Boundary-Layer Meteorology (2024) 190:43
https://doi.org/10.1007/s10546-024-00880-y

RESEARCH ARTICLE

Assessing the Performance of Flux Imbalance Prediction Models Using Large Eddy Simulations Over Heterogeneous Land Surfaces

Lijie Zhang[1] · **Stefan Poll**[1,2,3] · **Stefan Kollet**[1,2]

Received: 3 March 2024 / Accepted: 16 August 2024 / Published online: 24 September 2024

Abstract

Accurate representation of heat fluxes is crucial for understanding land–atmosphere interactions and improving atmospheric simulations. However, a common issue arises with flux imbalance, where the measured turbulent heat flux tends to be underestimated due to the non-local effects of atmospheric secondary circulations. This study evaluated four flux imbalance prediction models by analyzing data from large eddy simulations performed over heterogeneous land surfaces. For that, a checkerboard pattern of soil moisture was used to define the lower boundary conditions for the atmosphere, across heterogeneity scales ranging from 50 m to 2.4 km. The results show that the selected models can effectively predict flux imbalance when provided with proper semi-empirical factors. The presence of two distinct secondary circulations, thermally-induced mesoscale circulation and turbulent organized structures, account for the nonlinear effect of the heterogeneity scale on the flux imbalance, but it does not affect the performance of the selected models. This study suggests that the flux imbalance prediction models are useful for improving e.g. eddy-covariance measurements. Additionally, a quadrant analysis showed an increasing difference between ejections and sweeps with height, which explains the decrease and increase of the turbulent heat flux and flux imbalance, respectively, and underscores the importance of accounting for vertical variations in turbulent fluxes to represent atmospheric processes accurately.

Keywords Flux imbalance · Large-eddy simulation · Secondary circulations · Quadrant analysis

---

✉ Lijie Zhang
l.zhang@fz-juelich.de

[1] Institute of Bio and Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich, 52428 Jülich, Germany

[2] Center for High-Performance Scientific Computing in Terrestrial Systems, Geoverbund ABC/J, 52428 Jülich, Germany

[3] SimDataLab Terrestrial Systems, Jülich Supercomputing Centre (JSC), 52428 Jülich, Germany

1 Introduction

The eddy-covariance (EC) technique serves as a tool to determine the exchange of energy, water, and trace gases at the Land–Atmosphere (L–A) interface, improving our understanding of the dynamics and the feedback (Baldocchi 2003, 2020; Pastorello et al. 2020). To date, data from over 900 EC stations worldwide are available through the FLUXNET2015 Dataset (Pastorello et al. 2020), covering a diverse range of climates and land covers (https://fluxnet.org/sites/site-summary/, latest access on 30th May 2024). With the establishment of networks like AmeriFlux (Novick et al. 2018; Chu et al. 2023), EuroFlux (Aubinet et al. 1999; Franssen et al. 2010), and AsiaFlux (Yamamoto et al. 2005), numerous additional stations currently being installed, providing a benchmark for the development, validation, and prediction of Earth system models (Baldocchi et al. 2001; Rebmann et al. 2018; Pastorello et al. 2020).

A significant challenge in EC measurement arises due to the discrepancy between the measured sensible (H) and latent (LE) heat fluxes and the available energy (net radiation Rn minus the ground heat flux G), leading to the surface energy balance (SEB) closure problem (i.e., $H + LE < Rn$ - G) (Foken 2008; Foken et al. 2010; Mauder et al. 2020). The reason for this inconsistency is still debated. However, it is generally agreed upon that the inherent assumptions of the EC method, such as Taylor's frozen turbulence hypothesis (Mauder et al. 2013), statistical stationarity (Franssen et al. 2010), and horizontal homogeneity (Eder et al. 2015), are often not fulfilled, especially over complex terrains (Mauder et al. 2020). A degree of underestimation of sensible heat flux between 10–30% has been consistently reported across sites (Twine et al. 2000; Wilson et al. 2002; Barr et al. 2006; Franssen et al. 2010; Stoy et al. 2013). The possible sources of error remain a topic of discussion in the community, including but not limited to measurement errors, missed additional terms related to energy conversion and storage, inadequate sampling over a finite mean period, phase lag existing between vertical wind and water vapour, among others. For more details, interested readers are referred to the review papers (Foken 2008; Mauder et al. 2020; Sun et al. 2021).

Many studies aimed to identify the key factors impacting the flux imbalance (FI, a ratio of nonlocal scale flux to the total reference flux) through the analysis of EC data, considering direct measurement states and prognostic variables, including wind speed ($U = (u^2 + v^2)^{1/2}$, where u and v are the longitudinal and lateral velocity components along the cartesian coordinates x and y, respectively), time average interval (T), friction velocity $u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2\right)^{1/4}$, convective velocity scale ($w_* = \left[\frac{gz_i}{\overline{\theta_v}}\left(\overline{w'\theta'}\right)_s\right]^{1/3}$, g is the gravitational acceleration, θ_v is the potential temperature and z_i is the boundary layer height), stability parameters (z/L or z_i/L, L *is* the Obukhov length) and others (Zhou et al. 2019). A diminished FI is frequently observed with higher wind speeds or friction velocities, particularly during unstable weather conditions (Franssen et al. 2010). Furthermore, an increase in the averaging time also results in a decrease in the FI, as the averaging operator acts as a high-pass filter. However, the averaging time is limited by the requirement of Taylor's frozen turbulence hypothesis for the stationarity, typically in 15, 30, or 60 min (Mauder et al. 2013). Conversely, FI is often reported to increase with an increase in the measurement height and the atmospheric instability (Twine et al. 2000; Wilson et al. 2002; Finnigan et al. 2003; Franssen et al. 2010; Stoy et al. 2013; Zhou and Li 2019). It was also reported that there was a phase difference between the vertical wind velocity and scalar states. With a greater phase difference, there is a greater FI (Gao et al. 2016, 2017).

Extensive research efforts have been devoted to large eddy simulation (LES) to explore the L–A interactions (Shen and Leclerc 1994; Gopalakrishnan and Avissar 2000; Albertson et al. 2001; Maronga and Raasch 2013; Patton et al. 2016) and the *FI* (Kanda et al. 2004; Inagaki et al. 2006; Steinfeld et al. 2007; Huang et al. 2008; De Roo et al. 2018; Zhou et al. 2019, 2023; Margairaz et al. 2020a, b; Wanner et al. 2022a). Several semi-empirical *FI* prediction models have been proposed based on LES. For instance, Huang et al. (2008), hereafter H08, proposed a *FI* prediction model as a function of stability parameters (u_*/w_*), and the ratio of measurement height to the boundary layer height (z/z_i) over a homogeneous land surface. The H08 model is based on a simple shape function and can capture the *FI* in their simulation within the range of 0.3 z_i to 0.5 z_i. An evaluation study by Eder et al. (2014) based on observations in the Bavarian Alps and the Pre-Alps indicates that the H08 may not be able to accurately account for energy balance in complex terrain. Nevertheless, further testing of its validity is still required. In our study, we conducted an evaluation of the H08 based on the LES output, demonstrating the significance of measurement height in *FI* estimates.

With the advancement in computational power, De Roo et al. (2018) build on the work of H08 by using nesting techniques to increase the vertical resolution to 2 m (Hellsten et al. 2021), and fitting the *FI* prediction model based on the data within the atmospheric surface layer (ASL), which is roughly 0.1 z_i (Stull 1988), hereafter De18. Additionally, De18 has been successfully applied to EC data from three sites, including a beech forest site in Denmark and two pre-alpine grass sites in Germany, demonstrating a promising outcome for reducing the *FI* and closing the surface energy balance (Mauder et al. 2021). Wanner et al. (2022a, b) extended the model, hereafter W22, to account for land surface heterogeneity by incorporating the thermal heterogeneity parameter (Margairaz et al. 2020b). The W22 model was fitted to the LES data, accounting for varying heterogeneity scales. This process resulted in the creation of semi-empirical *FI* prediction equations, each with its own set of parameters corresponding to the different heterogeneity scales.

Based on the cospectral analysis of the vertical wind velocity and potential temperature, Zhou et al. (2019) proposed a conceptual model to explain the *FI* with a semi-empirical formation of four variables, including the stability parameter ($-z_i/L$, note that $-z_i/L = -\kappa \frac{w_*^3}{u_*^3}$ under unstable conditions, with κ being the von Kármán constant of 0.40), wind speed (U), time average interval (T) and integral length scale of vertical velocity ($l_w = \int_0^\infty R(r)dr$, with $R(r)$ as a autocorrelation function), providing a potential approach to diagnose the FI. A further study by Zhou et al. (2023) compared the performance of H08, De18, and Z19 using LES data over a 1-D stripe heterogeneity pattern with prescribed sensible and latent heat fluxes. Zhou et al. (2023) suggest that the *FI* prediction model exhibits inadequate performance at the single-pixel scale. However, integrating the footprint model enhances its performance. Among the three models (H08, De18, Z19), Z19 demonstrates superior performance compared to H08 and De18 with their data. In addition, Z19 can be applied under all atmospheric stability conditions, whereas the other two models are only valid under unstable conditions.

Recent studies have shown that the *FI* is positively related to atmospheric instabilities, associated with the imbalance between the ejecting and sweeping contributions to the heat flux. For instance, based on EC measurements from a dryland (Hanford Site, Washington, USA), Gao et al. (2020) show that as atmospheric instability increases, the flux contribution from sweeps ($w' < 0, \theta' < 0$) decreases, while the flux contribution from ejections ($w' > 0, \theta' > 0$) remains constant (Fig. 8 in their work). In a separate study, based on observations from a 225 m meteorological tower in an urban environment (Tianjin, China), Zhang et al. (2023b) reported that with the increase of atmospheric instability, the flux contribution from ejections increases slightly, and the flux contribution of sweeps remains constant (Fig. 3 in their study). Despite slight inconsistencies in each component, the trend remains the same: the difference in flux contribution between the sweeps and ejections increases as the instability increases. This asymmetric transport phenomenon provides insight into the underestimation of the turbulent heat fluxes (Li et al. 2018; Liu et al. 2021). Further analysis of the asymmetric transport over the LES data can shed light on the mechanics of the *FI* within the ASL (Liu et al. 2024).

