Boundary-Layer Meteorology
https://doi.org/10.1007/s10546-020-00513-0

RESEARCH ARTICLE

Can Data Mining Help Eddy Covariance See the Landscape? A Large-Eddy Simulation Study

Ke Xu[2,3] · Matthias Sühring[4] · Stefan Metzger[1,2] · David Durden[1] · Ankur R. Desai[2]

Received: 15 July 2018 / Accepted: 13 March 2020

Abstract
Eddy-covariance fluxes serve as an essential benchmark for Earth system models and remote sensing data. However, two challenges prevent model-data intercomparisons from fully utilizing eddy-covariance fluxes. The first challenge stems from the differing and variable spatial representativeness of the eddy-covariance measurements, or footprint bias and transience. The second originates from the phenomenon of a non-closed energy balance using eddy-covariance measurements, hypothesized to result from unaccounted mesoscale flows or under-sampling of hot spots by flux towers, among others. Previous studies have suggested that these two problems can be mitigated by either building multiple towers or by applying space–time rectification approaches, such as the environmental response function (ERF) approach. Here we ask: (1) How many eddy-flux towers do we need to sufficiently rectify location bias, close the energy budget, and sample the regional domain? (2) Can an advanced space–time rectification approach reduce the tower density, while still adequately sampling the regional flux domain? Furthermore, (3) How accurately can the ERF approach retrieve the surface-flux variation? To answer these questions, we used data from a large-eddy simulation of atmospheric flows above a heterogeneous surface as captured by an ensemble of virtual tower measurements. We calculated eddy-covariance fluxes by spatial and spatio-temporal methods. The spatial eddy-covariance method captured 89% of the prescribed total surface energy flux with about one tower per 15 km^2, while the spatio-temporal method required only one tower per 40 km^2 to capture 95% of surface energy. To capture 97% of energy, applying the ERF approach further reduced the required tower density to one tower per 200 km^2, as a result of space–time rectification and incorporating mesoscale flows. This approach also enabled retrieving the surface spatial variation of the sensible heat flux. The results provide a reference for informing and designing future observation systems based on flux tower clusters, and scale-aware data products.

Keywords Eddy covariance · Energy balance · Footprint · Large-eddy simulation · Upscaling

✉ Ke Xu
xuke2012abroad@gmail.com

Extended author information available on the last page of the article

Published online: 11 April 2020

1 Introduction

The eddy-covariance method, in theory, provides reliable, direct, spatially-distributed, and temporally continuous observations for surface–atmosphere exchange of carbon, water, and energy across contrasting eco-climate regions (Baldocchi et al. 2001). Now eddy-covariance observations, such as those collected by Fluxnet, AmeriFlux, Integrated Carbon Observation System (ICOS), OzFlux, and the National Ecological Observatory Network (NEON), and others, have become available at unprecedented temporal duration and over distributed spatial extents (Novick et al. 2018). Near continuous (30-min, 60-min) data on the exchanges of carbon, water, heat, and momentum are collected at eddy-covariance towers across the globe, where the longest running towers are now approaching three decades of observations (Baldocchi 2008). This makes eddy covariance to be one of the most important datasets for benchmarking Earth system models (ESMs) (Running et al. 1999; Bonan 2008). For example, extensive land-surface model intercomparisons diagnosed limitations in the models, including spring phenology (Richardson et al. 2012), light-use efficiency (Schaefer et al. 2012), and drought sensitivity (Schwalm et al. 2010). Wavelet coherence spectral analyses suggest that models have consistent biases at the diurnal and interannual time scales, for which eddy-covariance tower observations are well-suited to validate and improve land-surface parametrizations (Dietze et al. 2011; Stoy et al. 2013).

Two challenges continue to limit how well eddy covariance can reliably inform model-data comparison. The first one stems from its spatial representativeness, or footprint bias. The sampling footprint of eddy-covariance towers varies rapidly in time. In addition, the footprint spatial scale (0.1–10 km^2) is typically mismatched with the resolution of most Earth system models (10^1–10^4 km^2). The spatio-temporal mismatch complicates the interpretation of model evaluation against eddy-covariance observations. Previous studies demonstrated that towers tend to under-sample strong convective areas and over-sample cold subsidence areas, which can be another source of location bias (Wyngaard and Brost 1984; Moeng and Wyngaard 1984; Foken 2008; Kenny et al. 2017). The second longstanding challenge of eddy covariance is energy balance non-closure, whereby the observed sum of the turbulent sensible and latent heat fluxes is 10–30% less than the sum of the available energy, when it is expected to follow conservation of energy at sufficient averaging time or length scales (Foken et al. 2011). One leading hypothesis is that low-frequency mesoscale contributions are hidden in the advection term, and inherently not captured by the eddy-covariance method using traditional calculation methods (e.g. Finnigan et al. 2003; Kanda et al. 2004; Foken 2008; Eder et al. 2015). Atmospheric structures, i.e., organized turbulent structures (Finnigan et al. 2003; Kanda et al. 2004) and/or mesoscale flows associated with surface heterogeneity (Schlegel et al. 2014; Eder et al. 2015; De Roo and Mauder 2018), may produce low-frequency contributions.

Previous studies found that two approaches better represent the mesoscale flows in eddy-covariance measurements that help close the energy budget: calculation of the eddy-covariance flux with multiple towers (Steinfeld et al. 2007; Mauder et al. 2008) or applying advanced upscaling approaches to single or clusters of towers (Metzger et al. 2013; Xu et al. 2017a, b). Thus far, these proposed approaches have not been systematically evaluated with respect to their spatial mismatch and energy imbalance.

In this study, we performed a "virtual-flux-tower" investigation to evaluate the inferred fluxes against prescribed surface fluxes. We used large-eddy simulation (LES) over a parametrized heterogeneous surface and performed an ensemble of virtual tower measurements under controlled surface and meteorological conditions. This way, the projected

surface-flux maps can be directly evaluated against the prescribed surface fluxes in the LES. Domain surface fluxes were calculated by both spatial eddy-covariance and spatio-temporal eddy-covariance approaches (Steinfeld et al. 2007; Mauder et al. 2008), as well as with the environmental response function (ERF, Metzger et al. 2013; Xu et al. 2017a, b; Metzger 2017).

The ERF approach is a technique that has been previously shown to address location bias and energy imbalance underlying traditional implementations of the eddy-covariance method (Metzger et al. 2013; Xu et al. 2017a, b; Metzger 2017). The ERF approach relies on frequency-based decomposition of eddy fluxes, flux-footprint modeling, and machine-learning based upscaling. Previous studies have assessed its performance in retrieving the domain mean sensible heat flux under different situations over a heterogeneous surface (Xu et al. 2017a). But its ability to retrieve spatial variation in an idealized heterogeneous surface scheme and evaluation of its performance against previous upscaling methods has not been investigated, and are tested herein.

Here we ask:

- How many eddy-flux towers are needed to sufficiently rectify location bias, close the energy budget, and sample the "true" regional mean surface energy fluxes when applying the assorted spatial and spatio-temporal upscaling approaches?
- Can an advanced scaling approach reduce this observational requirement, while still adequately sampling the regional flux domain?
- How accurately can advanced upscaling approaches retrieve the spatial pattern of surface fluxes?

