JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2024MS004801

Key Points:

- This study proposes a surface-layer heat flux correction scheme for earth system models
- The new approach (FLIMK) describes the local heat flux with the K-theory and the nonlocal part with the flux imbalance prediction model
- The FLIMK scheme is tested in both large eddy simulation and numerical weather prediction models, and the results are satisfactory

Correspondence to:
L. Zhang,
lijie.zhang@wsu.edu

Citation:
Zhang, L., Poll, S., Weinkaemmerer, J., & Kollet, S. (2026). An alternative way to parameterizing the nonlocal scale sensible heat flux using the flux imbalance and K-theory approach. *Journal of Advances in Modeling Earth Systems*, *18*, e2024MS004801. https://doi.org/10.1029/2024MS004801

Received 5 NOV 2024
Accepted 12 FEB 2026

Author Contributions:
Conceptualization: Lijie Zhang, Stefan Poll, Stefan Kollet
Data curation: Lijie Zhang, Jan Weinkaemmerer
Formal analysis: Lijie Zhang, Stefan Poll, Jan Weinkaemmerer
Funding acquisition: Stefan Kollet
Investigation: Lijie Zhang, Stefan Poll, Jan Weinkaemmerer
Methodology: Lijie Zhang, Stefan Poll, Stefan Kollet
Project administration: Stefan Poll, Stefan Kollet
Software: Lijie Zhang, Stefan Poll, Jan Weinkaemmerer
Supervision: Stefan Poll, Stefan Kollet
Validation: Lijie Zhang, Stefan Poll, Jan Weinkaemmerer
Visualization: Lijie Zhang

An Alternative Way to Parameterizing the Nonlocal Scale Sensible Heat Flux Using the Flux Imbalance and K-Theory Approach

Lijie Zhang[1], Stefan Poll[1,2,3], Jan Weinkaemmerer[1], and Stefan Kollet[1,2]

[1]Institute of Bio and Geosciences, Agrosphere (IBG-3), Forschungszentrum Jülich, Jülich, Germany, [2]Center for High-Performance Scientific Computing in Terrestrial Systems, Geoverbund ABC/J, Jülich, Germany, [3]SimDataLab Terrestrial Systems, Jülich Supercomputing Centre (JSC), Jülich, Germany

Abstract Accurately representing the surface sensible heat flux (SHF) in Earth System Models (ESMs) is of paramount importance and challenge. This study introduces an alternative surface-layer SHF correction scheme by modeling nonlocal scale SHF through flux imbalance (FI) prediction models while employing the conventional gradient diffusion approach (K-theory) for the local scale part. This approach is termed the FLux IMbalance and K-theory (FLIMK). The FLIMK relies solely on atmospheric stability parameters (u_*/w_*) and the ratio of measurement height to the boundary layer height (z/zi). The FLIMK underwent testing using large eddy simulation (LES) under dry convective boundary layer conditions with a prescribed stripe-like heterogeneous SHF at the land surface. Additionally, the FLIMK scheme was incorporated into a numerical weather prediction (NWP) model and tested in the single-column mode with data from the Atmospheric Radiation Measurement test case. The results indicate that K-theory underestimates the SHF by approximately 15% due to the effects of the nonlocal scale. In contrast, the FLIMK scheme effectively reduces this imbalance, decreasing it from 15% (16%) to 6% (6.7%) for LES (NWP). The findings suggest that the FLIMK scheme can potentially enhance the parameterization of SHF in ESMs.

Plain Language Summary Understanding the heat flux between the land surface and the atmosphere is essential for studying their interactions. Traditional K-theory only captures the local part of this heat transfer, so an extra term is needed to represent the nonlocal part. This study suggests that using flux imbalance prediction models to account for the nonlocal part of the heat flux offers an alternative solution. The new method, FLux IMbalance and K-theory, has been tested with large eddy simulations and weather prediction models, showing significant improvements in representing the sensible heat flux.

1. Introduction

An accurate representation of the sensible heat flux (SHF) at the land–atmosphere (L–A) interface is essential not only for land surface modeling but also for capturing the full spectrum of atmospheric processes affected by surface heterogeneity. SHF governs surface energy partitioning (Lin et al., 2022; Mauder et al., 2018), shapes the development of the planetary boundary layer (PBL; Sullivan et al., 1998), and influences entrainment processes at the PBL top (VanZanten et al., 1999). These processes, in turn, affect the timing and structure of cloud formation (Stevens, 2007), the initiation of moist convection, and the distribution and intensity of precipitation (Hohenegger et al., 2009; Sedlar et al., 2022). Recent studies, for example, Lee et al. (2019) demonstrate that surface heterogeneity–induced circulations can impact cloud development and mesoscale organization, underscoring the need for surface flux parameterizations that capture nonlocal effects.

SHF is commonly estimated using the gradient diffusion approach, also known as K-theory, due to its simplicity. This method assumes that the SHF is always down-gradient, and proportional to the vertical potential temperature gradient, and is expressed as:

$$H = -\rho C_p K_h \frac{\partial \overline{\theta}}{\partial z}, \tag{1}$$

where H (W m^{-2}) is the SHF, ρ (kg m^{-3}) is the air density, C_p (J kg^{-1} K^{-1}) is the specific heat capacity of air at constant pressure, K_h (m^2 s^{-1}) is the eddy diffusivity, θ (K) is the potential temperature, and z (m) is the height.

Writing – original draft: Lijie Zhang
Writing – review & editing: Stefan Poll, Stefan Kollet

K-theory captures only local-scale turbulent transport, overlooking contributions from nonlocal processes such as secondary circulations (Kang et al., 2007; Meredith et al., 2014; Pleim, 2007). This conceptual limitation is shared by commonly used measurement techniques, including the eddy-covariance (EC) system, which similarly assume fluxes arise solely from local turbulence (Jung et al., 2019; Mauder et al., 2024; Wilson et al., 2002). Secondary circulations are increasingly recognized as a key driver of nonlocal SHF, as supported by both observational evidence (Eder et al., 2015; Paleri et al., 2022) and large-eddy simulations (Avissar & Schmidt, 1998; Raasch & Harbusch, 2001). These unaccounted nonlocal contributions are a likely cause of the systematic underestimation of local-scale SHF when compared to the total reference SHF (H_0) (Kanda, 2006; Y. Zhou & Li, 2019).

A consistent underestimation of SHF by approximately 10%–30% has been widely reported in eddy-covariance measurements, motivating extensive efforts to understand the underlying causes of this mismatch, which have been linked to land surface heterogeneity (Foken, 2008; Mauder et al., 2020). Studies have shown that the degree of underestimation strongly depends on factors such as friction velocity ($u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2\right)^{\frac{1}{4}}$, where u, v, and w are the longitudinal, lateral and vertical velocity components along the Cartesian coordinates x, y, and z, respectively) and atmospheric stability ($-z/L$, or $-\frac{zi}{L} = k\left(\frac{w_*^3}{u_*^3}\right)$, z is the measurement height, zi is the boundary layer height defined as the height of the maximum in the vertical potential temperature gradient, $L = \frac{u_*^3\theta_v}{kg\left(w'\theta_v'\right)_s}$ is the Obukhov length, $w_* = \left[\frac{g}{\theta_v}zi\left(\overline{w'\theta_v'}\right)_s\right]^{\frac{1}{3}}$ is the convective velocity scale, θ_v is the virtual potential temperature and k is the von Kármán constant). For example, Franssen et al. (2010) conducted a comprehensive analysis using data from 26 European FLUXNET sites and found that the surface energy balance non-closure decreased with increasing friction velocity and atmospheric stability. Similar conclusions were reached by Stoy et al. (2013) based on data from 173 FLUXNET sites worldwide. In the rest of this study, we use $-zi/L$ as the bulk atmospheric stability parameter. This choice follows earlier boundary-layer studies (Deardorff, 1972; Panofsky et al., 1977) and has been widely adopted in large-eddy simulation (LES) based investigations of non-local turbulence and convective boundary-layer dynamics (De Roo et al., 2018; Johansson et al., 2001; Khanna & Brasseur, 1997).

Around the same time, research using LES demonstrated that nonlocal effects—such as secondary circulations—can be quantified through similarity-based approaches. Specifically, the flux imbalance (FI), defined as the ratio of nonlocal SHF to the total reference SHF (i.e., H_0 = local + nonlocal), can be expressed using semi-empirical functions based on relevant scaling variables (De Roo & Mauder, 2018; J. Huang et al., 2008; Wanner et al., 2022; Y. Zhou et al., 2019). For example, J. Huang et al. (2008) conducted idealized LES experiments and proposed that FI can be described as a function of friction velocity and convective velocity scale, as well as the normalized height: $\mathrm{FI} = f_1(u_*/w_*)f_2(z/zi)$. Their model is valid within the mixed layer (approximately 0.3 zi −0.5 zi) and was later extended by De Roo et al. (2018) to the atmospheric surface layer (ASL) (<0.1 zi). The semi-empirical function introduced by De Roo et al. (2018) (hereafter referred to as De18) has shown robust performance under various conditions, including field measurements at the Soroe beech forest site in Denmark and two pre-Alpine grassland sites in Germany (Mauder et al., 2021), as well as in LES studies incorporating different land surface heterogeneity patterns (Zhang, Poll, & Kollet, 2024; Y. Zhou et al., 2023).

This study addresses the research question: can the representation of SHF at the ASL in ESMs be improved by incorporating FI prediction models to capture nonlocal effects? Building on existing local-scale approaches, we introduce a hybrid framework that combines conventional K-theory with semi-empirical FI prediction models to represent both local and nonlocal SHF contributions. It is noted that the focus of this work is placed entirely on improving the SHF at the surface layer, without modifying boundary layer parameterizations. Section 2 describes the FI and K-theory (FLIMK) scheme and introduces the numerical model and simulation setups. Section 3 presents the results of the FLIMK scheme in both LES and numerical weather prediction model (NWP). Finally, a summary and conclusion are offered in Section 4.

