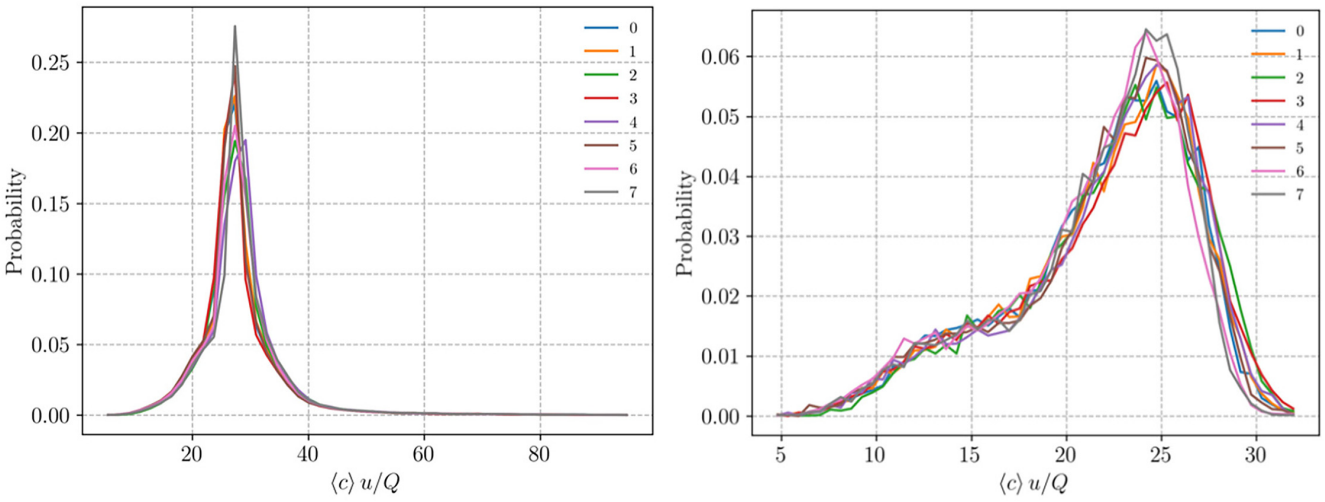

FIGURE A3 Probability distribution for spanwise normalized concentration within the canopy (left) and at the shear layer above the canopy (right). Each line corresponds to each realization [Colour figure can be viewed at wileyonlinelibrary.com]