The paper is organized as follows: Sect. 2 describes the LES model, the simulation set-up and virtual tower measurements, as well as the applied analysis techniques. Results and discussion are presented in Sect. 3, while Sect. 4 gives a summary and conclusions.

2 Methodology

2.1 Large-Eddy Simulation Set-Up and Virtual Tower Measurements

Previous LES studies have had difficulty detecting and distinguishing the source of systematic errors in tower measurements without prescribed surface boundary conditions (e.g. Sakai et al. 2001). Our study aims to reveal the source of different systematic errors and their magnitude to improve measurement strategies using a LES model with a heterogeneous surface forcing set-up.

For the numerical simulations the LES model PALM (Raasch and Schröter 2001; Maronga et al. 2015) is used. The surface forcing was set to be stripe-wise alternating patches of higher sensible heat flux H (and corresponding lower latent heat flux LE) and lower H (with higher LE flux) along the x-axis with patch width $L = 750$ m, hereafter referred to as warm–dry and cold–wet patches, respectively. As illustrated in Fig. 1, along the y-axis, the model surface was homogeneous; the domain-averaged surface heat fluxes were constant in time with a value of $H = 100$ W m^{-2} and $LE = 200$ W m^{-2}. The stripe-wise pattern of surface patches has H alternating between 150 and 50 W m^{-2}, with LE alternating between 150 and 250 W m^{-2} over the warm–dry and cold–wet patch, respectively. The LES model set-up was identical to that described in Sühring et al. (2019). The horizontal flow along the surface pattern stripes causes mesoscale flows to develop with the ascending flows over the warm–dry patches, and the subsiding flows over the cold–wet patches, as indicated in Fig. 1c.

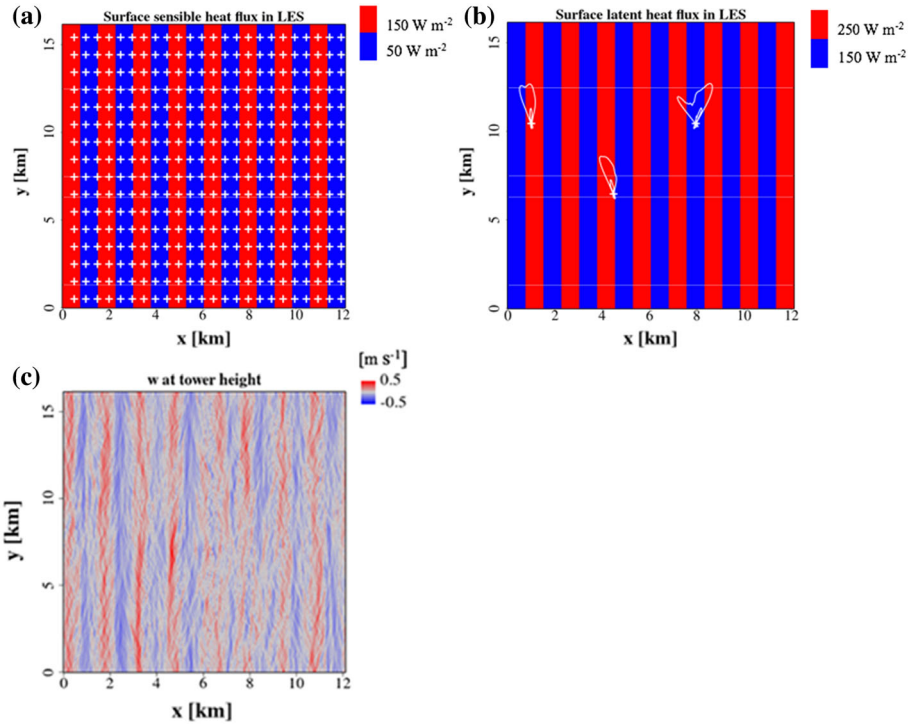

Fig. 1 Horizontal cross-sections of the heterogeneous surface fluxes of **a** sensible heat and **b** latent heat. Further, **c** shows the vertical velocity at virtual tower height of 49 m averaged over the analysis time. White crosses in **a** indicate virtual towers. White contour lines in **b** show the 90% and 30% cumulative flux footprint extent at different virtual tower positions above the center of the cold patch (upper-left sensor), at the transition between cold and warm patch (lower-central sensor), and above the center of the warm patch (upper-right sensor), all at $z = 49$ m

The PALM model was set up over a domain of 12 km × 16 km × 3.5 km at an isotropic grid spacing of 7-m and at a constant timestep of 0.3 s. The simulations ran for 5 h of simulation time, while data analysis started after the end of the second hour, with cyclic boundary conditions applied at the lateral boundaries. The detailed LES set-up and boundary-layer structure can be found in Sühring et al. (2019). A total of 384 virtual towers were placed 500 m apart on the x-axis and 1 km apart on the y-axis (see Fig. 1). Vertical wind speed, w, water vapour mixing ratio, q, and potential temperature, θ, were sampled at a height of 49 m above the model surface, i.e., the model's seventh vertical level at which the subgrid contribution to the total flux is less than 1%.

Footprints for the virtual towers were calculated following Kljun et al. (2004) and Metzger et al. (2013), where the footprint 90% upwind fetch can be as large as 2–3 km (Fig. 1). Several virtual towers see the local patch (e.g., the middle white footprint contour in Fig. 1b), while other towers see the mixture of multiple patches (e.g., the left and the right white footprint contour in Fig. 1b).

Storage fluxes were calculated using 1-h averaged x–z cross-sections of potential temperature and humidity. Averaging over the homogeneous y-axis was applied because storing potential temperature and humidity at each x–y–z point at 7-m spatial and 0.3-s temporal resolutions for the control volume creates data storage issues for the LES model. Homogeneity

along the y axis is a fair assumption in this LES set-up. We calculated the storage flux as follows: (1) Calculate the time rate-of-change at each (x, z) location using the 1-h averaged scalar at hour 5 minus the average at hour 3; and (2) Integrate the vertically-resolved time rate-of-change over the z-axis between the surface and the tower height at $z = 49$ m.

2.2 Analysis

We used three different techniques to determine domain-averaged and patch fluxes from the virtual tower measurements during the analysis time, i.e. from hour 3 to hour 5: spatial eddy-covariance method (Sect. 2.2.1), spatio-temporal eddy-covariance method (S07 and M08, i.e., Sect. 2.2.2), and a space–time rectification method, the ERF method (Sect. 2.2.3), respectively.