There is increasing scientific interest in understanding the mechanics of *FI* and utilizing this knowledge to improve EC measurements. However, few studies have focused on the validity of *FI* prediction models. For example, Zhou et al. (2023) compared three *FI* prediction models over a 1-D stripe heterogeneity surface with prescribed sensible and latent heat flux. Further evaluation of available *FI* prediction models over more complex terrain (i.e., 2-D heterogeneity) at different heterogeneity scales is needed to interrogate the robustness of these models. In order to achieve this objective, this study conducted a series of LES coupled with a land surface model (LES-LSM) over a checkerboard pattern of soil moisture heterogeneity. The focus of this study is to evaluate four selected *FI* prediction models and to investigate the role measurement height played in the *FI* through the quadrant analysis. The paper is organised as follows: Section 2 presents the *FI* prediction models, along with the model and simulation setup. Section 3 presents the results and a brief discussion. Section 4 presents the summary and conclusion.

2 Methods

2.1 Flux Imbalance

2.1.1 Flux Imbalance Estimation

To illustrate *FI*, we take the vertical wind speed (w) and potential temperature (θ), respectively, as examples, presenting them in both temporal and spatial domains using the Reynolds decomposition (Finnigan and Shaw 2008) as:

$$w(x,t) = \overline{w(x)} + w'(x,t), \tag{1a}$$

$$\theta(x,t) = \overline{\theta(x)} + \theta'(x,t), \tag{1b}$$

$$w(x,t) = \langle w\rangle(t) + \delta w(x,t), \tag{2a}$$

$$\theta(x,t) = \langle \theta\rangle(t) + \delta\theta(x,t). \tag{2b}$$

The temporal mean is indicated by an overline, and the spatial horizontal mean is indicated by the angle bracket. A prime symbol denotes the temporal fluctuation, and the symbol δ represents the spatial fluctuation. The total vertical heat flux at a given height (De Roo and Mauder 2018) is described as:

$$\langle\overline{w\theta}\rangle = \langle\overline{w}\rangle\langle\overline{\theta}\rangle + \langle\delta\overline{w}\delta\overline{\theta}\rangle + \langle\overline{w'\theta'}\rangle. \tag{3}$$

On the right-hand side of Eq. 3, the terms correspond to the mean vertical advection flux, the dispersive flux, and the eddy-covariance flux (i.e., turbulent heat flux), respectively. An important observation is that the vertical wind $\overline{w}$ does not exhibit a zero value at the point scale when a heterogeneous land surface experiences continuous heating under free convection conditions (Prabha et al. 2007).

Table 1 Different definitions of *FI* in the literature

Equations	References
$\langle I\rangle = \frac{\langle\overline{w\theta}\rangle - \left\langle\overline{w'\theta'}\right\rangle}{\langle\overline{w\theta}\rangle}$	Kanda et al. (2004), Inagaki et al. (2006), Steinfeld et al. (2007), Huang et al. (2008), Schalkwijk et al. (2016), Zhou et al. (2018)
$\langle I\rangle = \frac{\langle\overline{H}_{surf}\rangle - \left\langle\overline{w'\theta'}\right\rangle}{\langle\overline{H}_{surf}\rangle}$	De Roo et al. (2018), Zhou et al. (2019), Wanner et al. (2022a)
$\langle I\rangle = \frac{\left\langle\overline{w'\theta'}\right\rangle + \langle\delta\overline{w}\delta\overline{\theta}\rangle + \langle\overline{w}\rangle\langle\overline{\theta} - \overline{\theta}_{ref}\rangle}{\left\langle\overline{w'\theta'}\right\rangle}$	Morrison et al. (2022), (2023)

The dispersive flux has been consistently recognized as an important factor in achieving the closure of the surface energy balance (Kanda et al. 2004; Mauder et al. 2020). Research has shown that dispersive fluxes account for about 5–10% of the total sensible heat flux in the lower atmosphere, and this contribution increases with height, potentially reaching over 40% at heights above 100 m over an arid landscape (Margairaz et al. 2020a).

The *FI* is often defined as the nonlocal scale flux (i.e., the difference between the total reference heat flux and the turbulent heat flux) to the total reference heat flux (Kanda et al. 2004; Inagaki et al. 2006; Huang et al. 2008; De Roo et al. 2018; Zhou et al. 2019, 2023; Wanner et al. 2022b, 2024). In the literature, different selections of total reference heat flux have been proposed, including either the total heat flux within the atmosphere ($\langle\overline{w\theta}\rangle$, first row in Table 1), or surface heat flux ($\langle\overline{H}_{surf}\rangle$, second row in Table 1, often employed in studies with prescribed surface heat flux). In order to be consistent with the majority of previous research, we have followed the methodology outlined in the first row of Table 1 for the data processing. It can be reasonably assumed that the two definitions will have a negligible effect on the results, given that the sensible heat flux within the ASL is nearly constant (Baldocchi et al. 2001).

Furthermore, a novel definition of *FI* has been put forth by Morrison et al. (2022), which is the ratio of the total reference heat flux in the atmosphere to the turbulent heat flux. In the event of a perfect balance, the value is 1. Conversely, any underestimation yields a value greater than 1. The definition is also listed as the third row in Table 1 for reference.

2.1.2 Flux Imbalance Prediction Models

As mentioned in the introduction, numerous research efforts have been dedicated to determining the *FI* with statistical models considering turbulence characteristics. To gain a better understanding of the value and usefulness of these models, we selected four *FI* prediction models for a comprehensive evaluation. The models are H08, De18, Z19 and W22, with the semi-empirical formulas shown in Table 2. Notably, this study excluded the approach suggested by Panin and Bernhofer (2008) for using the effective roughness length, because a uniform roughness length of 0.1 m was applied across all simulations in our study.

In Table 2, $u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2\right)^{1/4}$ is the friction velocity, $w_* = \left[\frac{g z_i}{\theta_v}\left(\overline{w'\theta'}\right)_s\right]^{1/3}$ is the convective velocity scale, z is the measurement height, z_i is the boundary layer height defined

Table 2 *FI* prediction models in the literature

Notation	*FI* Prediction Model	Application Range	References
H08	$\langle I\rangle=\left[\exp\left(4.2-16\frac{u_*}{w_*}\right)+2.1\right]\left[1.1-8.0\left(\frac{z}{Z_i}-0.38\right)^2\right]^{0.5}$	$0.3\ z_i \sim 0.5\ z_i$	Huang et al. (2008)
De18	$\langle I\rangle=\left[0197\exp\left(-17.0\frac{u_*}{w_*}\right)+0.156\right][0.21+10.69\frac{z}{Z_i}]$	$< 0.1\ z_i$	De Roo et al. (2018)
Z19	$\langle I\rangle=1-[-1.46\frac{z}{Z_i}+1.0][-0.05\frac{Z_i}{L}\frac{l_w}{UT}+0.95]$	$0.03\ z_i \sim 0.1\ z_i$	Zhou et al. (2019)
W22	$\langle I\rangle=\left[a\exp\left(b\frac{u_*}{w_*}\right)+c\right]\left[20.05\frac{z}{Z_i}+0.157\right][H_{par}]$	$< 0.1\ z_i$	Wanner et al. (2022a)

as the height where the sensible heat flux profile has its minimum, $L=\frac{-u_*^3\theta_v}{kg\overline{w'\theta'}}$ is the Obukhov length, U is the mean horizontal wind speed, $l_w=\int_0^\infty R(r)dr$ is the integral length scale of vertical velocity, and R(r) is the autocorrelation function (Zhou et al. 2019), $H_{par}=\frac{gl_h}{U^2}\frac{\Delta T}{\overline{T}}$ is the thermal heterogeneity parameter (Margairaz et al. 2020b). Details of the scaling parameters are described in the Appendix 1.

It is important to note that the parameters a, b, and c in W22 are scale-dependent and were explicitly designed for heterogeneity scales of 200 m, 400 m, and 800 m in the original study. Therefore, they are labelled as W22-200 m, W22-400 m, and W22-800 m. To align closely with the heterogeneity scale in our simulation setting, we compared the simulations with heterogeneity scales of 2400 m (s2400), 1200 m (s1200) and 600 m (s600) to W22-800 m. However, this comparison may lead to uncertainties in the model's performance. The same approach applies to the remaining scenarios in which we compare the simulations with heterogeneity scale of 300 m (s300) to W22-400 m and heterogeneity scale of 150 m (s150) and 50 m (s50) to W22-200 m.

2.2 Data and Processing

2.2.1 Model Description

In this study, a dataset was generated using an LES model, ICON-LEM, coupled to a land surface model, TERRA-ML, at various scales of land surface heterogeneity. The ICON (Icosahedral Nonhydrostatic) framework is a collaboration between the German Weather Service (DWD), the Max Planck Institute for Meteorology (MPI-M), German Climate Computing Centre (DKRZ), Karlsruhe Institute of Technology (KIT) and the Centre for Climate Systems Modelling (C2SM). ICON was originally developed for research and operational purposes in numerical weather prediction and climate modelling (Zängl et al. 2015). The ICON framework enables the implementation of a LES mode, known as ICON-LEM (Dipankar et al. 2015). ICON-LEM uses the 3D Smagorinsky turbulence scheme to account for the SGS contribution. The performance of ICON-LEM has been extensively evaluated over a large spatial domain in Germany, as reported in the study of Heinze et al. (2017). Furthermore, ICON-LEM has been used in various scientific investigations, including studies on mesoscale secondary circulations (Han et al. 2019), high-CAPE (convective available potential energy)

(Rybka et al. 2021), Atmospheric Boundary Layer (ABL) tuning parametrization (Poll et al. 2022), cloud properties and precipitation (Verma and Burkhardt 2022), and many others. For a comprehensive explanation of the parameterisations of ICON-LEM, interested readers are referred to Dipankar et al. (2015).

The ICON framework comprises the land surface model TERRA-ML that enables two-way coupling between the land surface and the atmosphere (Grasselt et al. 2008; Schulz and Vogel 2020). TERRA-ML estimates the sensible and latent heat flux at the land surface using the following equations, and provides soil temperature and soil moisture as lower boundary layer conditions:

$$H = \rho C_p C_h |U| \big(\theta \pi_{sfc} - T_{sfc}\big), \tag{4}$$

$$LE = \rho L_v \frac{q_v - q_{sav}}{r_a + r_s}, \tag{5}$$

where H (W m^{-2}) is the sensible heat flux at the land surface, ρ is the air density (kg m^{-3}), C_p and C_h are the heat capacity (J kg^{-1} K^{-1}) and the bulk transfer coefficient (–), respectively, θ is the air potential temperature (K), and T_{sfc} is the ground surface temperature (K), with π as the scaled pressure at the ground surface (–). r_a denotes the aerodynamic resistance, and $r_s = r_{s,min}(\frac{\theta_1 - \theta_{adp}}{\theta_{fc} - \theta_{adp}})^{-1}$ is the soil resistance, with $r_{s,min}$=50 m s^{-1}, while θ_{adp} (K) and θ_{fc} (K) are the air dryness point and field capacity (Schulz and Vogel 2020).