2. Methods

2.1. Flux Imbalance and K-Theory Approach

The SHF in the ASL (roughly 10% of PBL height close to the surface) consists of the local and nonlocal parts, as shown in Figure 1 (De Roo & Mauder, 2018). The local scale turbulent heat flux is typically described as the covariance of two scalar measurement fluctuations, especially those obtained using the eddy-covariance system

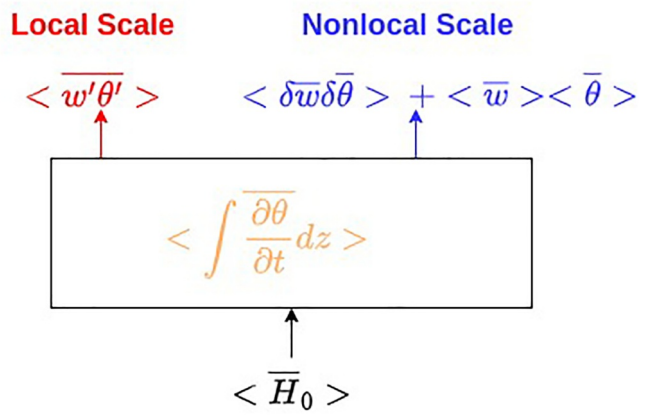

Figure 1. Graphical representation of the sensible heat flux in the lowest atmospheric layer, with an influx from the surface ($\langle \overline{H}_0 \rangle$) and an outflux into higher layers. The local scale component is shown in red, and the nonlocal component is shown in blue. The storage term of θ is shown in orange and is negligible compared to the other terms. Only the heat flux in the vertical direction is considered. Figure adapted from De Roo and Mauder (2018).

(e.g., $\overline{w'\theta'}$, the overbar denotes the temporal mean and prime "′" represents the temporal fluctuation). In ESMs, the local scale flux is presented using K-theory as $-K_h \frac{\partial \overline{\theta}}{\partial z}$, where K_h is the vertical heat diffusivity, and its detailed formulation is provided in Section 2.2. We note that both the Smagorinsky closure used in LES and the MOST-based flux parameterization used in the NWP configuration have known limitations under gray-zone resolutions and over heterogeneous surfaces, where local similarity assumptions may break down. The FLIMK scheme is specifically designed to address these limitations by providing a correction to surface layer SHF to account for unresolved nonlocal transport. The nonlocal processes include the mean vertical advection flux ($\langle \overline{w} \rangle \langle \overline{\theta} \rangle$), and the dispersive flux ($\langle \delta\overline{w}\delta\overline{\theta} \rangle$), here, the angle brackets indicate the spatial mean, and "δ" denotes the spatial fluctuation. In this paper, the definition of the FI follows Y. Zhou et al. (2019) as the ratio of the nonlocal part to the total reference SHF (i.e., the sum of local and nonlocal part) in the surface layer

$$\mathrm{FI} = \frac{\langle \delta\overline{w}\delta\overline{\theta} \rangle + \langle \overline{w} \rangle \langle \overline{\theta} \rangle}{\langle \overline{H}_0 \rangle} = 1 - \frac{\langle \overline{w'\theta'} \rangle}{\langle \overline{H}_0 \rangle}. \tag{2}$$

The FI has received attention both in field measurements (Barr et al., 2006; Foken et al., 2011; Franssen et al., 2010; Paleri et al., 2022; Stoy et al., 2013; Wilson et al., 2002) and in the LES (De Roo et al., 2018; J. Huang et al., 2008; Wanner et al., 2022; Y. Zhou et al., 2019, 2023), leading to several semi-empirical functions being proposed to quantify the FI. For example, by assuming that SHF in the atmosphere arises from a combination of bottom-up transport from the surface and top-down transport associated with entrainment at the boundary-layer top, a FI prediction model has been proposed using a shape function of the form: $\mathrm{FI} = f_1(\frac{u_*}{w_*}) f_2(\frac{z}{zi})$. The physical rationale underlying this formulation is based on the work of J. Huang et al. (2008), for convenience, is also provided in the Appendix B.

In this study, the FI prediction model from De Roo et al. (2018) is selected, hereafter FLIMK_De18, shown in Equation 3

$$\mathrm{FI} = \left[0.197 \exp\left(-17.0 \frac{u_*}{w_*}\right) + 0.156\right]\left[0.21 + 10.69 \frac{z}{zi}\right], \tag{3}$$

where u_* (m s^{-1}) is the friction velocity, w_* (m s^{-1}) is the convective velocity scale, z (m) is the measurement height, and zi (m) is the boundary layer height. Equation 3 is designed to predict the FI of the domain-averaged SHF. Equation 3 is primarily validated and applied within the surface layer, from several meters above the ground up to approximately 0.1 zi. The FI may also be directly diagnosed from LES output based on the definition in Equation 2, which is referred to as FLIMK_LES. It is noted that although Equation 3 was originally fitted using LES data over a homogeneous land surface, it has shown potential for application over more complex heterogeneous conditions. For example, a recent evaluation by Zhang, Poll, and Kollet (2024) demonstrated that Equation 3 can accurately capture near-surface FI even over two-dimensional checkerboard-patterned soil moisture fields with varying heterogeneity scales.

Based on the FI prediction model, a novel approach, the FLIMK scheme, is proposed to account for the nonlocal effect of SHF in the surface layer, as shown in Figure 2. Specifically, the SHF in the lowest atmospheric layer will be modified using the FI prediction model, as specified in Equation 4. The derivation of Equation 4 is presented in Appendix A.

$$\overline{H}_0 = \frac{-K_h \frac{\partial \overline{\theta}}{\partial z}}{1 - \mathrm{FI}}. \tag{4}$$

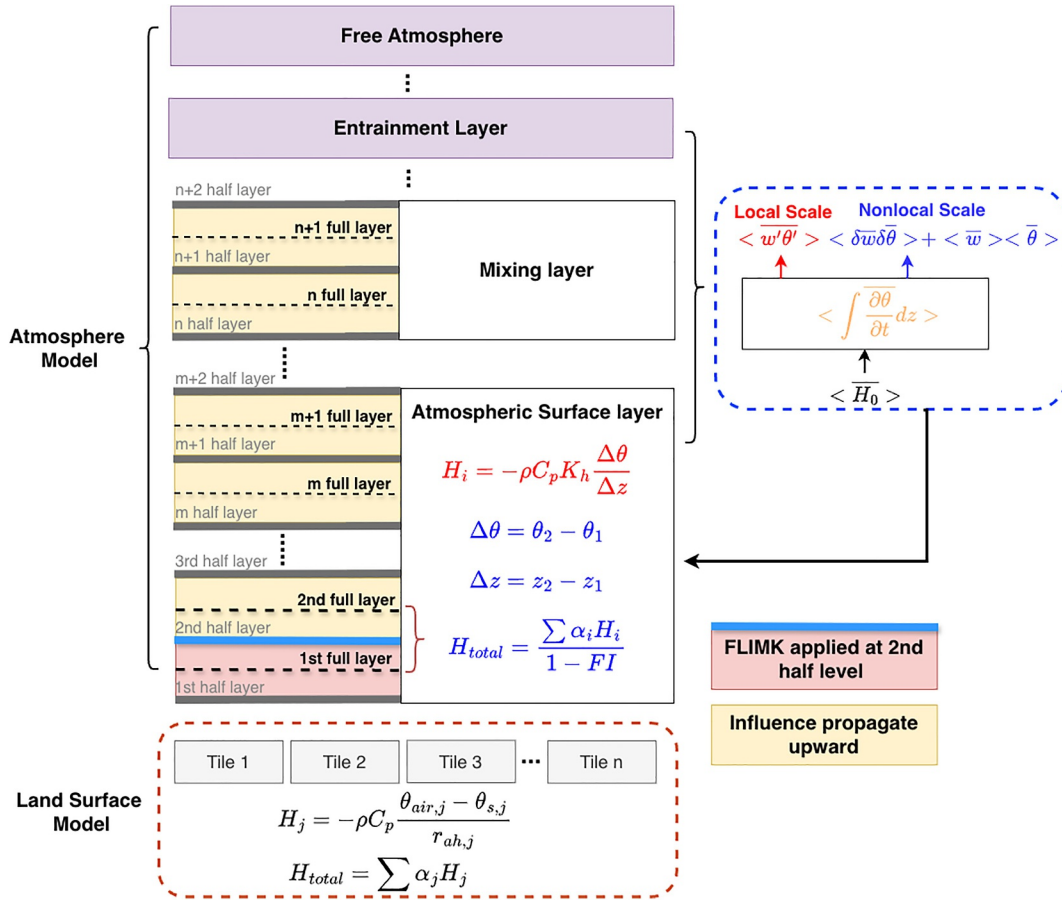

Figure 2. Schematic of the FLux IMbalance and K-theory. The *flux imbalance* prediction models represent a nonlocal effect that modifies the upper boundary of the lowest atmospheric layer (the second half layer). The potential temperature gradient is computed between two adjacent full layers. Note, layer indexing is inverted relative to the original ICON model for illustration.

2.2. Model and Simulation Setup

In this study, the FLIMK scheme is tested in both the idealized LES with prescribed SHF at the surface and in the NWP setup using real test cases from the Atmospheric Radiation Measurement (ARM) campaign (Brown et al., 2002; Ďurán et al., 2021). The Icosahedral Nonhydrostatic Weather and Climate Model (ICON) was selected, as ICON is the new generation of fully compressible atmospheric models, covering all scales including the climate model, NWP, and LES (Dipankar et al., 2015; Hohenegger et al., 2023; Jungclaus et al., 2022; Van Pham et al., 2020; Zängl et al., 2015).