2.2.1 Spatial Eddy Covariance

For a single tower measurement, traditional eddy-covariance fluxes are calculated in a temporal framework, where turbulent fluctuations are defined as deviations from the temporal mean state. With multiple towers, however, the mean state used for eddy-covariance representative fluxes can also be determined in a spatial framework, which we refer to as the spatial eddy-covariance approach throughout, viz.

$$F_{spatial\ EC} = \overline{\left[w_s'(x_m, y_m, z_1, t)c_s'(x_m, y_m, z_1, t)\right]}, \tag{1}$$

where

$$w_s'(x_m, y_m, z_1, t) = w(x_m, y_m, z_1, t) - [w], \tag{2}$$

$$c_s'(x_m, y_m, z_1, t) = c(x_m, y_m, z_1, t) - [c]. \tag{3}$$

Here, c represents a scalar, either potential temperature, θ, or water vapour mixing ratio, q, while $c_s{}'$ indicates the deviation from the reference state. Squared brackets denote the horizontal average of selected towers over the horizontal domain. In contrast the overbar, e.g. $\bar{c}$ in Eq. 5 denotes the temporal average of a quantity over 3 h; e.g.,

$$[c] = \frac{\sum_{m=1}^{m=n} c(x_m, y_m, z_1, t)}{n}, \tag{4}$$

$$\bar{c} = \frac{\sum_{t=2h}^{t=5h} c(x_m, y_m, z_1, t)}{\frac{3h}{0.3s}}, \tag{5}$$

where x_m, y_m indicates the tower location at domain, horizontal dimension, z_1 is the virtual tower height at $z = 49$ m, $n = \{1…14\}$ is the number of towers used for a given analysis, and t is time; 0.3 s is the temporal resolution, and 3 h is the analysis time, from the end of second hour to the fifth hour.

For each number of towers $n = \{1…14\}$ we evaluated an ensemble of 50,000 runs, each using a randomly-chosen subset of n towers from evenly and unevenly distributed virtual towers to obtain a stable ensemble flux. Here, "evenly distributed" means, that of the 320 towers, 50% were located in warm–dry patches, and 50% in cold–wet patches. On the other hand, "unevenly distributed" or "stratified" means that only 40% of the towers were chosen from the warm–dry patches, and 60% from the cold–wet patches.

2.2.2 Spatio-Temporal Eddy Covariance

Different from the spatial eddy-covariance method subtracting the spatial average of w and c, the spatio-temporal eddy-covariance method calculated fluxes by subtracting the spatio-temporal average of w and c (Eq. 6). Note that the subtraction term is the spatio-temporal average over each 30-min, as typical eddy covariance computes 30-min fluxes. We then averaged the six 30-min fluxes over the 3-h analysis time to obtain the representative flux over the whole analysis period,

$$F_{spatio-temporal\ EC} = \overline{\left[\overline{\overline{w'_{st}(x_m, y_m, z_1, t) \cdot c'_{st}(x_m, y_m, z_1, t)}}\right]}, \tag{6}$$

where

$$w'_{st}(x_m, y_m, z_1, t) = w(x_m, y_m, z_1, t) - \left[\overline{\overline{w}}\right], \tag{7}$$

and,

$$c'_{st}(x_m, y_m, z_1, t) = c(x_m, y_m, z_1, t) - \left[\overline{\overline{c}}\right], \tag{8}$$

and where

$$\left[\overline{\overline{w}}\right] = \frac{\sum_{m=1}^{m=n} \frac{\sum_{t=t_0}^{t=t_0+0.5h} w(x_m, y_m, z_1, t)}{\frac{0.5h}{0.3s}}}{n}. \tag{9}$$

Note that the double overbar, e.g. $\overline{\overline{w}}$, denotes the temporal average over each 30 min, while the single overbar denotes the temporal average over the entire 3-h analysis time. $\left[\overline{\overline{w}}\right]$ is thus the spatio-temporal average over each 30-min. t_0 is the start time of the 30-min window, i.e. t_0 can be hour 2, 2.5, 3, 3.5, 4, or 4.5. When using Reynolds decomposition for Eq. 6,

$$F_{spatio-temporal\ EC} = \left[\overline{\overline{\left(w(x_m, y_m, z_1, t) - \left[\overline{\overline{w}}\right]\right) \cdot \left(c(x_m, y_m, z_1, t) - \left[\overline{\overline{c}}\right]\right)}}\right]. \tag{10}$$

After expanding,

$$\begin{aligned} F_{\text{spatio - temporal EC}} &= \overline{\left[\overline{\overline{wc - \left[\overline{\overline{w}}\right]c - w\left[\overline{\overline{c}}\right] + \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]}}\right]} \\ &= \overline{\left[\overline{\overline{wc}}\right] - \left[\overline{\overline{\left[\overline{\overline{w}}\right]c}}\right] - \left[\overline{\overline{w\left[\overline{\overline{c}}\right]}}\right] + \left[\overline{\overline{\left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]}}\right]}. \end{aligned} \tag{11}$$

In each run, $\left[\overline{\overline{w}}\right]$ is w spatially averaged over the same virtual towers and temporally averaged over the same half an hour, therefore $\left[\overline{\overline{w}}\right]$ is a constant in each run. This makes $\left[\overline{\overline{\left[\overline{\overline{w}}\right]c}}\right] = \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]$. Similarly, $\left[\overline{\overline{c}}\right]$ is a constant in each run, therefore $\left[\overline{\overline{w\left[\overline{\overline{c}}\right]}}\right] = \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]$. Equation 10 then becomes

$$\begin{aligned} F_{\text{spatio - temporal EC}} &= \overline{\left[\overline{\overline{wc - \left[\overline{\overline{w}}\right]c - w\left[\overline{\overline{c}}\right] + \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]}}\right]} \\ &= \overline{\left[\overline{\overline{wc}}\right] - \left[\overline{\overline{\left[\overline{\overline{w}}\right]c}}\right] - \left[\overline{\overline{w\left[\overline{\overline{c}}\right]}}\right] + \left[\overline{\overline{\left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]}}\right]} \\ &= \overline{\left[\overline{\overline{wc}}\right] - \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right] - \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right] + \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]} \\ &= \overline{\left[\overline{\overline{wc}}\right] - \left[\overline{\overline{w}}\right]\left[\overline{\overline{c}}\right]}. \end{aligned} \tag{12}$$

This is the same equation introduced by Steinfeld et al. (2007), and from Eq. 23 in Steinfeld et al. (2007), the representative flux is calculated as

$$\begin{aligned} F_{S07} &= \overline{\left[w\left([c]-\left[\bar{c}\right]\right)\right]} + \overline{[w(c-[c])]} \\ &= \left[\overline{w[c]}\right] - \left[\overline{w}\left[\bar{c}\right]\right] + \left[\overline{wc}\right] - \left[\overline{w[c]}\right] \\ &= \overline{[wc]} - \left[\overline{w}\right]\left[\bar{c}\right]. \end{aligned} \tag{13}$$

As F_{S07} and the $F_{spatio-temporalEC}$ are identical, we call the S07 approach "spatio-temporal eddy covariance". This equation has been applied in several studies, e.g. Bohrer et al. (2009) and Mauder et al. (2008). Mauder et al. (2008) built on S07, but only one tower had high-rate samples for eddy-covariance-derived fluxes, while the remainder of the towers measured mean atmospheric state quantities only (wind velocity, temperature, moisture, etc.). Therefore, here we call the S07 and M08 approaches by a joint name "spatio-temporal eddy covariance". Note that when only one tower is used, S07 is identical to the traditional eddy-covariance approach. The averaging time period is 30 min for S07 and M08, as typical eddy-covariance computes 30-min fluxes. We averaged the six 30-min fluxes over the 3-h analysis time. The calculation was repeated 50,000 times by randomly choosing different subsets of towers each time from evenly- and unevenly-placed virtual towers to obtain a stable ensemble representative flux.