2.2.2 Simulation Setup

In this study, the spatial resolution of the simulation was set to 50 m in both the x and y directions, covering an area of 4.8 km × 4.8 km with a double periodic boundary setting. This spatial configuration is considered suitable for capturing the mesoscale secondary circulation while ensuring computational efficiency (Han et al. 2019; Zhang et al. 2023a). The vertical grid consists of 420 levels with a resolution of 10 m, extending to a height of 4.2 km. A sponge layer with Rayleigh damping of the vertical velocity component was incorporated into the model at an altitude of 3700 m and above, with a depth of 500 m (the upper 8.3%), in order to nudge the atmosphere at the upper atmospheric boundary layer. The initial height of the boundary layer was set to 1000 m, with a constant potential temperature of 290 K below. Above 1000 m, the potential temperature was set to increase at a lapse rate of 0.006 K m^{-1}. Each simulation was conducted for a duration of 8 h, with a temporal step of 0.5 s for both ICON-LEM and TERRA-ML that coupled in every time step. A fixed incoming radiation of 700 W m^{-2} at the top of the atmosphere and a flat terrain ensure the tractability of the problem, as incorporating the diurnal variations in radiation (Zhou et al. 2018) or accounting for complex terrain (Rihani et al. 2015) would introduce additional complexity into to the system. It should be noted that the Coriolis effect and the background wind are not included in the simulation.

The use of a two-dimensional checkerboard pattern has been widely used in various studies, including the application of prescribed heat fluxes (Lee et al. 2019), temperature (Margairaz et al. 2020b; Wanner et al. 2022a) and soil moisture (Huang and Margulis 2009). In this work, a classical checkerboard pattern of soil moisture was used (Fig. 1), with the same horizontal grid size as the atmosphere (i.e., 50 m). Patch sizes (i.e. heterogeneity scales) varied from 50 m to 2,400 m. Specifically, a simulation with a patch size of 2,400 m is named s2400; similarly, s1200, s600, s300, s150 and s50 are named for simulations with patch sizes of 1,200 m, 600 m, 300 m, 150 m and 50 m, respectively. The soil moisture content distribution

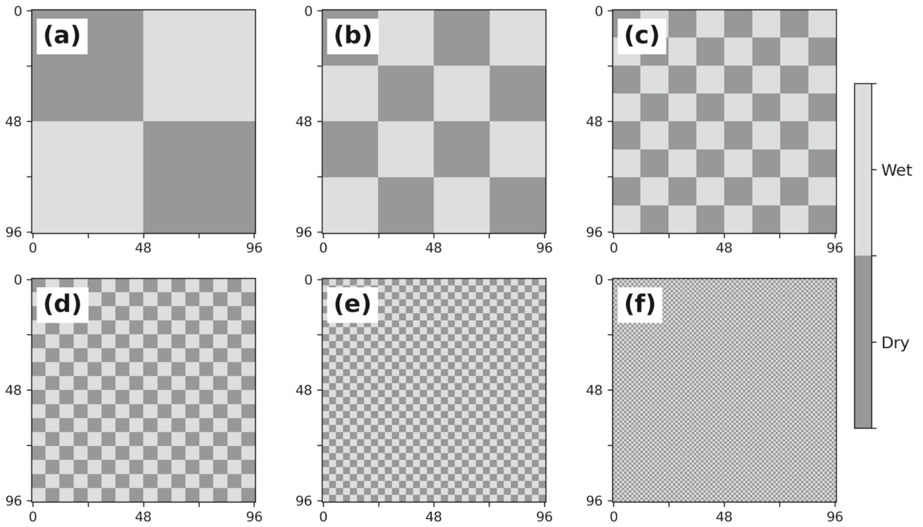

Fig. 1 Initial distribution of soil moisture. **a** heterogeneity scale = 2400 m, **b** heterogeneity scale = 1200 m, **c** heterogeneity scale = 600 m, **d** heterogeneity scale = 300 m, **e** heterogeneity scale = 150 m, **f** heterogeneity scale = 50 m. The simulations are named s2400, s1200, s600, s300, s150 and s50, respectively. The dark colour has a value of 0.25 $cm^3\ cm^{-3}$, and the light colour has a value of 0.55 $cm^3\ cm^{-3}$. The spatial average is 0.40 $cm^3\ cm^{-3}$. Additionally, an equivalent homogeneous simulation (Ho) was performed, see Appendix 2

included 0.55 $cm^3\ cm^{-3}$ for the wet region and 0.25 $cm^3\ cm^{-3}$ for the dry region, with a domain-average value of 0.40 $cm^3\ cm^{-3}$. In addition, an equivalent homogeneous simulation was conducted with a relatively lower soil moisture value but comparable heat fluxes. This is due to the fact that a simple soil moisture loss function is employed in the TERRA-ML model, which assumes that the evapotranspiration (i.e., latent heat flux) remains constant under low soil moisture conditions, increases at a constant rate within the transitional zone and becomes constant again when the soil becomes sufficiently wet. This is illustrated in Figs. 8 and 9 in the Appendix.

Although TERRA-ML comprises eight soil layers reaching a maximum depth of 15 m, the variation in the soil moisture profile was not considered due to the simulation's limited time of 8 h. Note, ICON uses unstructured grids instead of the traditional longitude-latitude grid (i.e., structured orthogonal grids); a checkerboard distribution of soil moisture can still be achieved through a remapping process (Wan et al. 2013). The model outputs the instantaneous variables at a 15-min interval. In addition, a time step of 0.5 s is saved for the wind components and potential temperatures of the lowest ten layers during the last hour.

2.2.3 Quadrant Analysis

The quadrant analysis method decomposes two turbulent fluctuations (e.g. w' and θ') into four quadrants, thereby determining the contribution and time fraction of turbulent heat flux from each quadrant. This approach has also been employed to investigate the dissimilarity between heat and momentum transport (Schmutz and Vogt 2019; Gao et al. 2020).

This study follows Li and Bou-Zeid's (2011) definition of quadrant decomposition, which includes the ejections ($w' > 0, \theta' > 0$), the sweeps ($w' < 0, \theta' < 0$), the inward interaction ($w' > 0, \theta' < 0$), and the outward interaction ($w' > 0, \theta' < 0$). The average fluxes within each quadrant can be expressed in Eq. 6, as:

$$\overline{w'\theta'}_k = \frac{1}{N}\sum_{t=1}^{N} w'\theta'\ I_k(t), \tag{6}$$

where $I_k(t)$ is an indicator variable with a weight of one if the sample is in the corresponding quadrant; otherwise, it has a weight of zero. The flux contribution $S(i)$ and time fraction $T(i)$ can then be calculated using the following equations (Schmutz and Vogt 2019),

$$S(i) = \frac{\overline{w'\theta'}_k}{\overline{w'\theta'}}, \tag{7}$$

$$T(i) = \frac{\sum_{n=1}^{N} I_k(t)}{N}. \tag{8}$$

3 Results and Discussions

3.1 Horizontal Cross-Section and Vertical Profiles

Figure 2 displays a clear cellular structure for the vertical wind velocity, characterised by narrow and intense updrafts and weak downdrafts. The air temperature over dry soil is higher than over wet soil with uniform incoming radiation. The air parcels over dry soil move upward due to buoyancy, which reduces and increases air pressure at the land surface and top of the ABL. Conversely, the air parcels over wet soil sinks, increasing air pressure at the land surface, leading to secondary circulations due to the horizontal pressure gradients.

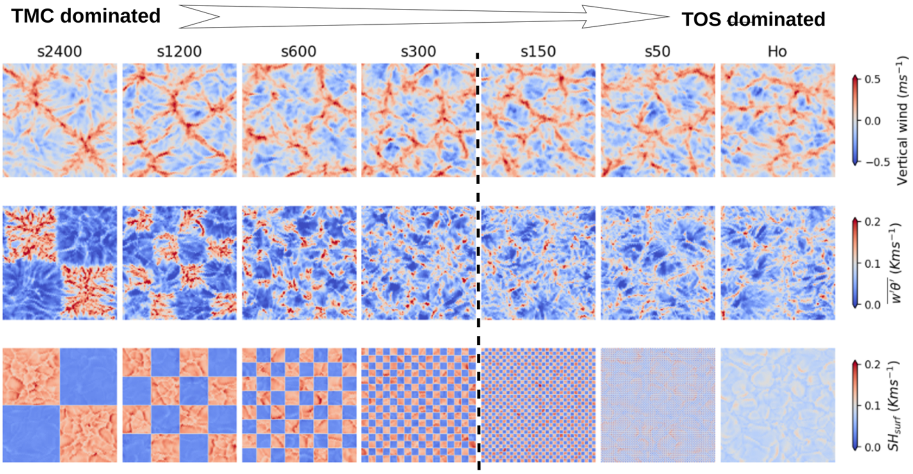

Fig. 2 Cross section of temporal-averaged vertical wind at 50 m (half level), turbulent heat flux, $\overline{w'\theta'}$ at 45 m (full level) and surface sensible heat flux, $\mathrm{SH_{surf}}$, at the land surface for the last hour of the simulation. At a large heterogeneity scale (on the left side), the atmospheric structure is dominated by the thermally-induced mesoscale circulation (TMC). At a small heterogeneity scale (on the right side), the atmospheric structure is dominated by the turbulent organised structure (TOS). The black dashed line represents the homogeneous equivalent line

This phenomenon is commonly known as thermally-induced mesoscale circulation (TMC) and has been extensively documented in the literature (Wanner et al. 2022a, b). A systematic discussion can be found in Salesky et al. (2017). Simulations s2400 and s1200 show a prevalence of updrafts in the dry region. However, as the scale of heterogeneity decreases to 600 m and 300 m (i.e., s600 and s300), the connection between the updraft pattern and the dry region pattern becomes less pronounced, as also shown in Fig. 11 in Appendix 3, that a point-to-point Pearson correlation coefficient between the turbulent heat flux at 45 m and the land surface heat flux decreased from 0.76 for s2400 to 0.27 for s300. Simulations s150 and s50 show a quasi-homogeneous random structure, similar to the homogeneous simulation, which has been previously reported as a slow-moving turbulent organised structure (TOS) in Kanda et al. (2004). According to the classification by Bou-Zeid et al. (2020), land patches of s150 and s50 could be considered statistically homogeneous if they exhibit locally variable surface characteristics at a small scale, yet remain relatively uniform at the regional scale. The turbulent energy spectrum in Fig. 10 of Appendix 3 also indicates that the difference between s150, s50 and the equivalent homogeneous case is negligible.