In the LES, the land surface heterogeneity was configured with three stripes (Figure 3), with the middle stripe representing a SHF of 0.2 K m s^{-1}, and the side stripes representing 0.1 K m s^{-1}. The domain, with double-periodic boundary conditions, covers an area of 9.6 × 9.6 km^2 and has a spatial resolution of 50 × 50 m^2. It extends vertically to 4.2 km with a vertical grid size of 10 m, with a sponge layer set at the top 500 m of the model. The simulation was run for 8 hr, with analysis focused on the final hour. During this period, the lowest 10 model levels were saved at a 0.5-s time step. It is worth noting that the FLIMK scheme is not run online; instead, it is evaluated by calculating a *priori* of the LES output. The SHF at 45 m is calculated as the sum of the resolved vertical flux $\overline{w'\theta'}$ and the subgrid-scale (SGS) heat flux, which is based on the K-theory through the 3D sub-grid Smagorinsky scheme with Lilly's stability correction. The underestimation is computed as the difference between the prescribed surface SHF and this diagnosed value. The 45 m level is selected to balance two goals: staying within the surface layer and minimizing the influence of SGS contributions. Prescribing the SHF at the surface

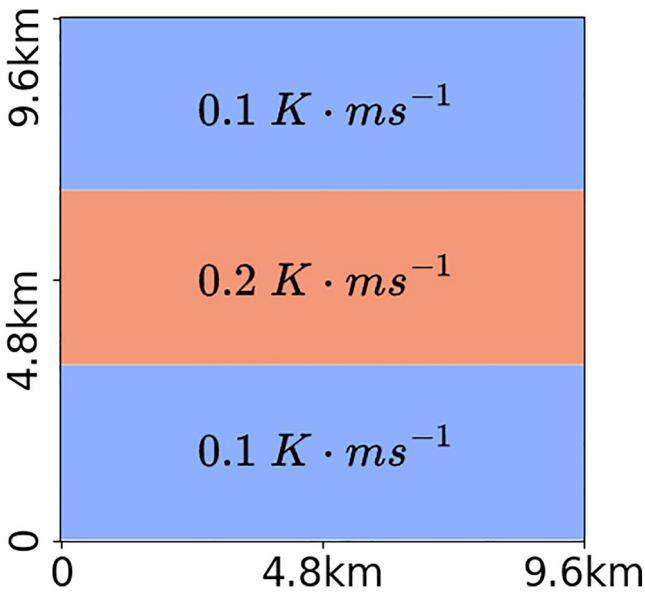

Figure 3. Prescribed sensible heat flux (SHF) in the large eddy simulation. The middle strip has a larger SHF of 0.2 K m s^{-1} and the side stripes have a smaller value of 0.1 K m s^{-1}. Each stripe has a width of 3.2 km and length of 9.6 km.

allows us to isolate the effects of the FLIMK parameterization on the representation of SHF in the ASL, without the added complexity of feedback from L-A coupling.

To assess how surface heterogeneity affects SHF representation across spatial scales, we applied FLIMK to a series of sub-regions within the LES domain, ranging from the native 50 m grid scale up to 9.6 km. Required factors in the Equation 3 were calculated as the spatial-mean value of the sub-region. In this study, w_* computed using potential temperature (θ) rather than virtual potential temperature (θ_v), under the assumption of dry conditions where their difference is negligible. For moist environments, θ_v should be used to better capture buoyancy-driven convective dynamics. The spatially averaged heat flux over the sub-region in the lower atmosphere functionally mimics what tile-based schemes aim to represent, and the scale-aware analysis helps bridge the gap between turbulence-resolving LES and coarser-resolution models. While FLIMK is ultimately intended for use as a subgrid-scale correction in ESMs, this intermediate evaluation provides a step toward understanding its behavior under varying surface conditions and spatial aggregation scales.

In the ICON-NWP, forcing data were obtained from measurements conducted over the ARM program site in Oklahoma on 21 June 1997 (Brown et al., 2002; Lenderink & Holtslag, 2004). The model is initialized with profiles of temperature, moisture, wind, and turbulent kinetic energy (TKE) over 12 vertical levels extending through the lower troposphere. Time-varying large-scale forcings—advection of heat and moisture, radiative cooling, and subsidence—are prescribed, along with geostrophic wind components. Surface boundary conditions include prescribed sensible and latent heat fluxes. This configuration provides a realistic yet controlled framework for evaluating turbulence parameterizations under diurnally forced boundary-layer conditions. A detailed description of the data set can be found in Ďurán et al. (2021) and the data are publicly available through the repository referenced therein on Zedono (Bastak Duran et al., 2021). Additionally, the performance of ICON in single-column mode (ICON-SCM) has been evaluated using a well-designed LES model (MicroHH, Van Heerwaarden et al., 2017), which yielded satisfactory results. The ICON-SCM configuration followed that of Ďurán et al. (2021), with an improved vertical grid size of 10 m and the incorporation of the FLIMK scheme in an online simulation approach in the upper boundary of the lowest atmosphere layer (i.e., second half layer in Figure 2). An additional simulation without incorporating the FLIMK scheme is carried out as a reference run.

In the ICON-NWP configuration, the SHF at the upper boundary of the lowest atmospheric layer is calculated using the local K-theory approach based on the potential temperature gradient between the two lowest model layers. The eddy diffusivity for heat is computed using a turbulence scheme based on the prognostic equation for TKE, following (Raschendorfer, 2001). The eddy diffusivity is given by:

$$K_h = c_k \cdot l \sqrt{\mathrm{TKE}}, \tag{5}$$

where K_h is the vertical eddy diffusivity for heat (ICON variable TKVH), c_k is a stability-dependent coefficient, l is the mixing length (limited by stability functions), and $\mathrm{TKE} = \frac{1}{2}\left(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}\right)$ is the turbulent kinetic energy. Note this formulation differs fundamentally from the SGS closure used in the LES mode.

3. Results and Discussions

Section 3.1 analyzes LES with prescribed striping SHF to examine the emergence of secondary circulations and to evaluate the resulting FI. FLIMK is applied diagnostically over sub-regions of varying spatial scales, where the spatially averaged SHF in the lower atmosphere reflects the influence of surface heterogeneity—similar to what tile-based schemes aim to represent at the L-A interface. Section 3.2 presents results from the ICON-SCM, configured using the ARM case with prescribed large-scale forcing. In this setup, FLIMK is implemented online to modify the SHF in the lowest atmospheric layer. The analysis focuses on how this correction influences the

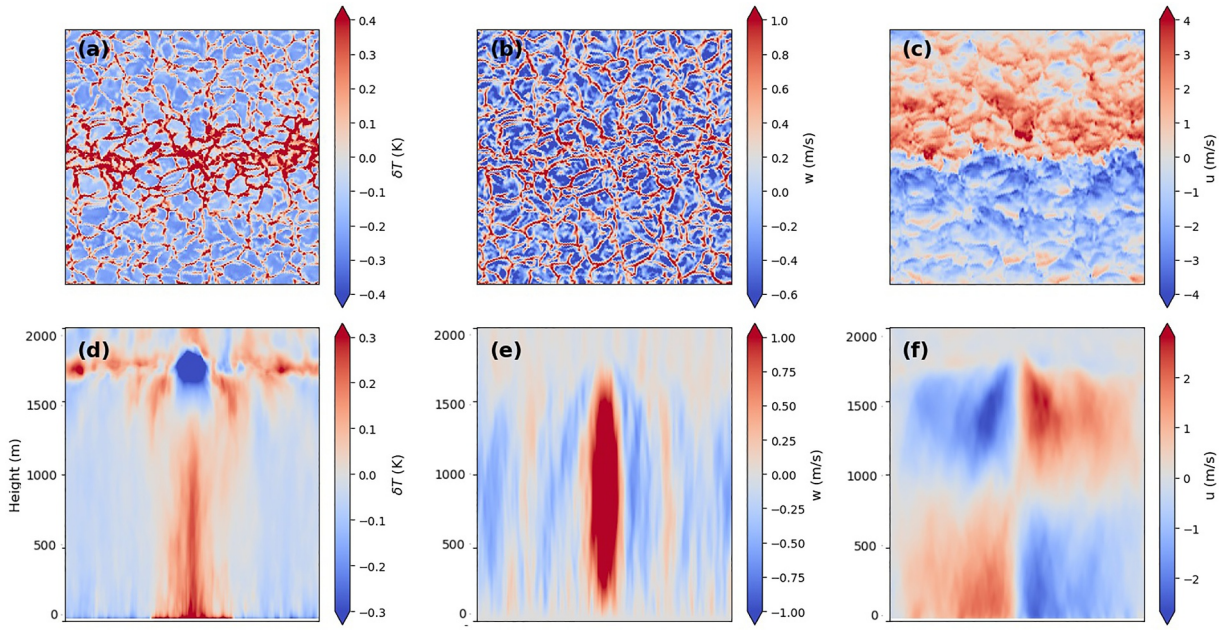

Figure 4. Horizontal cross-sections of potential temperature anomaly at 45 m (a), vertical wind speed at 50 m (b), and horizontal wind speed (u-component) at 45 m (c). Averaged vertical cross-sections of potential temperature anomaly (d), vertical wind speed (e), and horizontal wind speed (f).

surface energy balance and the vertical profile of potential temperature under realistic atmospheric conditions. Section 3.3 discusses the potential of using other FI prediction models, the numerical model stability and the limitations.

3.1. Large Eddy Simulation

Figure 4a shows the horizontal cross-section of the potential temperature anomaly at a height of 45 m. Figure 4d displays the vertical cross-section of the potential temperature anomaly, averaged along the *y*-direction. In the region with a higher SHF (see Figure 3), a greater potential temperature anomaly is observed in the lower part of the boundary layer, corresponding to a lower potential temperature anomaly at the top of the boundary layer. Figure 4e depicts a pattern of the vertical velocity characterized by a narrow region of strong updraft, compensated by a broader area of weak downdraft. Similar patterns have been well presented in the previous works (Salesky et al., 2017; B. Zhou et al., 2018). Figures 4c and 4f illustrate that horizontal wind (u-component) converge (diverge) at the lower (upper) portion of the boundary layer, indicating a well-defined secondary circulation structure. Additionally, the horizontal wind induced by the secondary circulation can reach speeds exceeding 2 m s^{-1}, comparable with the findings reported in previous LES studies (Lee et al., 2019; Zhang et al., 2023).

A series of sub-regions of varying sizes, ranging from the grid scale (50 m) to larger scales of up to 9.6 km, are selected to investigate the effect of spatial scale on the FI. Each sub-region is selected at a random location with 50 samples to ensure a broad representation of land surface heterogeneity patterns across spatial scales. Figure 5 provides an example of the 1.6 km sub-region sampling.

As illustrated in Figures 5b–5e, even overlapping sub-regions can exhibit distinct surface configurations. While some sub-regions may share similar surface conditions (e.g., uniform SHF values), they can still develop different atmospheric structures due to variations in local dynamics and turbulence characteristics.