2.2.3 Environmental Response Function

One purpose of our study is to test how accurately the ERF approach can retrieve the mean and spatial pattern of surface fluxes. Here, we describe the key principles and steps of the ERF approach and refer to Metzger et al. (2013) and Xu et al. (2017a, b) for more detailed descriptions. The underlying principle of the ERF approach is to relate responses to drivers, e.g. to relate sensible heat fluxes to potential temperature and land-surface temperature. The essential steps are: (1) identifying and measuring flux responses alongside key environmental drivers including meteorological forcings and surface ecological properties; (2) inferring process relationships between environmental responses and drivers using artificial intelligence bound by first principles; and (3) projecting flux responses in space and time.

Flux responses are calculated from wavelet-decomposed eddy covariance. The traditional eddy-covariance method requires averaging turbulent fluctuations over a 30-min or 60-min window for tall towers, making it susceptible to changing turbulent conditions and footprint variations that occur on shorter time scales. By using a wavelet-based spectral average instead, the flux measurement period can be reduced substantially (5-min window size at 1-min resolution), permitting clear spatial attribution without neglecting long-wavelength flux contributions (Metzger et al. 2013; Xu et al. 2017a). The drivers in the ERF approach are tower-measured air temperature and moisture, and footprint-weighted surface temperature and moisture. Footprint modelling (Kljun et al. 2004; Metzger et al. 2013) outputs reflect that towers located within warm–dry and cold–wet patches as well as at their boundaries, at various times measure both the warm–dry and the cold–wet patches in their flux footprints, therefore capturing the surface variability (Fig. 1c). Machine learning is then applied to extract the environmental response functions between flux responses and environmental drivers (Elith et al. 2008). The extracted relationships in conjunction with surface temperature and moisture maps are then utilized for spatio-temporal projection of the fluxes to the whole domain.

We calculated fluxes using the global wavelet covariance over the whole LES analysis time between hour 2 and hour 5. This includes energy contributions from the whole 3-h period. Instead of using a rectangular wavelet cut-off, which can exclude the artifacts due to edge effects (Misztal et al. 2014; Vaughan et al. 2016), we chose global wavelet covariance in order to conserve the covariance as calculated from time-domain covariance. Wolfe et al. (2018) found that for short time periods, neglecting cospectral power within the edge effect could create systematic errors due to the exclusion of larger-scale flux contributions. This was also observed in our analysis with the 3-h dataset whereby the rectangular cut-off underestimated fluxes by about 10–20% compared to time-domain eddy covariance, while global wavelet covariance was only biased by 6% and 5% for H and LE, respectively. This probably will not be a problem in 'real-world' continuous tower measurements, since real towers can ingest longer time series (e.g. five hours or longer) and thus allow targeting the diurnal or synoptic spectral gap when considering the appropriate analysis time period.

The sub-hourly wavelet-decomposed flux had a larger sample size and higher signal-to-noise ratio compared to standard eddy covariance, which enables machine learning to extract relationships between these atmospheric flux 'responses', and land-surface and meteorological 'drivers' as the footprint varies from minute to minute. Wavelet decomposition can produce 180 fluxes at 1-min resolution within a 3-h analysis time for each virtual tower. With n virtual tower(s), where n ranges from 1 to 14, the total training dataset for machine learning is $180n$.

For physics-guided machine learning we selected drivers in line with flux-gradient theory: to express observed sensible and latent heat fluxes, the learning algorithm was presented with driver pairs representing the temperature and moisture gradients between the surface and atmosphere, alongside information on vertical flux divergence. Specifically, potential temperature and water vapour mixing ratio measured by virtual towers were used as meteorological drivers. As surface drivers, near-land surface temperature and near-land surface humidity were used. They were calculated from the y-axis averaged potential temperature and mixing ratio at the lowest vertical grid level at 3.5 m. Near surface quantities were used as an alternative for the surface properties, since these were not defined in our LES set-up. This is a because we have not applied a land-surface model, which would have complicated the surface-flux pattern (Patton et al. 2005) and thus would have further complicated the analysis of systematic errors. Lastly, the learning algorithm was provided with the relative measurement height within the boundary layer (z_m/z_i) to express vertical flux divergence. Here, z_i is the boundary-layer height, calculated as the height of the minimum heat flux in the capping inversion layer.

For each number of towers ($n = \{1 \ldots 14\}$) 1000 ERF runs were made, each using a randomly-chosen subset of n towers from evenly- and unevenly-placed virtual towers to obtain a stable ensemble representative flux. The 1000 iterations were chosen for the ERF approach instead of 50,000 iterations as with other methods because of the longer ERF processing time and relatively stable ensemble representative flux obtained after 1000 times.

3 Results and Discussion

3.1 How Many Flux Towers are Needed to Sufficiently Sample the Flux Domain Mean?

Figure 2 shows how different approaches represent the domain-scale energy fluxes as a function of the number of towers. When using the spatial eddy-covariance approach, adding each

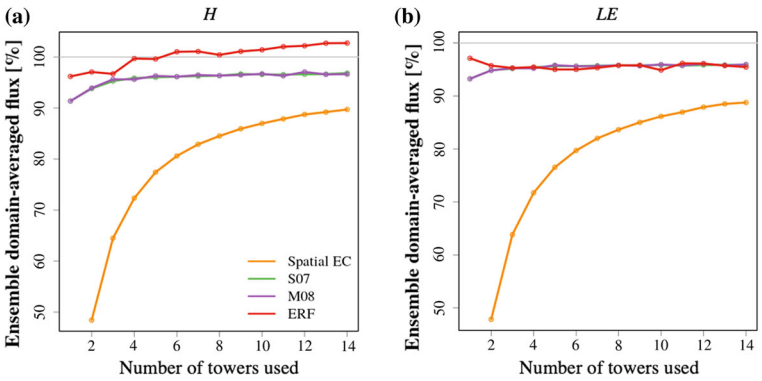

Fig. 2 Ensemble-averaged heat fluxes at 49 m using different numbers of 1–14 randomly chosen evenly-sampled virtual towers, for different upscaling approaches: spatial eddy covariance, spatio-temporal eddy covariance (S07 and M08), and environmental response function (ERF) for **a** sensible heat flux (H), and **b** latent heat flux (LE). Note that the fluxes shown here are the ensemble domain-averaged fluxes relative to the respective LES surface flux minus storage flux

tower helps the ensemble-averaged sensible heat flux (H) to approach the representative flux. However, a relatively high number of towers is needed to sufficiently retrieve the representative flux. When 14 towers are used per 12 km × 16 km, (one tower per 14 km^2), the resulting flux only reaches 89% of the representative flux, calculated as the prescribed surface flux minus a storage flux of 5 W m^{-2}. This indicates that the spatial eddy-covariance approach is still unable to capture and explain the remaining 11% of the total flux $H + LE$ regardless of additional towers.

In comparison, the spatio-temporal eddy-covariance approach (S07 and M08 in Fig. 2) shows improved performance: H starts from 91.3% even with one single-tower ensemble average. This is probably a result of capturing more energy transport by mesoscale flows: these permeate not only the space domain, but also propagate in time, e.g. throughout each 30-min averaging period. In contrast, spatial eddy covariance only includes the coherence of energy transport across space for each instantaneous sampling time.