In Fig. 3, the potential temperature above the ASL remains constant until the boundary layer height is reached, as indicated by an almost vertical line. The potential temperature for simulations with larger heterogeneity scales (e.g. s2400, s1200) are slightly higher than those with smaller heterogeneity scales (e.g. s150, s50), which is consistent with the findings of previous studies (Raasch and Harbusch 2001; Huang and Margulis 2009; Zhou et al. 2019;

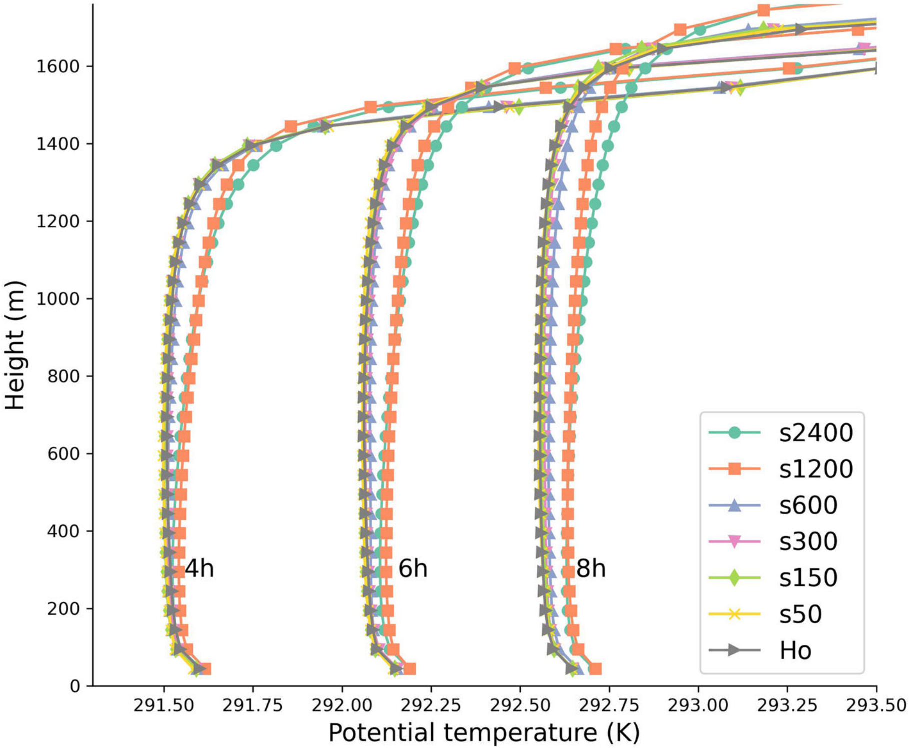

Fig. 3 Vertical profile of the horizontal mean potential temperature at 4 h, 6 h and 8 h

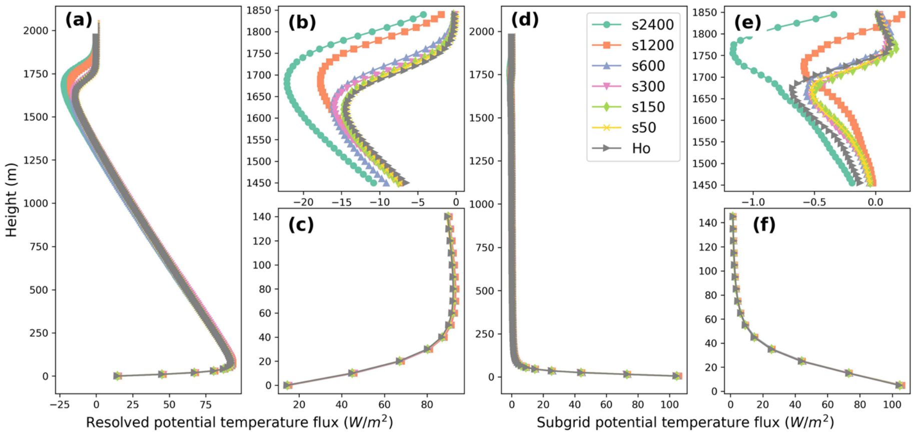

Fig. 4 Profile of the resolved and subgrid scale potential temperature flux (i.e. turbulent heat flux) for the last hour of the simulation. Panels **a** and **d** show the resolved and subgrid scale potential temperature flux within the convective boundary layer. Panels **b** and **c** are magnifications of **a** showing details in the entrainment layer and ASL. The same applies to panels **e** and **f** to **d**

Han et al. 2019). Enhanced entrainment effects in simulations with larger heterogeneity scales may explain this behaviour, as shown in Fig. 4.

In Fig. 4, the vertical profile of the potential temperature flux is depicted, showing a typical shape in accordance with Tennekes (1973), where the resolved potential temperature flux decreases as height increases, reaching its maximum negative value at the entrainment layer. For simulation s2400, the maximum negative value observed was -23 Wm^{-2}, which decreased to -17 Wm^{-2} for simulation s1200, and further decreased to -13 Wm^{-2} for the remaining cases up to a heterogeneity scale of 50 m. The significant influence of the subgrid-scale (SGS) contribution on the potential temperature flux near the surface layer is evident in Fig. 4f, which gradually decreases with increasing height. The lowest levels of the LES output are often excluded from the analysis because the flux is usually purely parameterized (De Roo and Mauder 2018; Zhou et al. 2019). However, surface-level data has been included in this study to provide readers with a complete perspective on the role of SGS in atmospheric heat flux characteristics.

3.2 Impact of Heterogeneity scale on the Atmospheric Structure

Table 3 shows the bulk statistics of the different simulations. Simulations with larger heterogeneity scales (s2400 and s1200) show a slightly higher boundary layer height (z_i, defined as the height where the sensible heat flux has its minimum), reaching 1695 m compared to the simulations with smaller heterogeneity scales (e.g., 1635 m for s300). A comparable trend is documented in the study by Zhou et al. (2019), which reports a decline in the mean boundary layer height from 1417 m to 1287 m for a decrease in heterogeneity scale from 2000 m to 240 m. However, z_i of s150 is slightly higher than s300 in our study, leading to a non-monotonic trend of the heterogeneity scale on the boundary layer height at small heterogeneity scales. One possible explanation is that the TMC dominates the atmosphere at sufficiently large heterogeneity scales. In contrast, at smaller heterogeneity scales, the formation of the TMC appears to be less pronounced, and the TOS dominates the mesoscale

Table 3 Bulk statistics of various simulations

Case	Heterogeneity Scale (m)	z_i(m)	u_* (m s^{-1})	w_* (m s^{-1})	t_* (s)	Bowen ratio (-)	$-z_i/L$(-)	H_{par}(-)
s2400	2400	1695	0.297	1.60	1056	0.747	74.6	135.3
s1200	1200	1695	0.323	1.60	1057	0.739	57.1	55.2
s600	600	1665	0.326	1.58	1050	0.730	54.8	30.3
s300	300	1635	0.348	1.57	1040	0.726	44.5	13.4
s150	150	1675	0.342	1.58	1061	0.722	47.6	6.7
s50	50	1635	0.345	1.56	1045	0.720	45.4	2.1
Ho	–	1635	0.344	1.57	1041	0.722	46.8	–

circulation pattern. Han et al. (2019) reported a non-monotonic trend in boundary layer height with a much larger heterogeneity scale ranging from 600 m to 76.8 km, which is attributed to surface heating but originates from secondary circulations. It is also noted that other factors, including the lapse rate (Peng et al. 2023) and geostrophic wind (Margairaz et al. 2020b), may affect the boundary layer height.

The values of u_* and w_* are calculated at a height of 45 m. The results indicate that as the heterogeneity scale decreases, u_* increases while w_* decreases, resulting in a slight decrease of the atmospheric stability ($-z_i/L$). The Bowen ratio decreased with a decreasing heterogeneity scale, shifting from 0.747 to 0.720. Additionally, the difference in large-eddy turnover time scale ($t_* = z_i/w_*$) among the simulations with different heterogeneity scales is negligible, with a value range between 1040 and 1061 s, roughly 17 min.

3.3 Flux Imbalance Prediction Model Evaluation

This section analyses the vertical sensible heat flux components, which comprise the turbulent heat flux, the mean vertical advection flux, and the dispersive flux at various measurement heights. Based on these components, we assess four *FI* prediction models selected from the literature, as listed in Table 2.

Figure 5 shows the different components of the vertical heat flux and the *FI* within the ASL. The turbulent heat flux (including the SGS contribution) decreases with height, while the dispersive flux increases with height. The results agree with Margairaz et al.'s (2020a) findings that the contribution of the dispersive component to the total heat flux increases from 10 to 40% as the height increases from near surface to above 100 m. The mean vertical advection flux calculation in this study includes a reference potential temperature, as suggested by Morrison et al. (2022). The mean vertical advection fluxes exhibit fluctuations in the lower levels, up to 45 m, before increasing with height. However, the mean vertical advection fluxes are several magnitudes smaller than the turbulent heat fluxes and the dispersive heat fluxes. This suggests that the mean vertical advection fluxes may be ignored in modelling efforts. For example, Wanner et al. (2024) also show that properly modelling the dispersive flux is the key to closing the energy balance.