The SHF at a height of 45 m is calculated using both the K-theory and the FLIMK and is then normalized to the prescribed SHF at the land surface, as shown in the boxplot in Figure 6. The underestimation of K-theory is consistent across all scales, with the mean value ranging from approximately 12%–16% (normalized heat flux of 0.84–0.88) for the majority of sub-regions and 15% for the entire area (i.e., 0.85). The variance of the normalized heat flux decreases with increasing size.

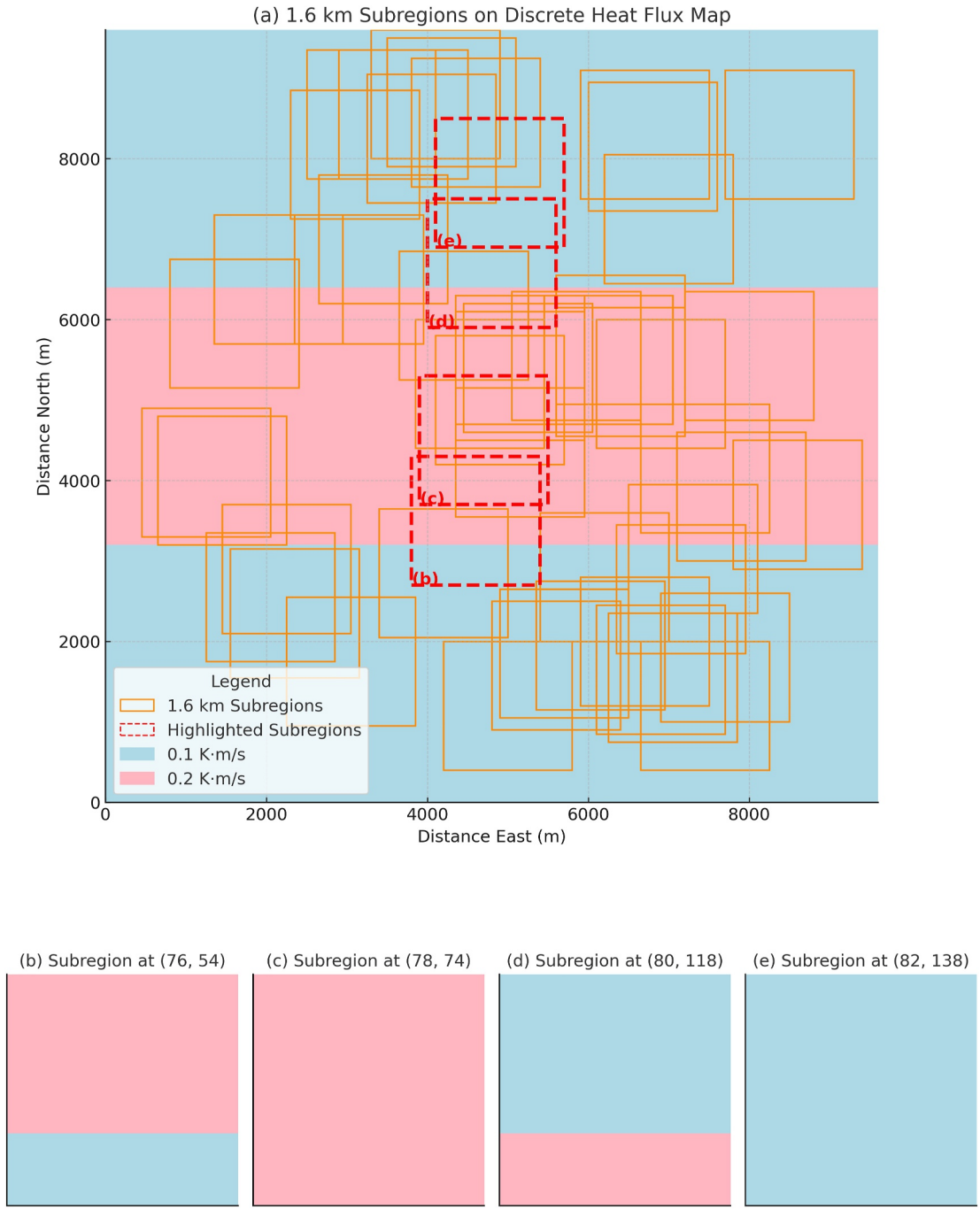

Figure 5. Example of 1.6 km sub-region sampling on a prescribed sensible heat flux (SHF) map (a). Orange boxes show all sampled sub-regions; red dashed boxes highlight selected sub-regions shown below. Background colors represent different prescribed SHF values (0.1 K m s^{-1} in blue, 0.2 K m s^{-1} in pink). Sub-region patterns corresponding to selected locations (b–e). Despite some spatial overlap, each sub-region exhibits a distinct configuration of surface heterogeneity.

Some extreme values (e.g., normalized SHF > 1) are associated with sub-regions dominated by updrafts, which only occupy a small fraction. As our focus is on improving domain-mean SHF, we did not use metrics such as root-mean square error here, which is highly sensitive to such localized extremes. The FI prediction model can be derived based on the definition of Equation 2 as FLIMK_LES or based on Equation 3 as FLIMK_De18. The underestimation of the SHF decreased from 15% to 8% (6%) for FLIMK_De18 (FLIMK_LES). It is crucial to highlight that the FLIMK does not lead to a perfect match of the fluxes. That is because the turbulent heat flux ($\overline{w'\theta'}$) is influenced by the sampling window size (i.e., averaging time), whereby a larger window size leads to a

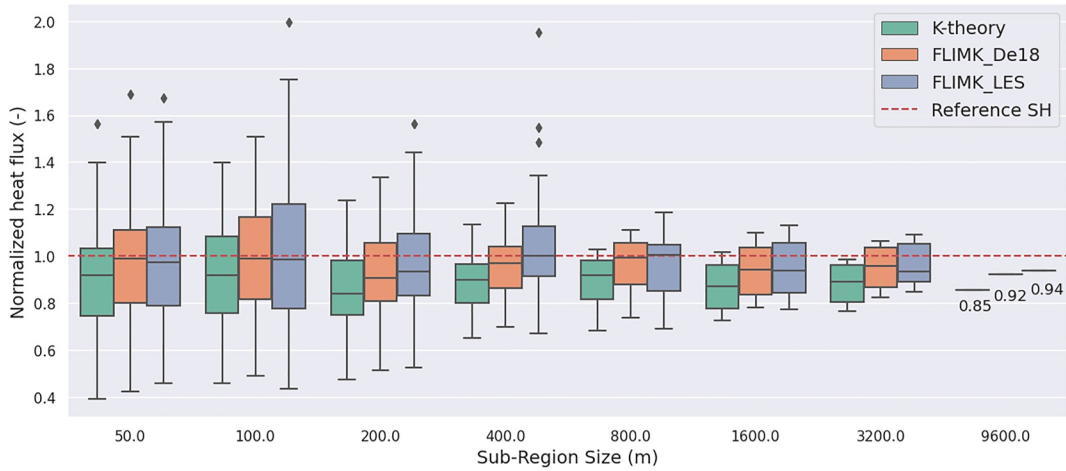

Figure 6. Normalized sensible heat flux (SHF) at 45 m height. Results are averaged over the final 60 min of the large eddy simulation to capture the quasi-steady state. Each boxplot summarizes the distribution of normalized SHF from 50 sampled sub-regions at each scale. The center line of each box indicates the median, the box spans the interquartile range (25th to 75th percentile), and the whiskers extend to 1.5 times the interquartile range. Points beyond this range are plotted as outliers. The value of unit one (red dashed line) indicates no flux imbalance, while any value smaller than unit one represents an underestimation.

larger heat flux value (Finnigan et al., 2003; Mauder et al., 2006, 2007, 2020). While the LES configuration used in this study is intentionally idealized to isolate and emphasize the role of nonlocal heat fluxes, this setup does not fully capture the complexity of real-world atmospheric conditions. To support the robustness of the FLIMK approach, we refer to complementary LES studies that evaluated the same formulation as Equation 3, originally proposed by De Roo et al. (2018). Zhang, Poll, and Kollet (2024) tested the same model under checkerboard heterogeneity with varying spatial scales, and Y. Zhou et al. (2023) explored stripe-based heterogeneity with different SHF contrasts under a fixed background wind. These studies offer further support for the applicability of the FI prediction model under diverse surface and flow regimes. Nonetheless, additional tests under more realistic or less favorable conditions are needed to fully evaluate the generality of FLIMK, and this represents an important direction for future work.

3.2. Numerical Weather Prediction

Figure 7a presents the temporal evolution of the SHF computed at 10 m (i.e., upper boundary of the lowest atmospheric layer) using both K-theory and the FLIMK scheme for the ARM single-column test case in ICON-NWP, starting approximately 1 hr after local sunrise. During the early morning hours, the convective boundary layer is not yet well developed, and nonlocal transport remains weak. Under these conditions (i.e., ∼1–2 hr after sunrise), both K-theory and FLIMK tend to overestimate SHF due to weak surface forcing. Moreover, applying FLIMK during this regime introduces unnecessary correction, potentially worsening SHF representation when local processes dominate. This highlights a key limitation of the FLIMK scheme: it is intended for use during well-developed convective conditions and should be applied with caution under stable or weakly forced regimes.

Once the convective boundary layer is established, the K-theory begins to systematically underestimate the SHF, with the bias gradually increasing and peaking around 7 hr (∼12:00 local time). FLIMK effectively reduces this FI, as shown in Figure 7b. However, FI becomes negative again after 10 hr, coinciding with a rapid decline in the prescribed SHF, perhaps due to the clouds. After 11 hr, the prescribed SHF becomes negative, but the K-theory estimate remains slightly positive—likely due to the vertical gradient of potential temperature being calculated between the two atmospheric levels rather than between the surface and atmosphere. In this regime, the FLIMK correction becomes less reliable, partly because the De18 FI formulation was derived under convective conditions and includes w_*, which is not well defined under stable or nocturnal conditions. This again illustrates that FLIMK should be applied only during periods with sufficient surface heating. Other FI prediction models that exists, will be discussed in the following section.