The spatio-temporal eddy-covariance representative flux increases with additional towers, reaching 94.8% of energy with fewer (five) towers (1 tower per 40 km^2), leaving only 5.2% of the sensible heat flux unexplained. The increased 3.5% ensemble domain mean H with multiple towers compared with one single tower indicates that the spatio-temporal eddy-covariance approach can account for a fraction of the mesoscale flow contributions (at long spatial scale but only up to 30-min temporal scale) in H. We also proved the M08 approach (Mauder et al. 2008) generates almost identical results with S07, while only requiring a single high-frequency tower and remaining towers measuring only averaged quantities, thereby reducing observational expense.

Similar results are seen in LE in Fig. 2b. The spatio-temporal eddy-covariance approach performs better than spatial eddy covariance for single- and multiple-tower estimates in terms of representing the domain-averaged latent heat flux. Ensemble-averaged LE derived from the spatio-temporal eddy-covariance approach increases from 94 to 96% of the prescribed surface flux as one goes from one to five towers.

Taking the storage flux into account, 5 W m^{-2} and 7 W m^{-2} for H and LE, respectively, the spatio-temporal eddy-covariance approach is able to reduce the missing flux contribution from 9 to 5% for H and from 6 to 4% for LE, with only five towers compared to traditional turbulent flux calculation (spatio-temporal eddy covariance with one tower).

Table 1 The energy contribution of turbulent flux by using traditional eddy-covariance turbulent flux, the mesoscale flows (calculated as the difference between S07 and ERF domain-mean flux derived with one tower), and atmospheric skewing (calculated as the difference between ERF domain mean derived with one tower and with 14 towers) relative to LES surface forcing minus storage flux of sensible heat flux (H), latent heat flux (LE), and $H+LE$

	Turbulent flux (%)	Energy at long transporting scales (%)	Atmospheric skewing (%)	Sum (%)
H	91.3	4.9	6.5	102.7
LE	93.8	3.8	− 1.7	95.3
$H+LE$	92.7	4.2	1	97.8

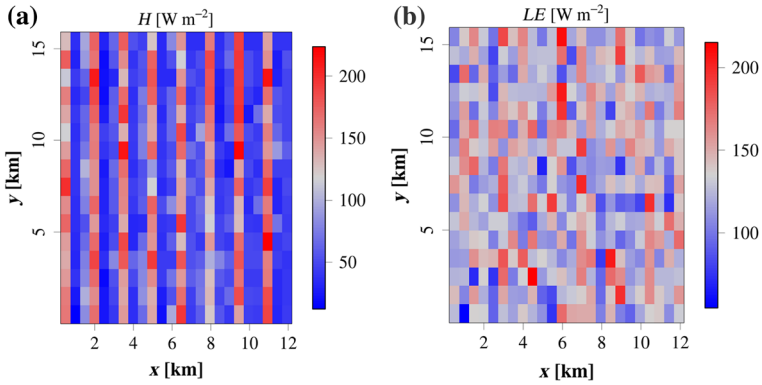

Fig. 3 24×16 virtual towers (one tower for each cell in Fig. 1) sensible heat and latent heat fluxes at 49 m calculated using standard eddy covariance during the analysis time, with one virtual tower shown as one cell. Note that each plot has a corresponding legend on the right side. Sensible heat flux map reflects distinct warm and cold patches, while the latent heat flux map is more mixed and homogenized without a strong wet–dry contrast

3.2 How Does the ERF Upscaling Method Help Flux Towers to Sample the Regional Domain?

Figure 2 shows how the environmental response function represents the regional domain with one and multiple towers. The ERF-projected H and LE values in Fig. 2a indicates that with only one tower, ERF represents 96% and 97% of the regional domain flux, separately, which is approximately what S07 and M07 reach with 14 towers. This is achieved by including the longer flux transporting scales (3-h) through wavelet-based flux computation in the environmental response function. Upscaling and wavelet-based flux computation improve the performance of energy flux measurements, even with a single tower.

Taking the difference between the resulting fluxes from the ERF approach with one tower and the resulting fluxes from spatio-temporal eddy covariance with one tower as a measure of the mesoscale flow contribution to the fluxes, we surmise that the mesoscale flow contributes to 5% and 4% of domain mean flux for H and LE, respectively (Table 1). Although 4% of energy in LE is potentially from longer transporting scales, the LE maps at virtual tower sites (Fig. 3b) and the humidity cross-section at the tower measurement height (Fig. 9b) do not show the pattern of LE mesoscale flows, as shown in the H field. The lack of any pattern

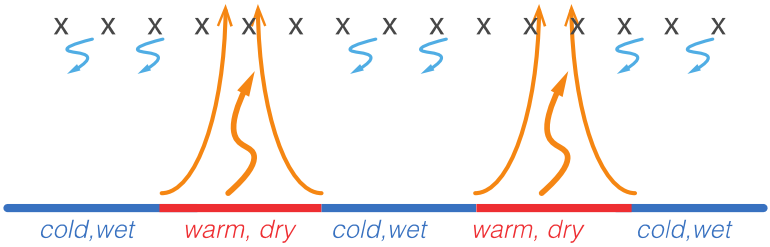

Fig. 4 Schematic illustration for convection-skewed atmospheric condition. In the updraft branch (orange lines) of the mesoscale flow, convective eddies gather more energy to overcome gravity, and become stronger and spatially-smaller as they ascend. The weaker downward flow (blue lines) occupies a larger horizontal area compared to the stronger upward flow so both compensate each other according to mass conservation. Because eddy-covariance towers (black crosses) are operating in this skewed atmospheric field, their measurements have a tendency to capture more of the weaker downward flow as opposed to the stronger upward flow

may indicate that the transport of *LE* is due to *H*-generated mesoscale flows, and would be consistent with previous studies (Charuchittipan et al. 2014; Eder et al. 2015).

With more towers the ERF-projected *H* increases and even slightly exceeds the reference flux, while the ERF-projected *LE* decreases (red line in Fig. 2b). This, we attribute to undersampling (oversampling) of warm–dry (cold–wet) patches by the virtual towers, which is described more below. Figure 5 shows probability density functions (p.d.f.) of the ERF-retrieved fluxes with different numbers of unevenly-distributed virtual towers (yellow to red lines in Fig. 5) and the p.d.f. of 320 evenly-distributed virtual tower measurements (green line in Fig. 5). It is striking that *H* projections demonstrate a bi-modal distribution, with a primary peak representing the cold–wet patch and a secondary peak representing the warm–dry partch. This unequal distribution of projected fluxes between the cold–wet and warm–dry patch indicates an undersampling (oversampling) of warm–dry (cold–wet) patches. As illustrated in Fig. 4, convective eddies gather more energy in the updraft branches, and become stronger and narrower as they ascend. Attributed to continuity, downdrafts are weaker and distributed over a larger area than the updrafts. The corresponding spatial distribution of atmospheric fields is skewed towards weaker, cooler subsidence areas (as opposed to stronger, warmer convective areas), even if cold–wet and warm–dry surface forcing is uniformly distributed in space (Fig. 4). Since the eddy-covariance tower is operating in this atmospheric field in a stationary frame of reference, its measurement is inherently "biased" to capture more of the weaker downward flow areas as opposed to the stronger upward flow areas (Fig. 4). The area bias is obviously seen in the p.d.f. of even-sampling virtual tower measurements (green line for *H* in Fig. 5), but the magnitude of upward flow is not obviously stronger than the surface forcing. This might be caused by stronger divergence in the upward branch, and overall is consistent with previous studies reflecting real world eddy-covariance towers that tend to undersample warm–dry spots (Wyngaard and Brost 1984; Foken 2008). With an increasing number of towers, the secondary peak of ERF-projected *H* values (yellow–red lines in Fig. 5a) shifts upward because, with multiple towers, the probability of fluxes originating from warm patches increases. This phenomenon allowed the environmental response function to rectify location bias from skewed atmospheric sampling by increasing the detection of warm–dry patches.