The *FI* shows non-monotonicity across the heterogeneity scale that increases from s150 to s1200, reaching a maximum, then decreasing again at s2400. A similar result has been reported by Zhou et al. (2019), who found that when the surface heterogeneity scale becomes

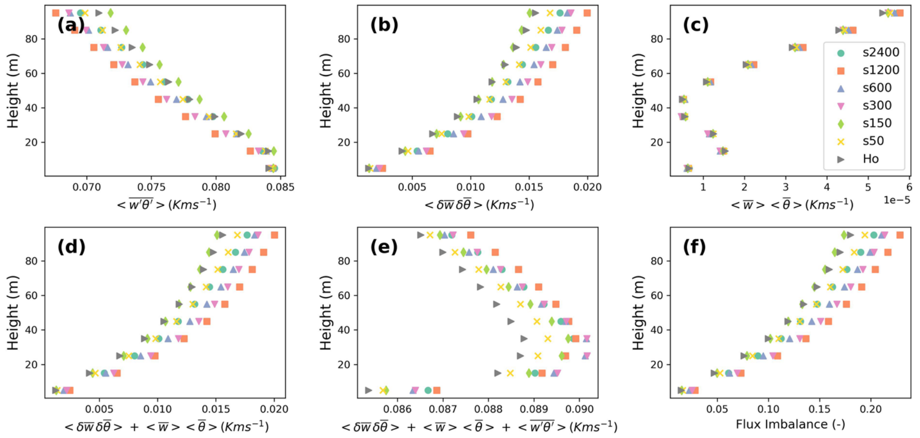

Fig. 5 Flux components and *FI* for the last hour of simulation. The first row shows the turbulent heat flux including the SGS contribution, the dispersive flux and the mean vertical advection flux, respectively. The second row shows the sum of the vertical advective and dispersive heat flux, the total reference heat flux and *FI*. Noted that the *FI* is dimensionless. A reference temperature is included to calculate the mean vertical advection flux based on Morrison et al. (2022), which is set as the mean value of the previous hour

larger than the boundary layer height, the *FI* starts to decrease. This may be due to the flows behaving like a homogeneous surface at scales below the heterogeneity scale, which minimizes the impact of the dispersive fluxes. Additionally, *FI* is more pronounced in the s50 case compared to the s150 case. This difference is likely because, in the s50 case, each patch corresponds to a single grid cell. Further investigation using varying grid sizes while maintaining the same patch size may help clarify the underlying causes, offering a promising direction for future research.

Figure 6 shows the performance of the four *FI* prediction models considered. The performance of H08 is relatively worse compared to the other models. This is because H08 is fitted with data obtained between 0.3 z_i to 0.5 z_i, which extends beyond the atmospheric surface layer (ASL) and limits its direct applicability to the ASL *FI* estimates. Similar conclusions have been drawn in Eder et al. (2014). Furthermore, if z/z_i is less than 0.01, (i.e., $z = 17$ m with $z_i = 1700$ m), the value within the square root in H08 formula becomes negative, resulting in a singularity. Therefore, the data below 15 m is excluded in Fig. 6a. Nevertheless, in our opinion, the concept of H08 is useful for quantifying *FI* based on atmospheric stability and relative measurement height in the mixing layer.

De18 produces results comparable to those obtained from the LES, even though it was originally developed from simulations over a homogeneous land surface. The results show a slight overestimation of *FI* near the surface and an underestimation with increasing height. Z19, compared to De18, shows a similar performance in predicting the *FI*. In the first layer near the land surface (i.e. z = 5 m), Z19 provides a negative imbalance, which is also reported by Zhou et al. (2019) (Fig. C2 in their paper). This can be explained by the fact that Z19 was constructed with a data range of 0.03 z_i to 0.1 z_i, which is larger than 5 m.

W22 performs best when the heterogeneity scale is well matched, such as s300 and s150. However, W22 tends to significantly overestimate the FI for simulations with a larger heterogeneity scale, resulting in unreasonable values (not shown in the figure). Consequently, directly applying W22 to the climate model presents a challenge, particularly for grid sizes at several kilometre scales.

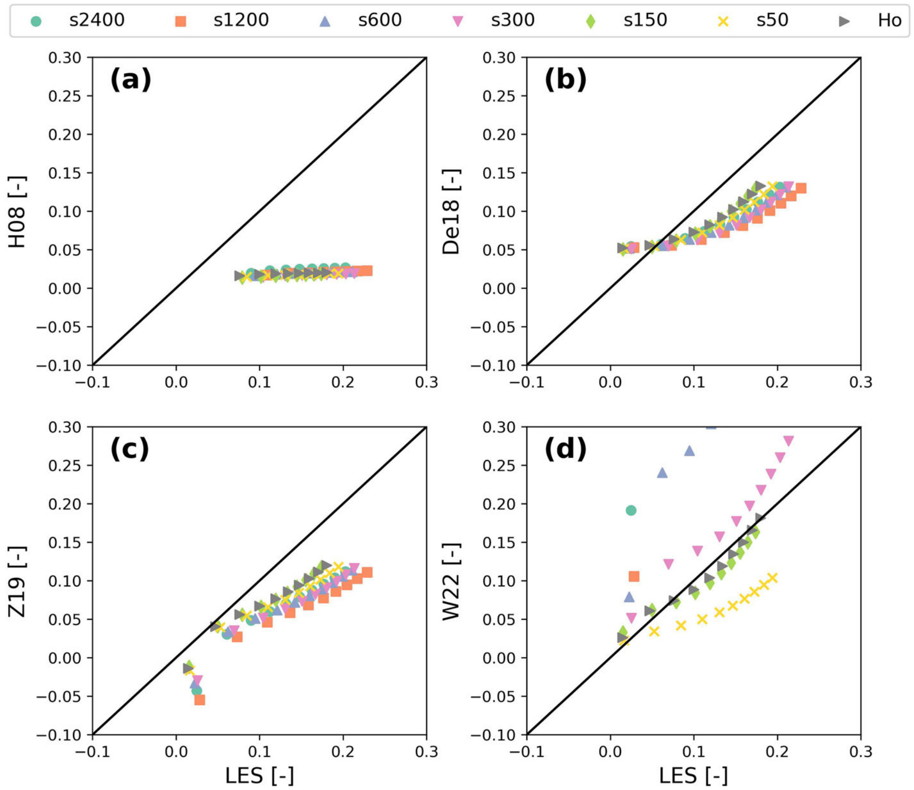

Fig. 6 Performance of the *FI* prediction models for the bottom ten layers. **a** H08, **b** De18, **c** Z19 **d** W22. The x-axis represents the *FI* calculated from LES data, and the y-axis displays the *FI* estimated by the *FI* prediction models. Note that *FI* increases consistently with height

3.4 Quadrant Analysis

The turbulent flux is mainly attributed to the contributions from the ejections ($w' > 0, \theta' > 0$) and sweeps ($w' < 0, \theta' < 0$) (Högström and Bergström 1996). Several studies indicate that the difference between the contributions from ejections and sweeps increases with atmospheric instability, leading to an increase in the *FI* (Gao et al. 2020; Liu et al. 2021). To investigate the potential correlation between *FI* and height, a quadrant analysis was conducted at different measurement heights.

As the height in the ASL increases from 5 to 95 m, the time fraction of ejection (Fig. 7a) decreases from a mean of 0.37 to 0.30. Meanwhile, the flux contribution of ejection (Fig. 7d) increases from a mean of 0.58 to 0.70, along with a higher standard deviation. This suggests that the ejections originate from the land surface and increase in intensity while becoming narrower with increasing height. In contrast, the time fraction of sweeps (Fig. 7b) fluctuates with increasing height, starting at 0.44 at 5 m, peaking at 0.50 at 35 m, and then decreasing to 0.43 at 95 m. The flux contribution of sweeps (Fig. 7e) decreases with increasing height, indicating that the sweeps originate from the atmosphere and head towards the land surface. These findings are consistent with the work of Li and Bou-Zeid (2011) and Gao et al. (2020).

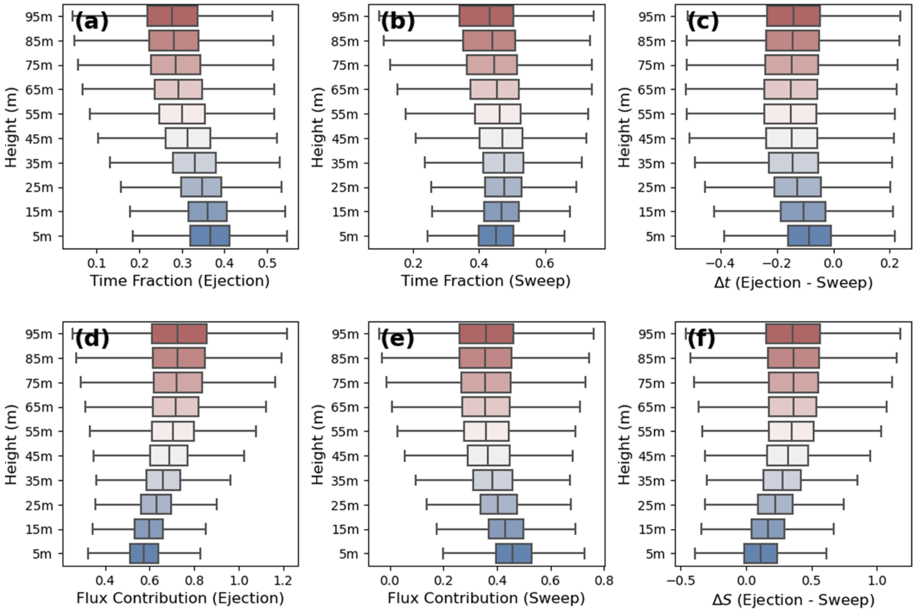

Fig. 7 Time fraction and flux contribution of ejections and sweeps for case s1200. Similar patterns are observed in other cases (not shown)

Another observation from Fig. 7f is that the contribution difference between ejections and sweeps (ΔS) increases with height, correlating with the decrease in turbulent heat flux in Fig. 5. Liu et al. (2021) also reported comparable results based on the measurement from the EC tower, that the increased ΔS links to the reduced turbulent heat flux, which explains the increased *FI* with height.

3.5 Discussion

The De18, Z19, and W22 show promising results for predicting *FI*. However, the barriers to directly applying *FI* prediction models in correcting EC measurements must be discussed. When revisiting the *FI* prediction models, it becomes clear that the boundary layer height (z_i) is a critical parameter. However, z_i is not always available from measurements (Mauder et al. 2021). The thermodynamic and aerodynamic properties of the atmosphere can be obtained through various methods, such as Ceilometer, Radiosonde, Doppler Sodar, and profiling radar, as discussed in Seibert et al. (2000). Moreover, these measurements are relatively expensive to install and maintain. Currently, only 28 stations are reported to have co-located atmospheric boundary layer measurements in the FLUXNET sites (Helbig et al. 2021). In addition, data from these stations may not be continued, and their distribution is generally concentrated in North America and Europe.