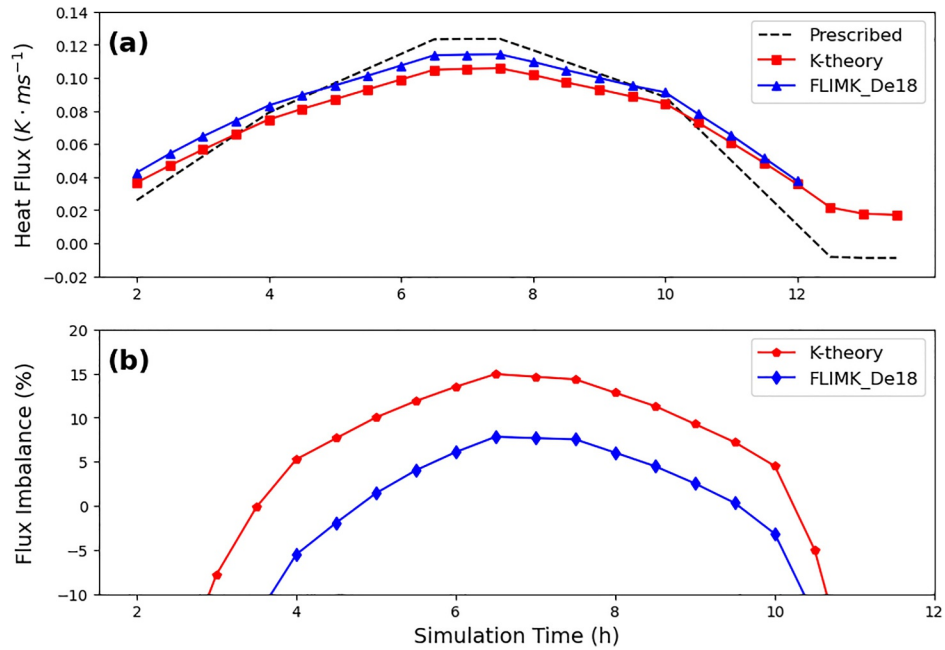

Figure 7. Temporal evolution of the sensible heat flux (a) and flux imbalance (FI) (b). The simulation begins at approximately 05:00 local time, 1 hr before sunrise. As FLux IMbalance and K-theory is designed for convective boundary layer conditions, results are shown starting 1 hr after sunrise. For visual clarity, FI values below −10% have been truncated.

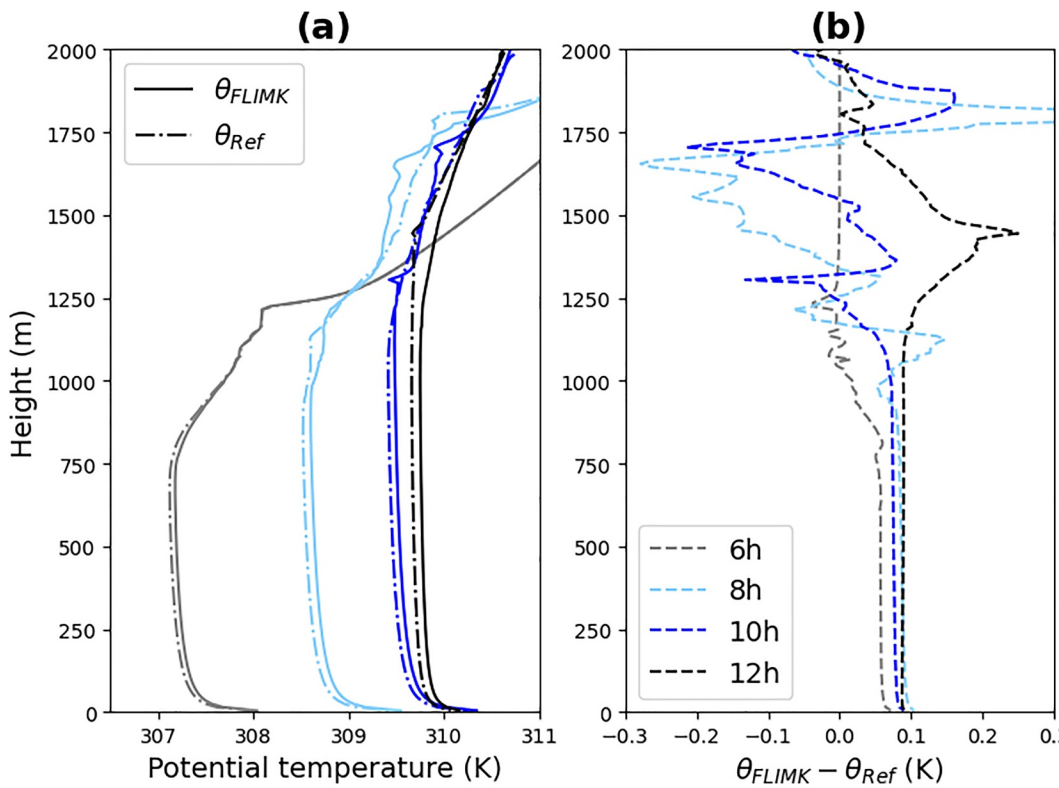

Figure 8. Vertical profile of potential temperature (a), and difference between the FLux IMbalance and K-theory runs and reference runs (b) for the ICON-NWP.

Table 1
Available Flux Imbalance (FI) Prediction Models in the Literature

FI prediction models	Reference
$\mathrm{FI} = \left[\exp\left(4.2 - 16\,\frac{u_*}{w_*}\right) + 2.1\right]\left[1.1 - 8.0\left(\frac{z}{zi} - 0.38\right)^2\right]^{0.5}$	J. Huang et al. (2008)
$\mathrm{FI} = \left[0.197\exp\left(-17.0\,\frac{u_*}{w_*}\right) + 0.156\right]\left[0.21 + 10.69\,\frac{z}{zi}\right]$	De Roo et al. (2018)
$\mathrm{FI} = 1 - \left[-1.46\,\frac{z}{zi} + 1.0\right]\left[-0.05\,\frac{zi}{\mathrm{L}}\,\frac{l_w}{\mathrm{U\,T}} + 0.95\right]$	Y. Zhou et al. (2019)
$\mathrm{FI} = \left[a\exp\left(b\,\frac{u_*}{w_*}\right) + c\right]\left[20.05\,\frac{z}{zi} + 0.157\right][H_{\mathrm{par}}]$	Wanner et al. (2022)

Note. u_* is the friction velocity, w_* is the convective velocity scale, z is the measurement height, zi is the boundary layer height, L is the Obukhov length, U is the horizontal wind speed, and T is the averaging period. $l_w = \int_0^\infty R(r)\,dr$ is the integral length scale for vertical velocity, with $R(r)$ as an autocorrelation function, $H_{\mathrm{par}} = \frac{gl_h}{U^2}\frac{\Delta\theta}{\theta}$, where l_h is the heterogeneity scale (set to 200, 400, and 800 m in the original work), and a, b, and c in the Wanner et al. (2022) are scale-dependent parameters. Table 1 adapted From Zhang, Poll, and Kollet (2024).

To assess the atmospheric response to the FLIMK correction, Figure 8a shows the vertical profile of potential temperature (θ) for both the control (i.e., Ref) and FLIMK runs. As expected, the two profiles remain nearly identical above the surface layer, indicating that FLIMK does not disrupt the broader thermodynamic structure of the boundary layer. While some oscillation appears above 1,250 m, these discrepancies diminish at higher altitudes, further indicating that the FLIMK correction has minimal influence beyond the surface layer. Figure 8b presents the difference in θ between the two runs. The difference in θ increased from 0.07 K to 0.1 K from 6 to 8 hr in the lower part of the PBL, then decreased again from 0.1 K to 0.08 K. This indicates that the model is stable, as the SHF increases the boundary layer height, leading to a decrease in the FI based on De18 (i.e., negative feedback).

3.3. Discussions

3.3.1. Flux Imbalance Prediction Models

Representing nonlocal processes, such as secondary circulations, within PBL schemes remains a major challenge. Numerous PBL schemes have been developed, each with its own strengths, but none explicitly resolve secondary circulations (H. Y. Huang et al., 2013; Milovac et al., 2016; Wang et al., 2016). To improve the representation of land surface heterogeneity, recent ESMs have incorporated surface statistical metrics such as the variance and covariance of plant functional types (Machulskaya & Mironov, 2018; M. Huang et al., 2022). However, Fowler et al. (2024) demonstrated that, even with these refinements, there is still a much weaker atmospheric response for ESMs compared to LES—highlighting that secondary circulations are governed by more complex and nonlinear dynamics than surface heterogeneity statistics alone can capture. A recent two-column framework proposed by Waterman et al. (2024) offers a promising approach to represent subgrid-scale circulations induced by surface heterogeneity, leveraging temperature contrasts between patches and vertical profiles of potential temperature and air density. This model demonstrates encouraging results for improving PBL representations, though its performance can be sensitive to parameter choices, suggesting opportunities for future refinement and broader evaluation.

In contrast, this study does not aim to develop or modify full PBL schemes. Instead, we focus specifically on improving the representation of surface-layer SHF by accounting for FI caused by secondary circulations. Motivated by previous findings that FI models can capture the impact of organized mesoscale transport (e.g., secondary circulations), this study evaluates whether incorporating such a model improves SHF representation. We test the FLIMK approach using both diagnosed FI from LES and a semi-empirical formulation from De Roo et al. (2018), both of which show promising performance. While several FI prediction models exist in the literature, a detailed intercomparison is beyond the scope of this study. We provide a brief overview to support future development and broader applications. Table 1 lists a few examples for illustrative purposes, and the interested readers are directed to Zhang, Poll, and Kollet (2024) for an evaluation work.