In contrast to *H*, ERF-retrieved *LE* values show a Gaussian-shaped distribution, even though surface fluxes were prescribed by a distinct bi-modal distribution. This indicates that the spatial pattern of *LE* at the 49-m measurement height is strongly blended and only weakly correlated to its surface pattern. Assuming that fluxes are fully sampled with 14 towers, and

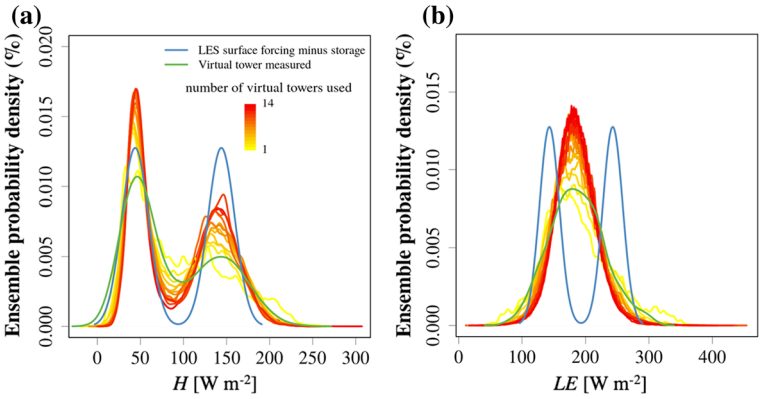

Fig. 5 Probability density functions of ensemble ERF-projected flux maps aggregated from 1000-fold ensemble for **a** H, and **b** LE. Yellow to red lines represent ERF-projected maps with 1–14 virtual towers, respectively. Blue lines are the LES surface forcing minus storage flux reference (5.3% and 3.4% for H and LE, respectively). Green lines are the 320 evenly-distributed virtual tower measured fluxes using standard eddy-covariance method

taking the difference in ERF-projected fluxes between one and 14 towers as a measure of the location bias incurred by skewed atmospheric sampling, we estimate the effect of location bias to be approximately 7% and − 2% for H and LE, respectively.

ERF-projected H and LE values with 14 towers approximately conserve the energy budget (101%) after adding the projected H and LE (98%) with the mean convection term (3%), which is calculated as $[\bar{\mathbf{w}}][\overline{\Theta}]$ (2.7%) and $[\bar{\mathbf{w}}][\bar{\mathbf{q}}]$ (0.04%).

Figure 6 shows the domain-averaged H and LE values for different number of towers, where the number of towers was randomly chosen from even as well as from uneven sampling. Here, even sampling implies that 1–14 towers were randomly chosen from 320 towers located in warm–dry and cold–wet patches evenly, while uneven sampling implies that 40% (60%) of the towers were located in warm–dry (cold–wet) areas. This shows how different upscaling approaches can reproduce the domain mean energy fluxes from even sampling (solid lines) and uneven (biased) sampling (dashed lines). This test was performed by comparing different sampling scenarios in order to estimate the sensitivity of the different approaches (spatio-temporal eddy covariance and the ERF method) on the location bias. For the ERF method, with an uneven sampling strategy the resulting H values deviate significantly from the reference flux for one tower, while an even sampling strategy is much closer to the reference flux, indicating that a location bias can have a large impact on the upscaled flux. For increasing number of towers the impact of location bias decreases by about 80% when 14 towers are used.

For S07 and M08, the even sampling strategies lead to significant improvements in the sampled H compared to the uneven sampling strategy, even for the large number of towers used. With respect to H, this suggests that S07 and M08 are much more sensitive to the chosen sampling strategy compared to ERF. For LE, however, the situation becomes different. Here, with an uneven sampling strategy ERF, S07 and M08 approaches give fluxes that are slightly closer to the reference flux compared to even sampling. At 'real-world' measurement sites with complex surface heterogeneity the locations of the updraft and downdraft branches of the mesoscale flows are hardly predictable and depend on wind speed and wind direction (Maronga and Raasch 2013), time of day, soil conditions, indicating a likely unable location

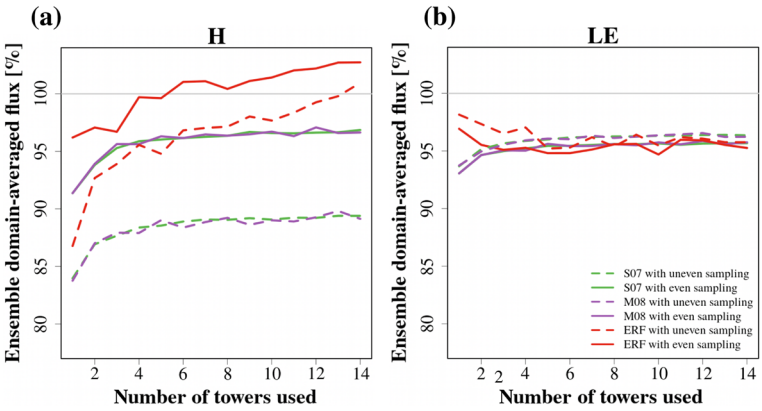

Fig. 6 Ensemble-averaged domain mean energy fluxes from randomly chosen 1–14 virtual towers from even sampling (solid lines) and uneven sampling (dashed lines) at 49 m. Even sampling indicates that 1–14 towers were randomly chosen from 320 towers so that each 50% of the towers were located in warm–dry and cold–wet patches, respectively. Uneven sampling indicates that 40% of the towers were located in warm–dry patches, while 60% of towers were chosen from cold–wet patches

bias, even with an elaborate tower set-up. Our findings show that any location bias in the tower set-up can introduce significant sampling errors if spatial and spatio-temporal eddy-covariance methods are used; whereas, the ERF approach is able to overcome most of the location bias.

3.3 How Accurately Can ERF Retrieve the Surface-Flux Variation?

After the previous ERF study comparing projected H against nearby tower observations (Xu et al. 2017a), we go one step further to assess the ability of ERF to retrieve the spatial pattern. Figure 7a shows the ERF-retrieved H with 14 towers projected onto the surface. The projected flux correlates well with the prescribed fluxes, revealing the ability of the ERF method to dis-aggregate measured fluxes into component fluxes appropriately. We found that with 14 towers ERF is able to reproduce the prescribed sensible-heat-flux pattern (Fig. 7a). Further, Fig. 7b shows the relative spatial coverage of the ERF-projected flux for different number of towers used. The ERF method can successfully rectify location bias by improving flux coverage from < 19% with one tower, to 33% with two towers, and to 74% with 14 towers (Fig. 7). Applying ERF over several months of "real-world" flux-tower data would build an ensemble of different atmospheric conditions resulting in a footprint climatology, which provides a similar improvement to the spatial representativeness, e.g. larger than 60% representativeness over a 10×10 km^2 area when conducted in a real-world test with a single tall tower (Xu et al. 2017a).