The *FI* prediction models under consideration primarily target convective (i.e., unstable) boundary layer conditions. However, it is important to correct EC data during neutral and stable conditions to achieve energy balance closure over extended time periods. The relative *FI* remains significant during nocturnal periods, even if absolute values may be small (Franssen et al. 2010). Here, the atmospheric stability parameters used in H08, De18, and W22 are inadequate because of the absence of a well-defined w_* under stable conditions (Zhou et al. 2019).

4 Summary and Conclusion

The accurate estimation of turbulent heat flux is a persistent challenge in Earth science. Numerous researches suggest that large-scale eddies significantly contribute to the flux imbalance (*FI*). The objective of this study is to investigate *FI* arising from mesoscale transport in convective boundary layer conditions, with a particular emphasis on assessing the performance of selected *FI* prediction models from the literature.

The dataset was generated using a large eddy simulation (ICON-LEM) coupled with a land surface model (TERRA-ML). A checkerboard pattern of wet-dry soil moisture was used to represent the two-dimensional heterogeneity of the land surface across a range of scales from 50 m to 2.4 km. The model was set up with a spatial coverage of 4.8 × 4.8 km^2 and a grid resolution of 50 m. It assumed flat terrain and employed a double periodic boundary setting. Simulations were conducted over an 8-h period with a time step of 0.5 s. The model was initialised with a constant incoming radiation of 700 W m^{-2} at the top of the atmosphere.

Two distinct sub-mesoscale circulations were identified: thermally induced mesoscale circulation (TMC) and turbulent organised structure (TOS). TMC primarily affects atmospheric dynamics on larger heterogeneity scales, reflecting a clear updraft pattern that mirrors the configuration of the arid land surface. As the heterogeneity scale decreases, the relationship between updraft position and soil moisture pattern weakens, and TOS becomes the dominant structure in atmospheric dynamics. Additionally, the domain-averaged *FI* shows a non-linear correlation with the heterogeneity scales, which could be attributed to the presence of two distinct sub-mesoscale circulations. This finding is consistent with previous research suggesting an optimal heterogeneity scale for the formation of secondary circulations.

Four selected *FI* prediction models were evaluated, including H08 (Huang et al. 2008), De18 (De Roo et al. 2018), Z19 (Zhou et al. 2019) and W22 (Wanner et al. 2022a). H08 showed a tendency to underestimate the *FI*, which can be attributed to its initial valid vertical height range of 0.3 z_i to 0.5 z_i, while the data used in our analysis are within the above surface layer (ASL), approximately within 0.1 z_i. De18 and Z19 show a similar performance that can roughly capture the *FI*, with a slight underestimation as the measurement height increases. W22 agrees well with our LES data when the heterogeneity scale matches its simulation setting. The performance of the *FI* prediction model provides a potential in correcting EC measurements.

A further observation derived from quadrant analysis is the notable discrepancy between the flux contributions from ejections and sweeps along the height. This difference correlates with a decrease in the turbulent heat flux and an increase in the *FI*. These findings emphasise the importance of considering the EC measurement height and the atmospheric model's vertical grid size when studying the vertical variations of turbulent fluxes.

Appendix 1: Scaling Parameters

The scaling parameters used in this work are listed in Table 4.

Table 4 Scaling Parameters

Scaling parameters	Symbols	Equations
Obukhov length	L	$L = \frac{-u_*^3\theta_v}{kg\overline{w'\theta'}}$
von Kármán constant	k	0.40
Friction velocity	u_*	$u_* = \left\lvert\frac{\tau}{\rho}\right\rvert^{\frac{1}{2}} = (\overline{u'w'}^2 + \overline{v'w'}^2)^{1/4}$
Convective velocity scale	w_*	$w_* = \left[\frac{gz_i}{\theta_v}(\overline{w'\theta'})_s\right]^{1/3}$
Characteristic temperature scale	θ_*	$\theta_* = \frac{\overline{w'\theta'}_s}{u_*}$
Large-eddy turnover time scale	t_*	$t_* = z_i/w_*$
Thermal heterogeneity parameter	H_{par}	$H_{par} = \frac{gl_h}{U^2}\frac{\Delta T}{T}$, l_h is the heterogeneity scale
Integral length scale of vertical velocity	l_w	$l_w = \int_0^\infty R(r)dr$

Appendix 2: Equivalent Homogeneous Simulation

An equivalent homogeneous simulation was performed in this study but with a different domain-averaged soil moisture value. A sensitivity analysis of soil moisture to the domain-averaged sensible heat flux, as shown in Fig. 8, indicates that a value of 0.325 cm^3 cm^{-3} is comparable to the heterogeneous case. This is because the land surface model used in this

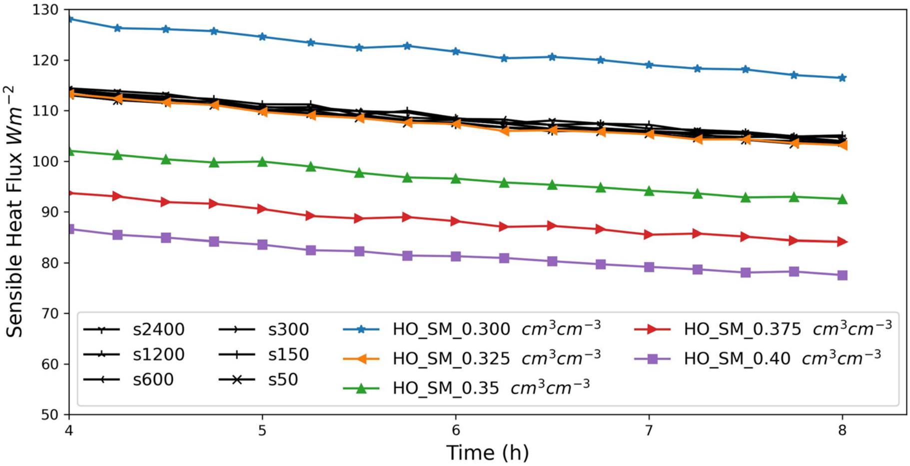

Fig. 8 Time-series of sensible heat flux at the land surface for different simulations. Simulation with homogeneous soil moisture (HO_SM) of 0.325 cm^3 cm^{-3} is suitable to represent a comparable homogeneous case in this study

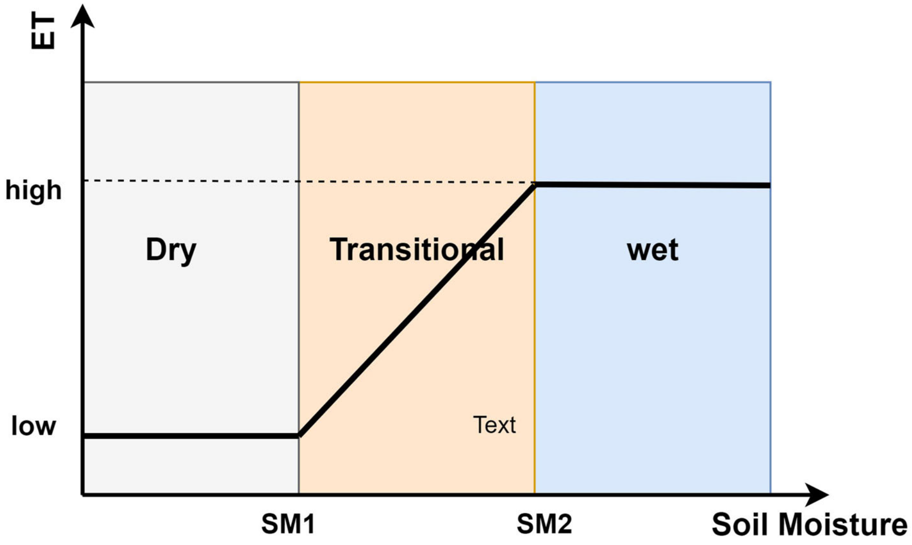

Fig. 9 Sketch of the evapotranspiration calculated in TERRA-ML

study (TERRA-ML) utilized a simple soil moisture loss function to calculate evapotranspiration (i.e., latent heat flux), as shown in Fig. 9. In this model, evapotranspiration remains constant when soil moisture is low, increases at a constant rate within the transitional zone, and becomes constant again when the soil is sufficiently wet.

Appendix 3: Statistical Homogeneous Simulation

Figure 10 indicates that the differences in the kinetic energy spectrum between different simulations are largest at the lowest wavenumber. Specifically, simulations s2400 and s1200 show significant differences compared to other cases. In the case of s600, these differences begin to decrease. For the s300 simulation, differences are still observable, particularly at wavenumbers 2 and 3. However, for simulations s150 and s50, the differences from the homogeneous case are negligible, indicating that a statistically homogeneous state is reached.

Furthermore, a point-to-point Pearson correlation coefficient is calculated between the land surface heat flux and the turbulent heat flux in the atmosphere at 45 m for the average of the last hour simulation. This is illustrated in Fig. 11. The Pearson correlation coefficient demonstrates a decline with a reduction in heterogeneity scale, exhibiting values of 0.74 for s2400, 0.62 for s1200, and 0.47 for s600. The coefficient for simulation s300 is 0.27, which is higher than those for simulations s150 and s50, where values are less than 0.1 and are similar to those observed in the homogeneous case.