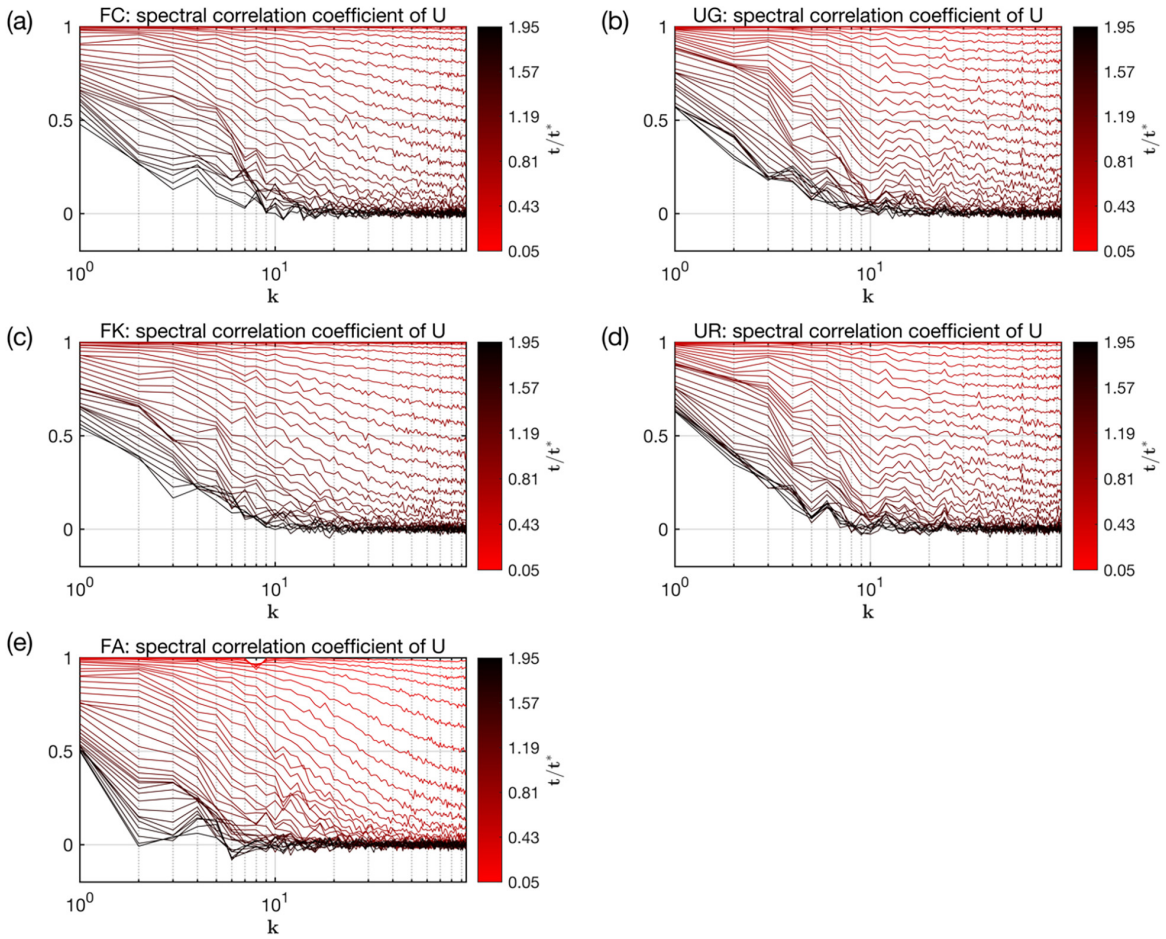

FIGURE A4 Vertically averaged (above roof level) spectral correlation coefficient of the streamwise velocity change with time for case (a) FC, (b) UG, (c) FK, (d) UR and (e) FA [Colour figure can be viewed at wileyonlinelibrary.com]