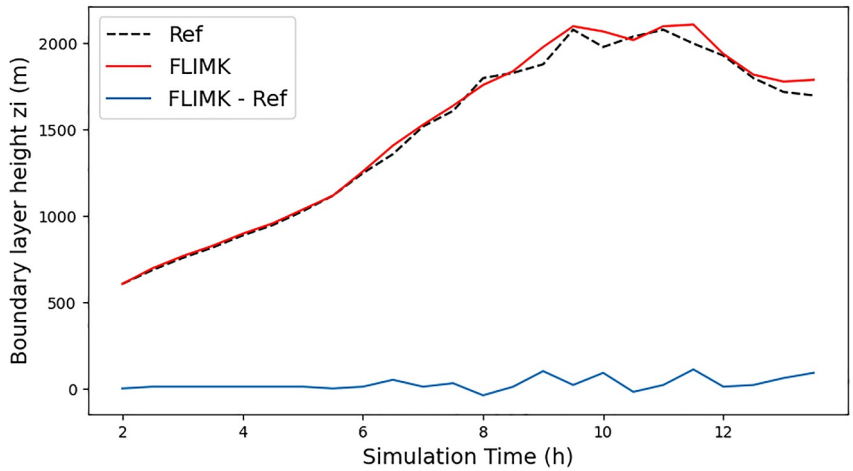

Figure 9. Boundary layer height as a function of simulation time. The boundary layer height is defined as the height where the potential temperature gradient has its maximum.

The first FI prediction model is that of J. Huang et al. (2008), which is suitable for calculating FI at heights between 0.3 *zi* to 0.5 *zi*. The work of De Roo et al. (2018) and Y. Zhou et al. (2019) is more appropriate for the lower range, below 0.1 *zi*. The heterogeneity parameters were incorporated in Wanner et al. (2022) to account for the impact of the land surface heterogeneities. Furthermore, FI prediction models for latent heat flux are documented in the literature, including the one by De Roo et al. (2018).

The subsequent phase of research will focus on the implementation of the FLIMK scheme, which will entail modifications to both sensible and latent heat fluxes. However, whether to maintain the Bowen ratio in a consistent state and the issue of reconciling FI prediction models remain open. A consistent body of evidence has reported discrepancies between the transport of heat and humidity in the atmosphere (H. Y. Huang & Margulis, 2009; Lamaud & Irvine, 2006).

3.3.2. Numerical Model Stability

Modifying the SHF in the surface layer raises the question of whether the energy input to the model might be excessively increased, potentially leading to model instability or failure. The revised SHF in the lower atmosphere through the FLIMK scheme gives rise to a modification in the potential temperature that propagates to a higher layer, as illustrated in Figure 8b. However, an increase in boundary layer height has reduced FI, which has led to negative feedback. This is further supported by the fluctuations in the difference between the boundary layer height between the K-theory and the FLIMK scheme, as demonstrated in Figure 9.

It is also important to note that in the SCM setup, the surface SHF is prescribed, which removes the interactive feedback between the land surface and atmosphere. In a fully coupled system, increased upward flux from the FLIMK correction would cool the surface, potentially reducing the SHF through a negative feedback mechanism. This effect is not represented in our current setup, which was intentionally designed to isolate the atmospheric response to FLIMK. Future work will involve testing FLIMK in a fully interactive L-A model to assess the role of surface feedback under realistic coupling conditions.

3.3.3. Limitations

While the FLIMK scheme shows promise in improving surface layer SHF representation under convective conditions, its current implementation has several limitations. First, it has only been tested under idealized surface heterogeneity (e.g., one-dimensional striping) without background wind. However, previous LES studies have shown that atmospheric responses to surface heterogeneity depend not only on spatial mean values but also on higher-order characteristics such as variance and pattern structure. For example, Han et al. (2019) demonstrated that changes in soil moisture variance—despite identical domain-mean soil moisture—can significantly impact domain-averaged surface fluxes and boundary layer structure via induced circulations. These circulations strongly influence vertical profiles of atmospheric characteristics, particularly cloud liquid water. Moreover, the

presence of a sufficiently strong background wind can suppress or disrupt secondary circulations altogether, as shown in studies by Lee et al. (2019) and Avissar and Schmidt (1998). As such, the performance of FLIMK under more realistic surface conditions and atmospheric forcings remains to be fully evaluated.

Second, non-local transport and secondary circulations influence not only the surface layer but also the upper boundary layer. In our single-column NWP test case, FLIMK produced only minor changes in the potential temperature profile, suggesting limited impact aloft. However, many existing PBL schemes already include some representation of non-local processes, for example, through counter-gradient or mass-flux terms. Applying FLIMK in parallel with such schemes raises the possibility of overlapping corrections. A key challenge moving forward is to ensure consistency and avoid double-counting of non-local contributions across different parts of the parameterization suite.

Third, FLIMK is specifically designed for convective boundary layer conditions. Like counter-gradient and mass-flux components in established PBL schemes, it targets regimes dominated by non-local transport and secondary circulations. However, applying FLIMK during stable conditions—such as nighttime, when surface heat flux becomes negative—can lead to spurious corrections and worsen the FI. This highlights the need for a regime-dependent activation mechanism, similar to what exists in hybrid PBL schemes (e.g., EDMF), where non-local terms are only applied under appropriate convective conditions. Future development should focus on establishing robust criteria for activating FLIMK in coupled long-term simulations, where boundary layer regimes evolve dynamically.

While FLIMK has been developed and evaluated in the context of nonlocal fluxes driven by land surface heterogeneity, its underlying conceptual framework—addressing FI resulting from horizontal thermal contrasts—may be applicable to other forms of mesoscale transport, such as sea–land breeze circulations. Extending FLIMK to coastal environments would require further evaluation using LES data over land–sea transition zones and may involve adapting the shape function or incorporating additional predictors. This represents a promising direction for future work.

4. Summary and Conclusions

Parameterizing the effects of secondary circulations in ESMs is important yet challenging. Building on the demonstrated robustness of the FI prediction model in quantifying nonlocal processes in heat flux calculations, this study proposes a novel approach that employs the FI prediction model to improve the representation of SHF at L-A interface. This alternative approach, FI and K-theory (FLIMK), combines the gradient diffusion approach (K-theory) for the local process and the FI prediction models for the nonlocal process.

Two different models were employed to evaluate FLIMK separately: a LES using prescribed strip SHF and a NWP model using field measurements from the ARM campaign. The results demonstrate that FLIMK reduces the FI from approximately 15% (16%) to 6% (6.7%) for LES (NWP). Moreover, the FLIMK scheme demonstrates a high level of performance across a range of scales, from 50 m to 9.6 km. Furthermore, the FLIMK scheme is numerically stable. The boundary layer height rises as the SHF increases, reducing the FI through a negative feedback mechanism.

The proposed FLIMK scheme is not intended to replace existing PBL schemes but to enhance the representation of surface layer heat fluxes. Due to its simple formulation, which requires only atmospheric stability (u_*/w_*) and relative height (z/zi), the FLIMK scheme can be easily integrated with existing PBL schemes. This offers a potential pathway for model parameterization and improvement. While FLIMK performs well under convective conditions, it tends to overcorrect in stable or weakly unstable regimes. Further refinement is needed to extend its applicability across varying stability conditions.

Appendix A

This section presents the mathematical derivation of the Flux IMbalance and the K-theory approach (FLIMK). The definition of the flux imbalance (FI), as proposed by Y. Zhou et al. (2019), is the ratio of the nonlocal part (i.e., advection, A, and dispersive flux D) to the reference heat flux at the land surface ($\overline{H}_0$) as illustrated in Equation A1. In the context of the assumption of a constant flux for the ASL, the reference heat flux can be defined as the sum of the local part ($\overline{w'\theta'}$ or $-K_h\,\partial\overline{\theta}/\partial z$) and the nonlocal part ($A + D$), see Equations A2a and

A2b. Consequently, the ratio of the local part to the reference heat flux can be derived in Equations A3, and A4 can be obtained by moving $\overline{H}_0$ to the left-hand side of the equation.

$$\mathrm{FI} = \frac{A + D}{\overline{H}_0}, \tag{A1}$$

$$\overline{H}_0 = \underbrace{\overline{w'\theta'}}_{\text{local}} + \underbrace{A + D}_{\text{nonlocal}}, \tag{A2a}$$

$$\overline{H}_0 = \underbrace{\frac{-K_h \partial \overline{\theta}}{\partial z}}_{\text{local}} + \underbrace{A + D}_{\text{nonlocal}}, \tag{A2b}$$

$$1 - \mathrm{FI} = \frac{-\frac{K_h \partial \overline{\theta}}{\partial z}}{\overline{H}_0}, \tag{A3}$$

$$\overline{H}_0 = \frac{-\frac{K_h \partial \overline{\theta}}{\partial z}}{1 - \mathrm{FI}}. \tag{A4}$$

Appendix B

To clarify the physical basis of the empirical form of the FI prediction model, we start from the definition of FI:

$$\mathrm{FI} = 1 - \frac{\langle \overline{w'\theta'} \rangle}{\langle \overline{H}_0 \rangle}, \tag{B1}$$

Where $\overline{w'\theta'}$ is the resolved vertical turbulent heat flux, and H_0 is the reference flux, the angle bracket represents the (sub-)domain mean. We can conceptualize $\langle \overline{w'\theta'} \rangle$ as the sum of contributions from two dominant transport mechanisms. We define $\langle \overline{w'\theta'} \rangle_{\text{bottom}}$ as the upward transport associated with surface-generated thermals, assumed to scale with the reference surface SHF ($\langle \overline{H}_0 \rangle_{\text{bottom}}$) and decrease linearly to zero at the boundary layer height (zi). Conversely, $\langle \overline{w'\theta'} \rangle_{\text{top}}$ represents downward transport due to entrainment at zi, assumed to scale with the heat flux at the inversion layer ($\langle \overline{H}_0 \rangle_{\text{top}}$) and decrease linearly to zero toward the surface. These linear profiles are idealized and reflect the typical vertical structure of nonlocal heat transport in convective boundary layers.

$$\langle \overline{w'\theta'} \rangle = \langle \overline{w'\theta'} \rangle_{\text{top}} + \langle \overline{w'\theta'} \rangle_{\text{bottom}}, \tag{B2}$$

Substituting into the FI definition:

$$\mathrm{FI} = 1 - \frac{\langle \overline{w'\theta'} \rangle_{\text{top}} + \langle \overline{w'\theta'} \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle} = 1 - \left(\frac{\langle \overline{w'\theta'} \rangle_{\text{top}}}{\langle \overline{H}_0 \rangle} + \frac{\langle \overline{w'\theta'} \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle} \right), \tag{B3}$$