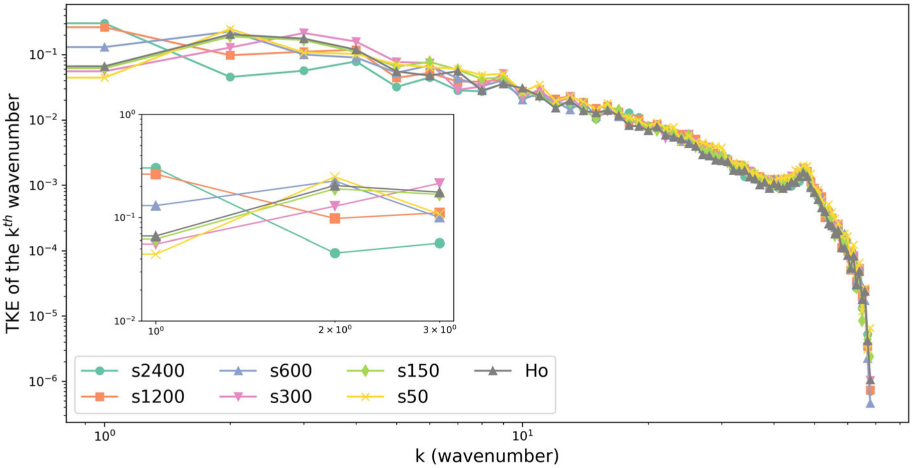

Fig. 10 Turbulent kinetic energy spectrum for the last hour simulation at a height of 45 m. The subfigure provides a zoomed-in view at the smallest wavenumber

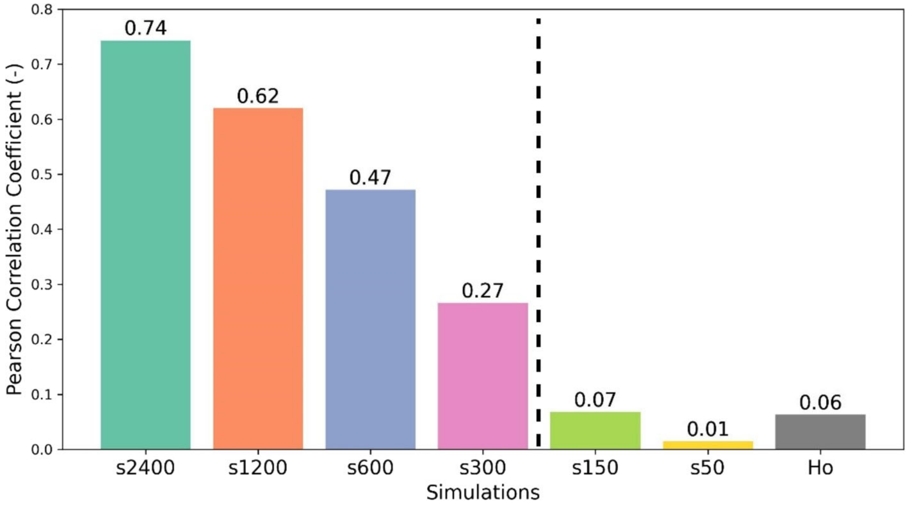

Fig. 11 Point-to-point Pearson correlation coefficient between the land surface heat flux and the turbulent heat flux in the atmosphere at 45 m, averaged over the last hour of simulation

Acknowledgements This study is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy-EXC 2070-390732324. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). The authors also thank the reviewers for their valuable feedback and constructive comments, which greatly contributed to improving the quality of this paper.

Author Contributions LZ and SP performed the model simulation. LZ analyzed the data, wrote the main manuscript and prepared Figs. 1–11. SK conceived the original idea and provided the funding. All authors revised the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability The numerical simulation data used in this paper could be made available upon direct request to the authors.

Declarations

Conflict of interest The authors declare no competing interests.

References

Albertson JD, Kustas P, Scanlon M (2001) Large-eddy simulation over heterogeneous terrain with remotely sensed land surface conditions and LES (in the absence of free parameters) yields regionally averaged land surface height in the atmospheric surface atmosphere is shown to be scale-dependent. Water Resour Res 37:1939–1953

Aubinet M, Grelle A, Ibrom A, et al. (1999) Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. In: Advances in ecological research. pp 113–175

Baldocchi DD (2003) Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob Chang Biol 9:479–492. https://doi.org/10.1046/j.1365-2486.2003.00629.x

Baldocchi DD (2020) How eddy covariance flux measurements have contributed to our understanding of global change biology. Glob Chang Biol 26:242–260. https://doi.org/10.1111/gcb.14807

Baldocchi D, Falge E, Gu L et al (2001) FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull Am Meteorol Soc 82:2415–2434. https://doi.org/10.1175/1520-0477(2001)082%3c2415:FANTTS%3e2.3.CO;2

Barr AG, Morgenstern K, Black TA et al (2006) Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurement of the CO_2 flux. Agric For Meteorol 140:322–337. https://doi.org/10.1016/j.agrformet.2006.08.007

Bou-Zeid E, Anderson W, Katul GG, Mahrt L (2020) The persistent challenge of surface heterogeneity in boundary-layer meteorology: a review. Bound-Layer Meteorol 177:227–245. https://doi.org/10.1007/s10546-020-00551-8

Chu H, Christianson DS, Cheah YW et al (2023) AmeriFlux BASE data pipeline to support network growth and data sharing. Sci Data 10:1–13. https://doi.org/10.1038/s41597-023-02531-2

De Roo F, Mauder M (2018) The influence of idealized surface heterogeneity on virtual turbulent flux measurements. Atmos Chem Phys 18:5059–5074. https://doi.org/10.5194/acp-18-5059-2018

De Roo F, Zhang S, Huq S, Mauder M (2018) A semi-empirical model of the energy balance closure in the surface layer. PLoS ONE 13:1–23. https://doi.org/10.1371/journal.pone.0209022

Dipankar A, Stevens B, Heinze R et al (2015) Large eddy simulation using the general circulation model ICON. J Adv Model Earth Syst 6:963–986. https://doi.org/10.1002/2015MS000431

Eder F, De Roo F, Kohnert K et al (2014) Evaluation of two energy balance closure parametrizations. Bound-Layer Meteorol 151:195–219. https://doi.org/10.1007/s10546-013-9904-0

Eder F, De Roo F, Rotenberg E et al (2015) Secondary circulations at a solitary forest surrounded by semi-arid shrubland and their impact on eddy-covariance measurements. Agric For Meteorol 211–212:115–127. https://doi.org/10.1016/j.agrformet.2015.06.001

Finnigan JJ, Shaw RH (2008) Double-averaging methodology and its application to turbulent flow in and above vegetation canopies. Acta Geophys 56:534–561. https://doi.org/10.2478/s11600-008-0034-x

Finnigan JJ, Clement R, Malhi Y et al (2003) A Re-evaluation of long-term flux measurement techniques part I: averaging and coordinate rotation. Bound-Layer Meteorol 107:1–48

Foken T (2008) The energy balance closure problem: an overview. Ecol Appl 18:1351–1367. https://doi.org/10.1890/06-0922.1

Foken T, Mauder M, Liebethal C et al (2010) Energy balance closure for the LITFASS-2003 experiment. Theor Appl Climatol 101:149–160. https://doi.org/10.1007/s00704-009-0216-8

Franssen HJH, Stöckli R, Lehner I et al (2010) Energy balance closure of eddy-covariance data: a multisite analysis for European FLUXNET stations. Agric For Meteorol 150:1553–1567. https://doi.org/10.1016/j.agrformet.2010.08.005

Gao Z, Liu H, Russell ES et al (2016) Large eddies modulating flux convergence and divergence in a disturbed unstable atmospheric surface layer. J Geophys Res d Atmos 121:4211–4232. https://doi.org/10.1002/2015JD024529.Received

Gao Z, Russell ES, Missik JEC et al (2017) A novel approach to evaluate soil heat flux calculation: an analytical review of nine methods. J Geophys Res 122:6934–6949. https://doi.org/10.1002/2017JD027160

Gao Z, Liu H, Chen X et al (2020) Enlarged nonclosure of surface energy balance with increasing atmospheric instabilities linked to changes in coherent structures. J Geophys Res d Atmos 125:1–13. https://doi.org/10.1029/2020JD032889

Gopalakrishnan SG, Avissar R (2000) An LES study of the impacts of land surface heterogeneity on dispersion in the convective boundary layer. J Atmos Sci 57:352–371. https://doi.org/10.1175/1520-0469(2000)057%3c0352:ALSOTI%3e2.0.CO;2

Grasselt R, Schüttemeyer D, Warrach-Sagi K et al (2008) Validation of TERRA-ML with discharge measurements. Meteorol Z 17:763–773. https://doi.org/10.1127/0941-2948/2008/0334

Han C, Brdar S, Raasch S, Kollet S (2019) Large-eddy simulation of catchment-scale circulation. Quart J Roy Meteor Soc 145:1218–1233. https://doi.org/10.1002/qj.3491

Heinze R, Dipankar A, Henken CC et al (2017) Large-eddy simulations over Germany using ICON: a comprehensive evaluation. Quart J Roy Meteor Soc 143:69–100. https://doi.org/10.1002/qj.2947

Helbig M, Gerken T, Beamesderfer ER et al (2021) Integrating continuous atmospheric boundary layer and tower-based flux measurements to advance understanding of land-atmosphere interactions. Agric For Meteorol. https://doi.org/10.1016/j.agrformet.2021.108509

Hellsten A, Ketelsen K, Sühring M et al (2021) A nested multi-scale system implemented in the large-eddy simulation model PALM model system 6.0. Geosci Model Dev 14:3185–3214. https://doi.org/10.5194/gmd-14-3185-2021

Högström U, Bergström H (1996) Organized turbulence structures in the near-neutral atmospheric surface layer. J Atmos Sci 53:2452–2464. https://doi.org/10.1175/1520-0469(1996)053%3c2452:OTSITN%3e2.0.CO;2

Huang HY, Margulis SA (2009) On the impact of surface heterogeneity on a realistic convective boundary layer. Water Resour Res 45:1–16. https://doi.org/10.1029/2008WR007175

Huang J, Lee X, Patton EG (2008) A modelling study of flux imbalance and the influence of entrainment in the convective boundary layer. Bound-Layer Meteorol 127:273–292. https://doi.org/10.1007/s10546-007-9254-x

Inagaki A, Letzel MO, Raasch S, Kanda M (2006) The impact of the surface heterogeneity on the energy imbalance problem using les. J Meteorol Soc Japan 84:187–198. https://doi.org/10.2208/prohe.49.343

Kanda M, Inagaki A, Letzel MO et al (2004) LES study of the energy imbalance problem with eddy covariance fluxes. Bound-Layer Meteorol 110:381–404. https://doi.org/10.1023/B:BOUN.0000007225.45548.7a

Lee JM, Zhang Y, Klein SA (2019) The effect of land surface heterogeneity and background wind on shallow cumulus clouds and the transition to deeper convection. J Atmos Sci 76:401–419. https://doi.org/10.1175/JAS-D-18-0196.1

Li D, Bou-Zeid E (2011) Coherent structures and the dissimilarity of turbulent transport of momentum and scalars in the unstable atmospheric surface layer. Bound-Layer Meteorol 140:243–262. https://doi.org/10.1007/s10546-011-9613-5

Li D, Katul GG, Liu H (2018) Intrinsic constraints on asymmetric turbulent transport of scalars within the constant flux layer of the lower atmosphere. Geophys Res Lett 45:2022–2030. https://doi.org/10.1002/2018GL077021