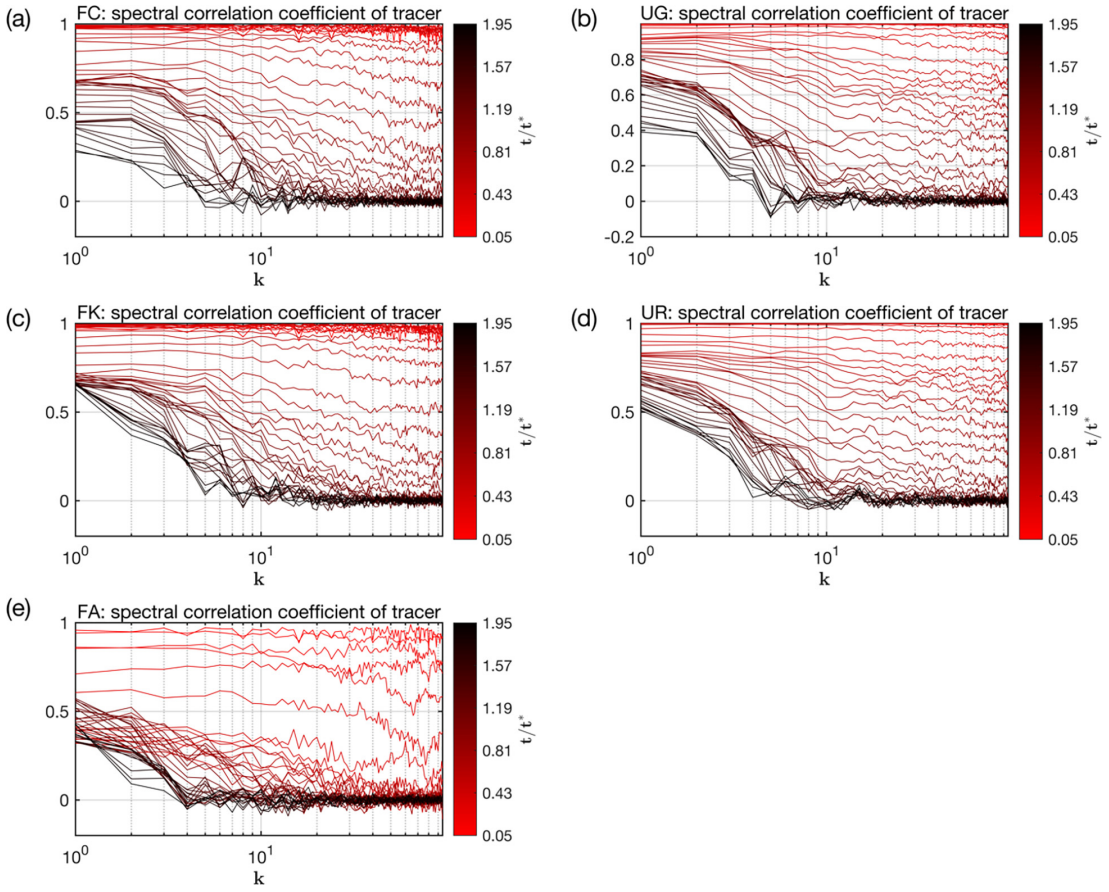

FIGURE A5 Vertically averaged (above roof level) spectral correlation coefficient of the scalar change with time for case (a) FC, (b) UG, (c) FK, (d) UR and (e) FA [Colour figure can be viewed at wileyonlinelibrary.com]

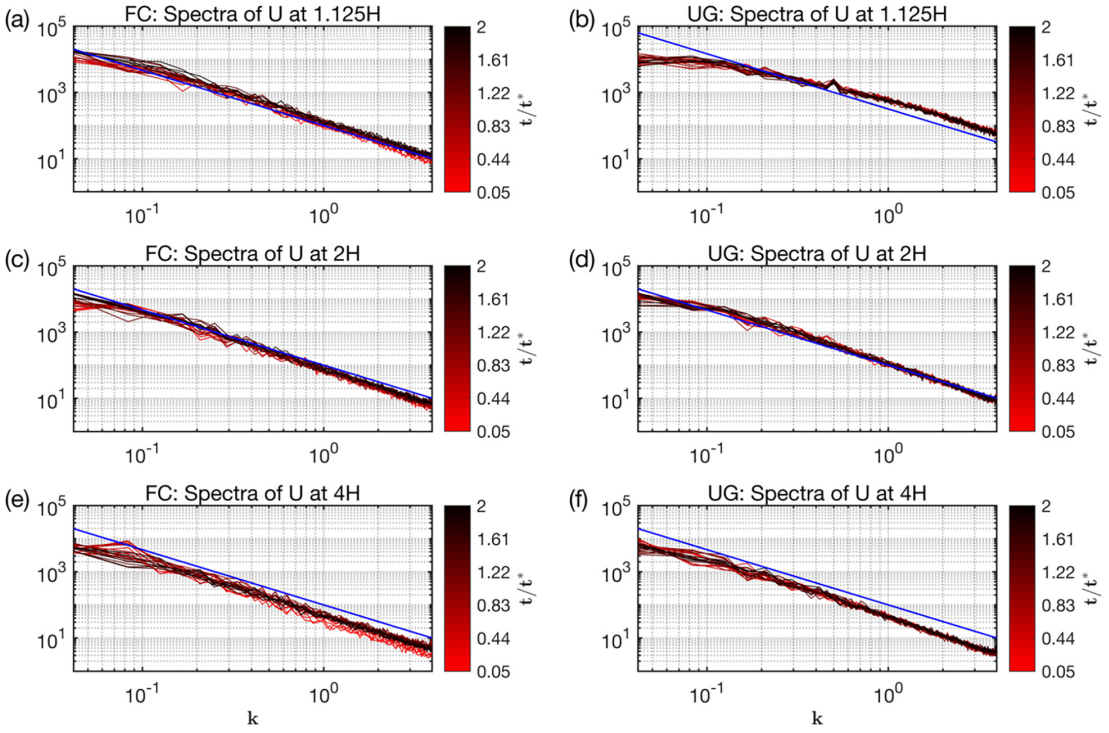

FIGURE A6 Power spectra of streamwise velocity at different heights and times for case (a,c,e) FC and (b,d,f) UG; eddy turnover time: $t^* = z_i/u^*$ [Colour figure can be viewed at wileyonlinelibrary.com]

We conducted a doubled resolution simulation for urban cases to highlight that the differences between velocity and tracer error growth are resolution independent. We name them UGH (urban ground high-resolution) case and URH (urban roof high-resolution). The results are as expected. The volume averaged error, and the error spectra are shown below:

The results for the same resolution are overlapped in Figure 6a, which make it looks like only two lines. The doubled resolution cases need less time to reach the plateau value.