The bottom-up and top-down contributions can be related to their respective reference fluxes as:

$$\frac{\langle \overline{w'\theta'} \rangle_{\text{top}}}{\langle \overline{H}_0 \rangle} = \frac{\langle \overline{H}_0 \rangle_{\text{top}}}{\langle \overline{H}_0 \rangle} \frac{\langle \overline{w'\theta'} \rangle_{\text{top}}}{\langle \overline{H}_0 \rangle_{\text{top}}} = \frac{\langle \overline{H}_0 \rangle_{\text{top}}}{\langle \overline{H}_0 \rangle} (1 - \mathrm{FI}_{\text{top}}), \tag{B4}$$

$$\frac{\langle \overline{w'\theta'} \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle} = \frac{\langle \overline{H}_0 \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle} \frac{\langle \overline{w'\theta'} \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle_{\text{bottom}}} = \frac{\langle \overline{H}_0 \rangle_{\text{bottom}}}{\langle \overline{H}_0 \rangle} (1 - \mathrm{FI}_{\text{bottom}}), \tag{B5}$$

that is

$$\left\langle \overline{w'\theta'} \right\rangle_{\text{top}} = \left\langle \overline{H}_0 \right\rangle_{\text{top}}(x), \tag{B6}$$

$$\left\langle \overline{w'\theta'} \right\rangle_{\text{bottom}} = \left\langle \overline{H}_0 \right\rangle_{\text{bottom}}(1 - \text{FI}_{\text{bottom}}), \tag{B7}$$

where $\left\langle \overline{H}_0 \right\rangle_{\text{top}}$ is $\left\langle \overline{H}_{\text{entrainment}} \right\rangle \left(\frac{z}{zi}\right)$ and $\left\langle \overline{H}_0 \right\rangle_{\text{bottom}}$ is $\left\langle \overline{H}_{\text{surf}} \right\rangle \left(1 - \frac{z}{zi}\right)$,

$$\left\langle \overline{w'\theta'} \right\rangle_{\text{top}} = \left\langle \overline{H}_{\text{entrainment}} \right\rangle \left(\frac{z}{zi}\right)(1 - \text{FI}_{\text{top}}), \tag{B8}$$

$$\left\langle \overline{w'\theta'} \right\rangle_{\text{bottom}} = \left\langle \overline{H}_{\text{surf}} \right\rangle \left(1 - \frac{z}{zi}\right)(1 - \text{FI}_{\text{bottom}}), \tag{B9}$$

In addition, previous studies have shown the FI is modulated by the atmospheric stability (u_*/w_*). Combining both influences leads to a shape function to the form: $\text{FI} = f_1(u_*/w_*) f_2(z/zi)$.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The Icosahedral Nonhydrostatic model (ICON) is available at https://gitlab.dkrz.de/icon/icon-model, and the version used in this study is available at https://icg4geo.icg.kfa-juelich.de/ModelSystems/tsmp_src/icon2.5.0_lagecy.git. The numerical weather prediction data (single-column mode) and the script for producing the figures for this manuscript are available from Zhang, Poll, Weinkaemmerer, and Kollet (2024).

Acknowledgments
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). We are grateful to the three anonymous reviewers for their valuable comments and suggestions, which have substantially improved the quality of this work. Open Access funding enabled and organized by Projekt DEAL.

References

Avissar, R., & Schmidt, T. (1998). An evaluation of the scale at which ground-surface heat flux patchiness affects the convective boundary layer using large-eddy simulations. *Journal of the Atmospheric Sciences*, *55*(16), 2666–2689. https://doi.org/10.1175/1520-0469(1998)055<2666:AEOTSA>2.0.CO;2

Barr, A. G., Morgenstern, K., Black, T. A., McCaughey, J. H., & Nesic, Z. (2006). Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurement of the CO2 flux. *Agricultural and Forest Meteorology*, *140*(1–4), 322–337. https://doi.org/10.1016/j.agrformet.2006.08.007

Bastak Duran, I., Köhler, M., Eichhorn-Müller, A., Maurer, V., Schmidli, J., Schomburg, A., et al. (2021). Data for the ICON single-column mode [Dataset]. *Zenodo*. https://doi.org/10.5281/ZENODO.5070234

Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J. C., Khairoutdinov, M., et al. (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, *128*(582), 1075–1093. https://doi.org/10.1256/003590002320373210

Deardorff, J. W. (1972). Numerical investigation of neutral and unstable planetary boundary layers. *Journal of the Atmospheric Sciences*, *29*(1), 91–115. https://doi.org/10.1175/1520-0469(1972)029<0091:NIONAU>2.0.CO;2

De Roo, F., & Mauder, M. (2018). The influence of idealized surface heterogeneity on turbulent flux measurements: A parameter study with large-eddy simulation. *Atmospheric Chemistry and Physics*, 5059–5074. https://doi.org/10.5194/acp-18-5059-2

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*(12), 1–23. https://doi.org/10.1371/journal.pone.0209022

Dipankar, A., Stevens, B., Heinze, R., Moseley, C., Zängl, G., Giorgetta, M., & Brdar, S. (2015). Large eddy simulation using the general circulation model ICON. *Journal of Advances in Modeling Earth Systems*, *6*(3), 963–986. https://doi.org/10.1002/2015MS000431.Received

Ďurán, I. B., Köhler, M., Eichhorn-Müller, A., Maurer, V., Schmidli, J., Schomburg, A., et al. (2021). The icon single-column mode. *Atmosphere*, *12*(7), 1–24. https://doi.org/10.3390/atmos12070906

Eder, F., De Roo, F., Rotenberg, E., Yakir, D., Schmid, H. P., & Mauder, M. (2015). Secondary circulations at a solitary forest surrounded by semi-arid shrubland and their impact on eddy-covariance measurements. *Agricultural and Forest Meteorology*, *211–212*, 115–127. https://doi.org/10.1016/j.agrformet.2015.06.001

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., & 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

Foken, T. (2008). The energy balance closure problem: An overview. *Ecological Applications: A Publication of the Ecological Society of America*, *18*(6), 1351–1367. https://doi.org/10.1890/06-0922.1

Foken, T., Aubinet, M., Finnigan, J. J., Leclerc, M. Y., Mauder, M., & Paw U, K. T. (2011). Results of a panel discussion about the energy balance closure correction for trace gases. *Bulletin of the American Meteorological Society*, *92*(4), 13–18. https://doi.org/10.1175/2011BAMS3130.1

Fowler, M. D., Neale, R. B., Waterman, T., Lawrence, D. M., Dirmeyer, P. A., Larson, V. E., et al. (2024). Assessing the atmospheric response to subgrid surface heterogeneity in the single-column community Earth system model, version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *16*(3), 1–18. https://doi.org/10.1029/2022MS003517

Franssen, H. J. H., Stöckli, R., Lehner, I., Rotenberg, E., & Seneviratne, S. I. (2010). Energy balance closure of eddy-covariance data: A multisite analysis for European FLUXNET stations. *Agricultural and Forest Meteorology*, *150*(12), 1553–1567. https://doi.org/10.1016/j.agrformet.2010.08.005

Han, C., Brdar, S., & Kollet, S. (2019). Response of convective boundary layer and shallow cumulus to soil moisture heterogeneity: A large-eddy simulation study. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4305–4322. https://doi.org/10.1029/2019ms001772

Hohenegger, C., Brockhaus, P., Bretherton, C. S., & Schär, C. (2009). The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *Journal of Climate*, *22*(19), 5003–5020. https://doi.org/10.1175/2009JCLI2604.1

Hohenegger, C., Korn, P., Linardakis, L., Redler, R., Schnur, R., Adamidis, P., et al. (2023). ICON-Sapphire: Simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, *16*(2), 779–811. https://doi.org/10.5194/gmd-16-779-2023

Huang, H. Y., Hall, A., & Teixeira, J. (2013). Evaluation of the WRF PBL parameterizations for marine boundary layer clouds: Cumulus and stratocumulus. *Monthly Weather Review*, *141*(7), 2265–2271. https://doi.org/10.1175/MWR-D-12-00292.1

Huang, H. Y., & Margulis, S. A. (2009). On the impact of surface heterogeneity on a realistic convective boundary layer. *Water Resources Research*, *45*(4), 1–16. https://doi.org/10.1029/2008WR007175

Huang, J., Lee, X., & Patton, E. G. (2008). A modelling Study of flux imbalance and the influence of entrainment in the convective boundary layer. *Boundary-Layer Meteorology*, *127*(2), 273–292. https://doi.org/10.1007/s10546-007-9254-x

Huang, M., Ma, P. L., Chaney, N. W., Hao, D., Bisht, G., Fowler, M. D., et al. (2022). Representing surface heterogeneity in land–atmosphere coupling in E3SMv1 single-column model over ARM SGP during summertime. *Geoscientific Model Development*, *15*(16), 6371–6384. https://doi.org/10.5194/gmd-15-6371-2022

Johansson, C., Smedman, A.-S., Högström, U., Brasseur, J. G., & Khanna, S. (2001). Critical Test of the validity of Monin–Obukhov similarity during convective conditions. *Journal of the Atmospheric Sciences*, *58*(12), 1549–1566. https://doi.org/10.1175/1520-0469(2001)058<1549:ctotvo>2.0.co;2

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific Data*, *6*(1), 1–14. https://doi.org/10.1038/s41597-019-0076-8

Jungclaus, J. H., Lorenz, S. J., Schmidt, H., Brovkin, V., Brüggemann, N., Chegini, F., et al. (2022). The ICON earth system model version 1.0. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002813. https://doi.org/10.1029/2021ms002813

Kanda, M. (2006). Large-eddy simulations on the effects of surface geometry of building arrays on turbulent organized structures. *Boundary-Layer Meteorology*, *118*(1), 151–168. https://doi.org/10.1007/s10546-005-5294-2

Kang, S. L., Davis, K. J., & LeMone, M. (2007). Observations of the ABL structures over a heterogeneous land surface during IHOP_2002. *Journal of Hydrometeorology*, *8*(2), 221–244. https://doi.org/10.1175/JHM567.1

Khanna, S., & Brasseur, J. G. (1997). Analysis of Monin-Obukhov similarity from large-eddy simulation. *Journal of Fluid Mechanics*, *345*, 251–286. https://doi.org/10.1017/s0022112097006277