Liu H, Gao Z, Katul GG (2021) Non-closure of surface energy balance linked to asymmetric turbulent transport of scalars by large eddies. J Geophys Res d Atmos 126:1–12. https://doi.org/10.1029/2020JD034474

Liu H, Liu C, Huang J et al (2024) Scalar flux profiles in the unstable atmospheric surface layer under the influence of large eddies: Implications for eddy covariance flux measurements and the non-closure problem. Geophys Res Lett. https://doi.org/10.1029/2023gl106649

Margairaz F, Pardyjak ER, Calaf M (2020a) Surface thermal heterogeneities and the atmospheric boundary layer: the relevance of dispersive fluxes. Bound-Layer Meteorol 175:369–395. https://doi.org/10.1007/s10546-020-00509-w

Margairaz F, Pardyjak ER, Calaf M (2020b) Surface thermal heterogeneities and the atmospheric boundary layer: the thermal heterogeneity parameter. Bound-Layer Meteorol 177:49–68. https://doi.org/10.1007/s10546-020-00544-7

Maronga B, Raasch S (2013) Large-eddy simulations of surface heterogeneity effects on the convective boundary layer during the LITFASS-2003 experiment. Bound-Layer Meteorol 146:17–44. https://doi.org/10.1007/s10546-012-9748-z

Mauder M, Cuntz M, Drüe C et al (2013) A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agric for Meteorol 169:122–135. https://doi.org/10.1016/j.agrformet.2012.09.006

Mauder M, Foken T, Cuxart J (2020) Surface-energy-balance closure over land: a review. Bound-Layer Meteorol 177:395–426. https://doi.org/10.1007/s10546-020-00529-6

Mauder M, Ibrom A, Wanner L et al (2021) Options to correct local turbulent flux measurements for large-scale fluxes using an approach based on large-eddy simulation. Atmospher Measur Techniq 14:7835–7850. https://doi.org/10.5194/amt-14-7835-2021

Morrison T, Pardyjak ER, Mauder M, Calaf M (2022) The heat-flux imbalance: the role of advection and dispersive fluxes on heat transport over thermally heterogeneous terrain. Bound-Layer Meteorol 183:227–247. https://doi.org/10.1007/s10546-021-00687-1

Morrison TJ, Calaf M, Pardyjak ER (2023) A full three-dimensional surface energy balance over a desert playa. Quart J Roy Meteor Soc 149:102–114. https://doi.org/10.1002/qj.4397

Novick KA, Biederman JA, Desai AR et al (2018) The AmeriFlux network: a coalition of the willing. Agric for Meteorol 249:444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

Panin GN, Bernhofer C (2008) Parametrization of turbulent fluxes over inhomogeneous landscapes. Izv Atmos Ocean Phys 44:701–716. https://doi.org/10.1134/S0001433808060030

Pastorello G, Trotta C, Canfora E et al (2020) The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data 7:225. https://doi.org/10.1038/s41597-020-0534-3

Patton EG, Sullivan PP, Shaw RH et al (2016) Atmospheric stability influences on coupled boundary layer and canopy turbulence. J Atmos Sci 73:1621–1647. https://doi.org/10.1175/JAS-D-15-0068.1

Peng S, Yang Q, Shupe MD et al (2023) The characteristics of atmospheric boundary layer height over the Arctic Ocean during MOSAiC. Atmos Chem Phys 23:8683–8703. https://doi.org/10.5194/acp-23-8683-2023

Poll S, Shrestha P, Simmer C (2022) Grid resolution dependency of land surface heterogeneity effects on boundary-layer structure. Quart J Roy Meteor Soc 148:141–158. https://doi.org/10.1002/qj.4196

Prabha TV, Karipot A, Binford MW (2007) Characteristics of secondary circulations over an inhomogeneous surface simulated with large-eddy simulation. Bound-Layer Meteorol 123:239–261. https://doi.org/10.1007/s10546-006-9137-6

Raasch S, Harbusch G (2001) An analysis of secondary circulations and their effects caused by small-scale surface inhomogeneities using large-eddy simulation. Bound-Layer Meteorol 101:31–59. https://doi.org/10.1023/A:1019297504109

Rebmann C, Aubinet M, Schmid H et al (2018) ICOS eddy covariance flux-station site setup: a review. Int Agrophys 32:471–494. https://doi.org/10.1515/intag-2017-0044

Rihani JF, Chow FK, Maxwell R (2015) Isolating effects of terrain and soil moisture heterogeneity on the atmospheric boundary layer: idealized simulations to diagnose land-atmosphere feedbacks. J Adv Model Earth Syst 7:915–937. https://doi.org/10.1002/2014MS000371

Rybka H, Burkhardt U, Köhler M et al (2021) The behavior of high-CAPE (convective available potential energy) summer convection in large-domain large-eddy simulations with ICON. Atmos Chem Phys 21:4285–4318. https://doi.org/10.5194/acp-21-4285-2021

Salesky ST, Chamecki M, Bou-Zeid E (2017) On the nature of the transition between roll and cellular organization in the convective boundary layer. Bound-Layer Meteorol 163:41–68. https://doi.org/10.1007/s10546-016-0220-3

Schalkwijk J, Jonker HJJ, Siebesma AP (2016) An investigation of the eddy-covariance flux imbalance in a year-long large-eddy simulation of the weather at Cabauw. Bound-Layer Meteorol 160:17–39. https://doi.org/10.1007/s10546-016-0138-9

Schmutz M, Vogt R (2019) Flux similarity and turbulent transport of momentum, heat and carbon dioxide in the urban boundary layer. Bound-Layer Meteorol 172:45–65. https://doi.org/10.1007/s10546-019-00431-w

Schulz JP, Vogel G (2020) Improving the processes in the land surface scheme TERRA: bare soil evaporation and skin temperature. Atmosphere 11:513. https://doi.org/10.3390/atmos11050513

Seibert P, Beyrich F, Gryning SE et al (2000) Review and intercomparison of operational methods for the determination of the mixing height. Atmos Environ 34:1001–1027. https://doi.org/10.1016/S1352-2310(99)00349-0

Shen S, Leclerc MY (1994) Large-eddy simulation of small-scale surface effects on the convective boundary-layer structure. Atmos Ocean 32:717–731. https://doi.org/10.1080/07055900.1994.9649519

Steinfeld G, Letzel MO, Raasch S et al (2007) Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study. Bound-Layer Meteorol 123:77–98. https://doi.org/10.1007/s10546-006-9133-x

Stoy PC, Mauder M, Foken T et al (2013) A data-driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity. Agric For Meteorol 171–172:137–152. https://doi.org/10.1016/j.agrformet.2012.11.004

Stull RB (1988) An introduction to boundary layer meteorology. Springer, Dordrecht

Sun J, Massman WJ, Banta RM, Burns SP (2021) Revisiting the surface energy imbalance. J Geophys Res d Atmos 126:1–31. https://doi.org/10.1029/2020JD034219

Tennekes H (1973) A model for the dynamics of the inversion above a convective boundary layer. J Atmos Sci 30:558–567. https://doi.org/10.1175/1520-0469(1973)030%3c0558:AMFTDO%3e2.0.CO;2

Twine TE, Kustas WP, Norman JM et al (2000) Correcting eddy-covariance flux underestimates over a grassland. Agric For Meteorol 103:279–300. https://doi.org/10.1016/s0168-1923(00)00123-4

Verma P, Burkhardt U (2022) Contrail formation within cirrus: ICON-LEM simulations of the impact of cirrus cloud properties on contrail formation. Atmos Chem Phys 22:8819–8842

Wan H, Giorgetta MA, Zängl G et al (2013) The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids—Part 1: Formulation and performance of the baseline version. Geosci Model Dev 6:735–763. https://doi.org/10.5194/gmd-6-735-2013

Wanner L, Calaf M, Mauder M (2022a) Incorporating the effect of heterogeneous surface heating into a semi-empirical model of the surface energy balance closure. PLoS ONE 17:e0268097. https://doi.org/10.1371/journal.pone.0268097

Wanner L, De Roo F, Sühring M, Mauder M (2022b) How does the choice of the lower boundary conditions in large-eddy simulations affect the development of dispersive fluxes near the surface? Bound-Layer Meteorol 182:1–27. https://doi.org/10.1007/s10546-021-00649-7

Wanner L, Jung M, Paleri S et al (2024) Towards energy-balance closure with a model of dispersive heat fluxes. Bound-Layer Meteorol 190:25. https://doi.org/10.1007/s10546-024-00868-8

Wilson K, Goldstein A, Falge E et al (2002) Energy balance closure at FLUXNET sites. Agric For Meteorol 113:223–243

Yamamoto S, Saigusa N, Gamo M et al (2005) Findings through the AsiaFlux network and a view toward the future. J Geogr Sci 15:142–148. https://doi.org/10.1007/bf02872679

Zängl G, Reinert D, Rípodas P, Baldauf M (2015) The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: description of the non-hydrostatic dynamical core. Quart J Roy Meteor Soc 141:563–579. https://doi.org/10.1002/qj.2378

Zhang L, Poll S, Kollet S (2023a) Large-eddy simulation of soil moisture heterogeneity-induced secondary circulation with ambient winds. Quart J Roy Meteor Soc. https://doi.org/10.1002/qj.4413

Zhang L, Zhang H, Zhang X et al (2023b) Dissimilarity of turbulent transport of momentum and heat under unstable conditions linked to convective circulations. J Geophys Res d Atmos 128:1–18. https://doi.org/10.1029/2022JD037997

Zhou Y, Li X (2019) Energy balance closures in diverse ecosystems of an endorheic river basin. Agric For Meteorol 274:118–131. https://doi.org/10.1016/j.agrformet.2019.04.019

Zhou Y, Li D, Liu H, Li X (2018) Diurnal variations of the flux imbalance over homogeneous and heterogeneous landscapes. Bound-Layer Meteorol 168:417–442. https://doi.org/10.1007/s10546-018-0358-2

Zhou Y, Li D, Li X (2019) The effects of surface heterogeneity scale on the flux imbalance under free convection. J Geophys Res d Atmos 124:8424–8448. https://doi.org/10.1029/2018JD029550

Zhou Y, Sühring M, Li X (2023) Evaluation of energy balance closure adjustment and imbalance prediction methods in the convective boundary layer—a large eddy simulation study. Agric for Meteorol 333:109382. https://doi.org/10.1016/j.agrformet.2023.109382