Figure 8 shows scatterplots of the ERF-projected fluxes versus the prescribed surface fluxes. For the sensible heat flux the fitted relationship (dashed line in Fig. 8) between ERF-projected H and the known LES surface fluxes is close to 1:1 line: the slope is 1.05, and offset is -5 W m^{-2} with $R^2 = 0.9$.

In contrast to H, the fitted relationship for LE deviates significantly from the 1:1 line, meaning that ERF is less successful in retrieving the two types of surface patches. ERF-projected LE maps were less successful in retrieving the two types of surface patches (Figs. 5b, 8b) compared to H. This is in contrast to previous studies that demonstrated successful

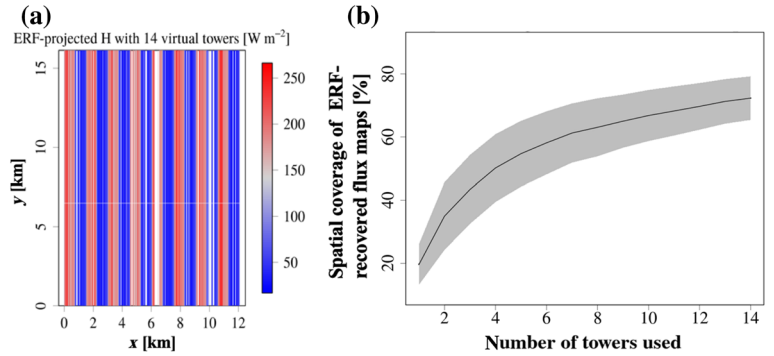

Fig. 7 **a** ERF-projected sensible heat fluxes using 14 towers. White areas are gaps that cannot be reproduced by ERF because their physical properties exceed the range of the observation dataset. The range of the ERF projecting limit was determined by the range of the driver inputs. If towers haven't seen the surface property of one area, e.g. the projected surface's land surface temperature is higher than any training data that towers have seen in their flux footprints, the ERF rejects projecting to this area (white area). **b** The percentage of projected area (i.e., colored space in Fig. 7a) of ERF-projected flux maps over the 12×16 km^2 target domain (solid line), and its double standard deviation (grey shaded area)

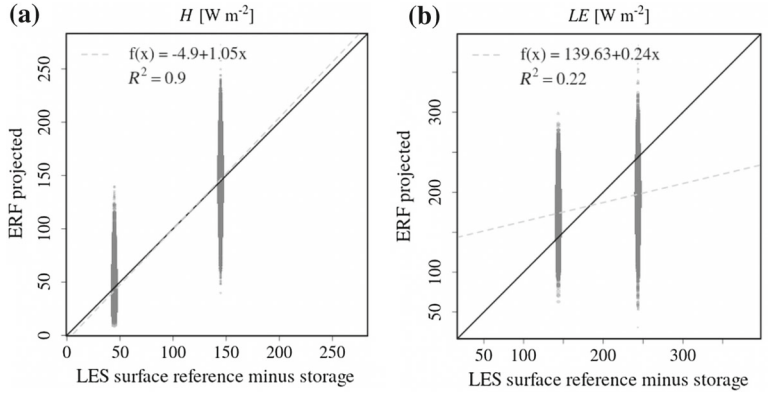

Fig. 8 Scatterplot of ensemble ERF-projected against the known LES surface reference fluxes minus the storage flux for **a** sensible heat flux and **b** latent heat flux, using 14 virtual tower observations. Solid black line is 1:1 line, and dashed black line is the least-squares regression

application of the ERF method on *LE* (Metzger et al. 2013; Xu et al. 2017a, b). We speculate that the "extreme" surface stripes result in different *LE* and *H* transport and circulations, which make it difficult for offline footprint model to accurately or sufficiently relate the turbulent-based *LE* back to the surface. This might relate to the fact that heterogeneity-induced mesoscale flows, which are dominated by sensible heat fluxes, mix the passive water vapour (Figs. 3b, 9b), making the atmospheric *LE* field more homogenized compared to *H* (compare Fig. 9a, b). While the potential temperature field is organized in rolls broadened along the vertical direction (Fig. 9a), the mostly passive water vapour is more horizontally mixed, and the rolls of water vapour are broadened horizontally. Thus, water vapour spreads laterally relative to the stripes of forcing (Fig. 9b). Although more investigations are needed, e.g. on comparing the relative importance of the horizontal mixing for q and θ, this argument is consistent with Bertoldi et al. (2012), who found, based on observations and LES, that

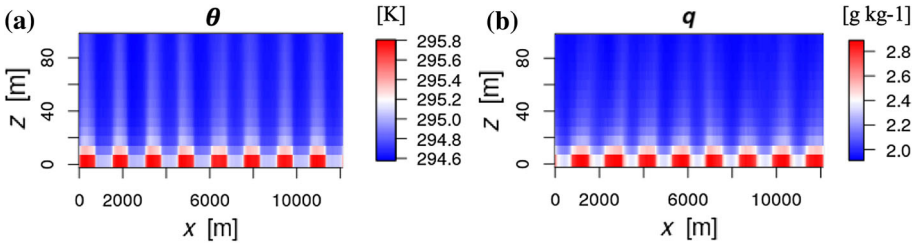

Fig. 9 *x*–*z* cross-section of **a** potential temperature and **b** water vapour in the study time period. The rolls of water vapour are broadened and more blended along the horizontal direction at elevations above 30 m, relative to potential temperature

the water vapour field is mostly controlled by horizontal advection. To further substantiate this hypothesis in follow-up research, we suggest applying Lagrangian particle footprint approaches, in order to improve understanding of how water vapour is mixed.

Another possible root cause for the less successfully ERF-projected *LE* values is that the offline footprint model used here have trouble relating the passive *LE* measured below the blending height back to the surface. The offline footprint models, such as used here, do not consider these scalar-specific properties in relation to the actual wind field, thus have insufficient spatial attribution of the passive (neutral density; *LE*) flux fields (Fig. 3b) to surface spatial patterning as opposed to active (self-buoyant; *H*) fluxes.

More analysis is needed to assess the ERF method's performance in other cases besides the idealized LES setting in this study, for example when the wind direction is perpendicular to the stripe pattern of the surface forcing as opposed to the paralleling of wind speed and stripe pattern in this LES setting. Here, the ERF method might perform better as the surface coverage and the variations in the footprint becomes better because of the larger footprint variation over heterogeneous surface, or worse because of more blending and mixing in the flux signal.

4 Summary and Conclusions

We used an ensemble of virtual towers in a large-eddy simulation (LES) over a prescribed striped surface forcing to investigate whether and how multiple towers with upscaling methods can sufficiently sample the mean and spatial variation of the regional energy fluxes in an unbiased manner. Our study demonstrates that the spatial eddy-covariance and the spatio-temporal eddy-covariance methods require one tower per 15 km^2 and one tower per 40 km^2 to capture 95% of the surface energy, while the environmental response function (ERF) upscaling approach is able to reduce the required tower density up to one order of magnitude to one tower per 200 km^2. ERF achieves this result by using Wavelet decomposition to incorporate low-frequency flux contributions (here: up to 3 hours) that are neglected by traditional eddy-covariance methods. In our striped-patched LES simulation, mesoscale flows contribute an average of 4% to the energy among virtual flux towers.