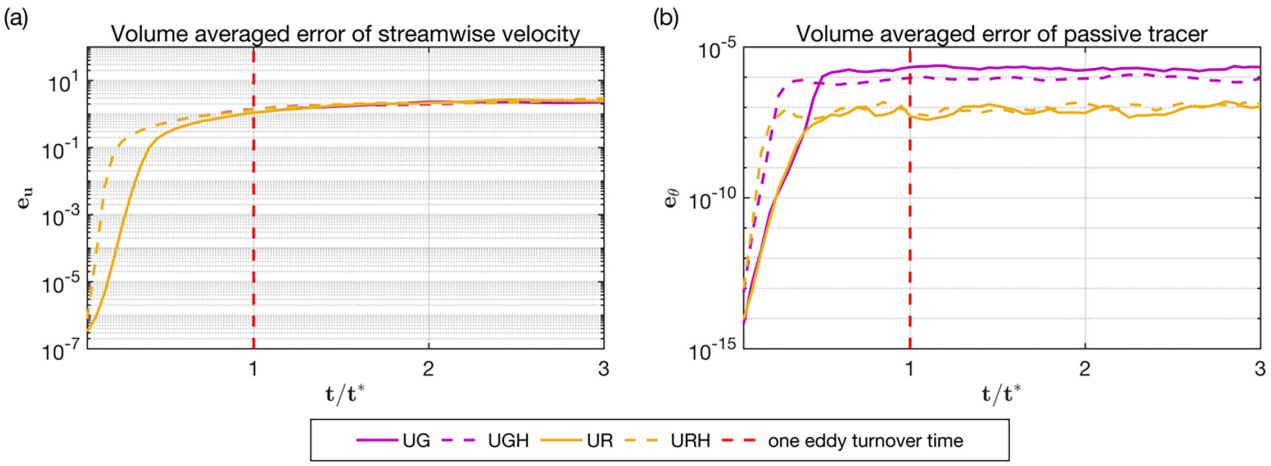

FIGURE A7 Volume-averaged error for urban cases [Colour figure can be viewed at wileyonlinelibrary.com]