Lamaud, E., & Irvine, M. (2006). Temperature-humidity dissimilarity and heat-to-water-vapour transport efficiency above and within a pine forest canopy: The role of the Bowen ratio. *Boundary-Layer Meteorology*, *120*(1), 87–109. https://doi.org/10.1007/s10546-005-9032-6

Lee, J. M., Zhang, Y., & Klein, S. A. (2019). The effect of land surface heterogeneity and background wind on shallow cumulus clouds and the transition to deeper convection. *Journal of the Atmospheric Sciences*, *76*(2), 401–419. https://doi.org/10.1175/JAS-D-18-0196.1

Lenderink, G., & Holtslag, A. A. M. (2004). An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. *Quarterly Journal of the Royal Meteorological Society*, *130 C*(604), 3405–3427. https://doi.org/10.1256/qj.03.117

Lin, H., Li, Y., & Zhao, L. (2022). Partitioning of sensible and latent heat fluxes in different vegetation types and their spatiotemporal variations based on 203 FLUXNET sites. *Journal of Geophysical Research, D: Atmospheres*, *127*(21), e2022JD037142. https://doi.org/10.1029/2022jd037142

Machulskaya, E., & Mironov, D. (2018). Boundary conditions for scalar (Co)Variances over heterogeneous surfaces. *Boundary-Layer Meteorology*, *169*(1), 139–150. https://doi.org/10.1007/s10546-018-0354-6

Mauder, M., Foken, T., & Cuxart, J. (2020). Surface-Energy-Balance closure over land: A review. *Boundary-Layer Meteorology*, *177*(2), 395–426. https://doi.org/10.1007/s10546-020-00529-6

Mauder, M., Genzel, S., Fu, J., Kiese, R., Soltani, M., Steinbrecher, R., et al. (2018). Evaluation of energy balance closure adjustment methods by independent evapotranspiration estimates from lysimeters and hydrological simulations. *Hydrological Processes*, *32*(1), 39–50. https://doi.org/10.1002/hyp.11397

Mauder, M., Ibrom, A., Wanner, L., De Roo, F., Brugger, P., Kiese, R., & Pilegaard, K. (2021). Options to correct local turbulent flux measurements for large-scale fluxes using an approach based on large-eddy simulation. *Atmospheric Measurement Techniques*, *14*(12), 7835–7850. https://doi.org/10.5194/amt-14-7835-2021

Mauder, M., Jung, M., Stoy, P., Nelson, J., & Wanner, L. (2024). Energy balance closure at FLUXNET sites revisited. *Agricultural and Forest Meteorology*, *358*(110235), 110235. https://doi.org/10.1016/j.agrformet.2024.110235

Mauder, M., Liebethal, C., Göckede, M., Leps, J. P., Beyrich, F., & Foken, T. (2006). Processing and quality control of flux data during LITFASS-2003. *Boundary-Layer Meteorology*, *121*(1), 67–88. https://doi.org/10.1007/s10546-006-9094-0

Mauder, M., Oncley, S. P., Vogt, R., Weidinger, T., Ribeiro, L., Bernhofer, C., et al. (2007). The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post-field data processing methods. *Boundary-Layer Meteorology*, *123*(1), 29–54. https://doi.org/10.1007/s10546-006-9139-4

Meredith, L. K., Commane, R., Munger, J. W., Dunn, A., Tang, J., Wofsy, S. C., & Prinn, R. G. (2014). Ecosystem fluxes of hydrogen: A comparison of flux-gradient methods. *Atmospheric Measurement Techniques*, *7*(9), 2787–2805. https://doi.org/10.5194/amt-7-2787-2014

Milovac, J., Warrach-Sagi, K., Behrendt, A., Späth, F., Ingwersen, J., & Wulfmeyer, V. (2016). Investigation of PBL schemes combining the WRF model simulations with scanning water vapor differential absorption LiDAR measurements. *Journal of Geophysical Research*, *121*(2), 624–649. https://doi.org/10.1002/2015JD023927

Paleri, S., Desai, A. R., Metzger, S., Durden, D., Butterworth, B. J., Mauder, M., et al. (2022). Space-Scale resolved surface fluxes across a heterogeneous, mid-latitude forested landscape. *Journal of Geophysical Research, D: Atmospheres*, *127*(23), e2022JD037138. https://doi.org/10.1029/2022JD037138

Pleim, J. E. (2007). A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, *46*(9), 1383–1395. https://doi.org/10.1175/JAM2539.1

Raasch, S., & Harbusch, G. (2001). An analysis of secondary circulations and their effects caused by small-scale surface inhomogeneities using large-eddy simulation. *Boundary-Layer Meteorology*, *101*(1), 31–59. https://doi.org/10.1023/A:1019297504109

Raschendorfer, M. (2001). The new turbulence parameterization OfLM. *COSMO Newsl*, *1*, 89–97. Retrieved from https://www.cosmo-model.org/content/model/documentation/newsLetters/newsLetter01/newsLetter_01.pdf

Salesky, S. T., Chamecki, M., & Bou-Zeid, E. (2017). On the Nature of the transition between roll and cellular Organization in the convective boundary layer. *Boundary-Layer Meteorology*, *163*(1), 41–68. https://doi.org/10.1007/s10546-016-0220-3

Sedlar, J., Riihimaki, L. D., Turner, D. D., Duncan, J., Adler, B., Bianco, L., et al. (2022). Investigating the impacts of daytime boundary layer clouds on surface energy fluxes and boundary layer structure during CHEESEHEAD19. *Journal of Geophysical Research, D: Atmospheres*, *127*(5), 1–24. https://doi.org/10.1029/2021JD036060

Stevens, B. (2007). On the growth of layers of nonprecipitating cumulus convection. *Journal of the Atmospheric Sciences*, *64*(8), 2916–2931. https://doi.org/10.1175/JAS3983.1

Stoy, P. C., Mauder, M., Foken, T., Marcolla, B., Boegh, E., Ibrom, A., et al. (2013). A data-driven analysis of energy balance closure across FLUXNET research sites: The role of landscape scale heterogeneity. *Agricultural and Forest Meteorology*, *171–172*, 137–152. https://doi.org/10.1016/j.agrformet.2012.11.004

Sullivan, P. P., Moeng, C. H., Stevens, B., Lenschow, D. H., & Mayor, S. D. (1998). Structure of the entrainment zone capping the convective atmospheric boundary layer. *Journal of the Atmospheric Sciences*, *55*(19), 3042–3064. https://doi.org/10.1175/1520-0469(1998)055<3042:SOTEZC>2.0.CO;2

Van Heerwaarden, C. C., Van Stratum, B. J. H., Heus, T., Gibbs, J. A., Fedorovich, E., & Mellado, J. P. (2017). MicroHH 1.0: A computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows. *Geoscientific Model Development*, *10*(8), 3145–3165. https://doi.org/10.5194/gmd-10-3145-2017

Van Pham, T., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., et al. (2020). ICON in Climate Limited-area Mode (ICON Release version 2.6.1): A new regional climate model. *Geoscientific Model Development Discussions*. https://doi.org/10.5194/gmd-2020-20

VanZanten, M. C., Duynkerke, P. G., & Cuijpers, J. W. M. (1999). Entrainment parameterization in convective boundary layers. *Journal of the Atmospheric Sciences*, *56*(6), 813–828. https://doi.org/10.1175/1520-0469(1999)056<0813:EPICBL>2.0.CO;2

Wang, W., Shen, X., & Huang, W. (2016). A comparison of boundary-layer characteristics simulated using different parametrization schemes. *Boundary-Layer Meteorology*, *161*(2), 375–403. https://doi.org/10.1007/s10546-016-0175-4

Wanner, L., Calaf, M., & Mauder, M. (2022). Incorporating the effect of heterogeneous surface heating into a semi-empirical model of the surface energy balance closure. *PLoS One*, *17*(6), e0268097. https://doi.org/10.1371/journal.pone.0268097

Waterman, T., Bragg, A. D., Hay-Chapman, F., Dirmeyer, P. A., Fowler, M. D., Simon, J., & Chaney, N. (2024). A two-column model parameterization for subgrid surface heterogeneity driven circulations. *Journal of Advances in Modeling Earth Systems*, *16*(5), 1–22. https://doi.org/10.1029/2023ms003936

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., et al. (2002). Energy balance closure at FLUXNET sites. *Agricultural and Forest Meteorology*, *113*(1–4), 223–243. https://doi.org/10.1016/s0168-1923(02)00109-0

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. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 563–579. https://doi.org/10.1002/qj.2378

Zhang, L., Poll, S., & Kollet, S. (2023). Large-eddy simulation of soil moisture heterogeneity-induced secondary circulation with ambient winds. *Quarterly Journal of the Royal Meteorological Society*, *149*(vember 2022), 1–17. https://doi.org/10.1002/qj.4413

Zhang, L., Poll, S., & Kollet, S. (2024). Assessing the performance of flux imbalance prediction models using large eddy simulations over heterogeneous land surfaces. *Boundary-Layer Meteorology*, *190*(10), 43. https://doi.org/10.1007/s10546-024-00880-y

Zhang, L., Poll, S., Weinkaemmerer, J., & Kollet, S. (2024). An alternative way to parameterizing the nonlocal scale sensible heat flux using the flux imbalance and K-theory approach [Dataset]. *Figshare*. https://doi.org/10.6084/m9.figshare.27304923.v1

Zhou, B., Sun, S., Yao, K., & Zhu, K. (2018). Reexamining the gradient and countergradient representation of the local and nonlocal heat fluxes in the convective boundary layer. *Journal of the Atmospheric Sciences*, *75*(7), 2317–2336. https://doi.org/10.1175/JAS-D-17-0198.1

Zhou, Y., Li, D., & Li, X. (2019). The effects of surface heterogeneity scale on the flux imbalance under free convection. *Journal of Geophysical Research, D: Atmospheres*, *124*(15), 8424–8448. https://doi.org/10.1029/2018JD029550

Zhou, Y., & Li, X. (2019). Energy balance closures in diverse ecosystems of an endorheic river basin. *Agricultural and Forest Meteorology*, *274*, 118–131. https://doi.org/10.1016/j.agrformet.2019.04.019

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. *Agricultural and Forest Meteorology*, *333*, 109382. https://doi.org/10.1016/j.agrformet.2023.109382