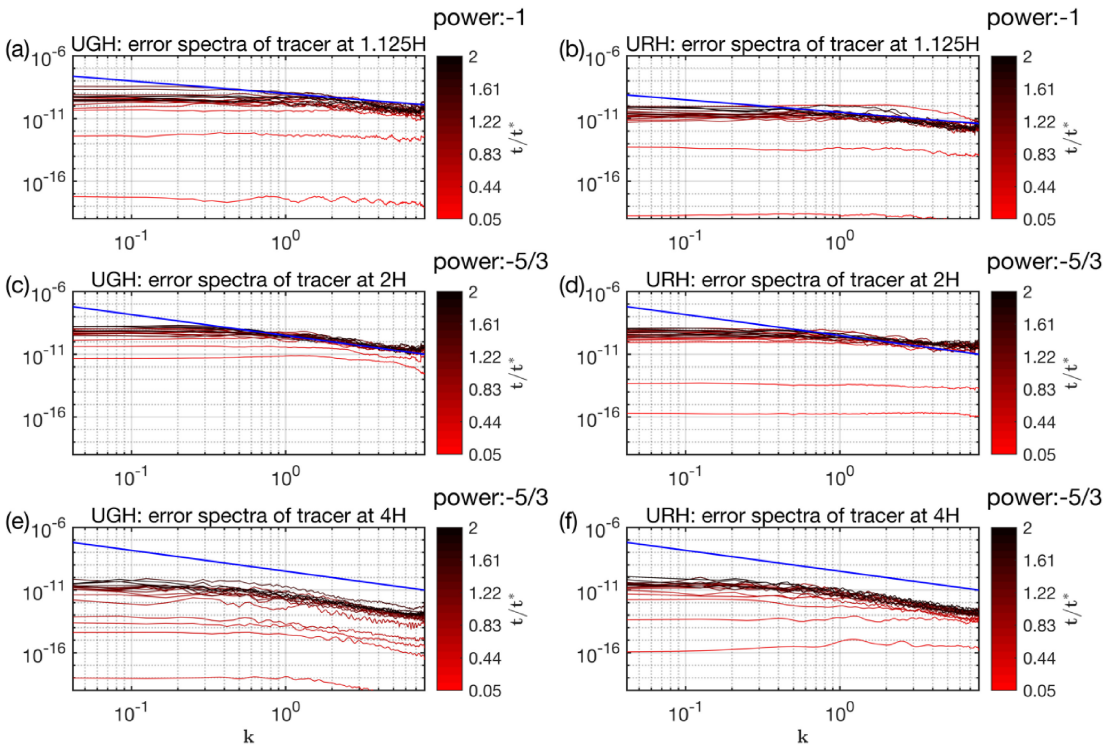

FIGURE A8 Error spectra of the scalar at different heights and times for case (a,c,e) UGH and (b,d,f) URH; eddy turnover time: $t^* = z_i/u^*$. The blue solid line indicates k^{power} [Colour figure can be viewed at wileyonlinelibrary.com]

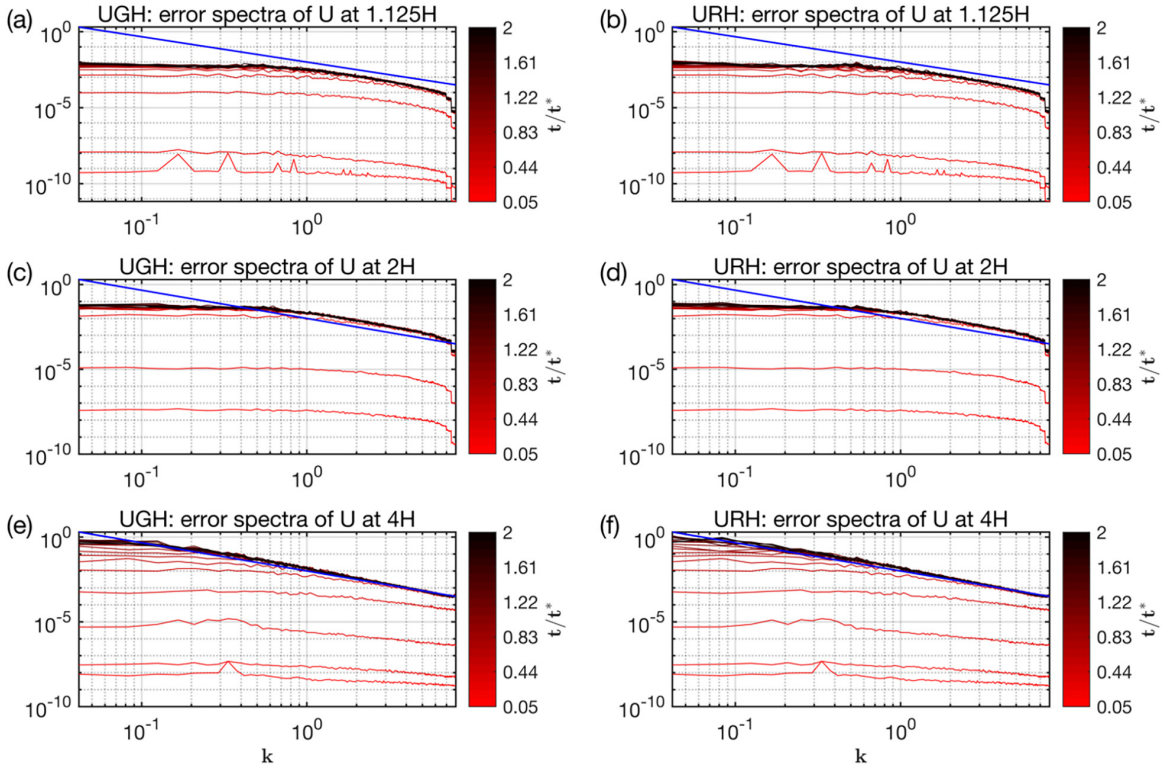

FIGURE A9 Error spectra of the streamwise velocity at different heights and times for case (a,c,e) UGH and (b,d,f) URH; eddy turnover time: $t^* = z_i/u^*$. The blue solid line indicates $k^{-5/3}$ [Colour figure can be viewed at wileyonlinelibrary.com]