The ERF approach enables detecting and rectifying the location bias introduced by both location bias and atmospheric skewing, something which the spatial eddy-covariance and the spatio-temporal eddy-covariance approaches cannot. We showed that the sensible heat flux (*H*) resulting from spatial and spatio-temporal eddy-covariance methods is stronger affected by atmospheric skewness as well as location bias, meaning that the hot spots with strong but

narrow updrafts (which significantly contribute to the vertical transport) are underrepresented by the tower measurements. In contrast, the ERF method is able to rectify the location bias, making ERF-processed flux measurements less prone to skewed/location-biased tower setups.

In addition, the theoretical soundness and statistical significance of the ERF approach for H are supported by the good agreement of mean and spatial variation among ERF-projection and the known LES surface reference. Environmental response function enables rectifying location bias by improving flux coverage over the regional domain from < 1% tower footprint area, to 19%, 33%, and 74%, respectively, with one, two, and 14 virtual towers, respectively. Previous studies (Metzger et al. 2013; Xu et al. 2017a, b; Metzger 2017) have assessed ERF requirements as follows: (1) significant flux footprint variation due to varying wind speed and direction above the surface, (2) the accuracy of flux footprint modelling for different stability regimes and at different measurement heights, (3) physics-guided machine learning to represent important environmental processes in adherence to first principles, (4) the spatial and temporal resolution of drivers, and (5) the degree of the surface heterogeneity within the tower footprints. Together with previous studies, the results in this study support the suitability of the ERF approach for H under differing turbulent conditions and surface heterogeneity.

A possible root cause for the less successfully ERF-projected LE is confounded by differing atmospheric blending mechanisms for passive (neutral density; LE) scalars as opposed to active (self-buoyant; H) scalars. Simple footprint parametrizations, such as used here, do not consider these scalar-specific properties in relation to the actual wind field. As a result, spatial attribution of LE to surface properties is insufficient to enable the machine learning algorithm to distinguish areas of different surface forcings. In particular, two future research directions could help overcome this issue: (1) considering scalar-specific blending height to estimate a scalar-specific measurement height most suitable for existing algorithmic solutions, and (2) utilization of Lagrangian and scalar-specific footprint models for those measurement heights where existing algorithmic solutions are shown to be insufficient.

Based on this work, the National Science Foundation (NSF) supports the CHEESEHEAD (Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors) intensive field campaign which evaluates these approaches in a large flux-tower cluster (20 towers) over a 10×10 km^2 heterogeneous surface. The CHEESEHEAD experiment will comprehensively explore the performance of ERF, and work towards mitigating the constraints in the current LES-based study and previous single tower analyses.

We have demonstrated through a LES-based study that advanced scaling techniques can capture the effect of the mesoscale flows, map fluxes across a regional scale, and order-of-magnitude reduce the tower density required for capturing both the mean and the spatial variation of a heterogeneous surface. These improvements promise to decrease experimental expense, while providing eddy-covariance data products that are more suitable for model evaluation. This is an important validation step towards scaling the ERF approach to multiple towers and tower networks, and towards informing and designing future observation systems using simulations. Among other scientific advances our study underlines the potential of ERF for improving the quality assessment and uncertainty quantification of eddy-covariance flux data products, alongside model benchmarking and and data assimilation.

Acknowledgements K. Xu and A.R. Desai acknowledge support from Contract #3010-0401-000 from Battelle Ecology, Inc. to UW-Madison, DOE Ameriflux Network Management Project support to the ChEAS Core site cluster, and National Science Foundation AGS-1822420. The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by Battelle.

This material is based in part upon work supported by the National Science Foundation (Grant DBI-0752017). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. M. Sühring is funded by the German Federal Ministry of Education and Research (BMBF) (Grant 01LP1601A) within the framework of Research for Sustainable Development (FONA; www.fona.de). All simulations were performed on the Cray XC40 at The North-German Supercomputing Alliance (HLRN), Hannover/Berlin. All the input and output data are available in this link: http://co2.aos.wisc.edu/data/kxu/ERF-LES/. The permission of usage of these data is not required, but we do kindly ask you contact us of your intentions, most likely because we can provide help on data analysis. Our routines were developed in GNU R version 3.1 (R Development Core Team 2012), and code and examples are being developed for a public repository. Corresponding Docker compute images, R-packages and workflow examples are being developed for a public repository, and are available upon request in the meantime.

References

Baldocchi D (2008) 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust J Bot 56(1):1–26. https://doi.org/10.1071/BT07151

Baldocchi D, Falge E, Gu L, Olson R, Hollinger D, Running S, Anthoni P, Bernhofer C, Davis K, Evans R, Fuentes J (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(11):2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Bertoldi G, Kustas WP, Albertson JD (2012) Evaluating source area contributions from aircraft flux measurements over heterogeneous land using large-eddy simulation. Boundary-Layer Meteorol 147(2):261–279. https://doi.org/10.1007/s10546-012-9781-y

Bohrer G, Katul GG, Walko RL, Avissar R (2009) Exploring the effects of microscale structural heterogeneity of forest canopies using large-eddy simulations. Boundary-Layer Meteorol 132(3):351–382. https://doi.org/10.1007/s10546-009-9404-4

Bonan GB (2008) Ecological climatology: concepts and applications. Cambridge University Press, Cambridge

Charuchittipan D, Babel W, Mauder M, Leps J-P, Foken T (2014) Extension of the averaging time in eddy-covariance measurements and its effect on the energy balance closure. Boundary-Layer Meteorol 152:303–327. https://doi.org/10.1007/s10546-014-9922-6

De Roo F, Mauder M (2018) The influence of idealized surface heterogeneity on virtual turbulent flux measurements. Atmos Chem Phys 18(7):5059–5074. https://doi.org/10.5194/acp-18-5059-2018

Dietze MC, Vargas R, Richardson AD, Stoy PC, Barr AG, Anderson RS, Arain MA, Baker Ian T, Black TA, Chen JM, Ciais P, Flanagan LB, Gough CM, Grant RF, Hollinger D, Izaurralde RC, Kucharik CJ, Lafleur P, Liu S, Lokupitiya E, Luo Y, Munger JW, Peng C, Poulter B, Price DT, Ricciuto DM, Riley WJ, Sahoo A, Schaefer K, Suyker AE, Tian H, Tonitto C, Verbeeck H, Verma Shashi B, Wang W, Weng E (2011) Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American Carbon Program site-level synthesis. J Geophys Res Biogeosci. https://doi.org/10.1029/2011jg001661

Eder F, De Roo F, Rotenberg E, Yakir D, Schmid HP, Mauder M (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

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77(4):802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Finnigan JJ, Clement R, Malhi Y, Leuning R, Cleugh HA (2003) A re-evaluation of long-term flux measurement techniques—part I: averaging and coordinate rotation. Boundary-Layer Meteorol 107(1):1–48. https://doi.org/10.1023/a:1021554900225

Foken T (2008) The energy balance closure problem: an overview. Ecol Appl 18(6):1351–1367. https://doi.org/10.1890/06-0922.1

Foken T, Aubinet M, Finnigan JJ, Leclerc MY, Mauder M, Paw UKT (2011) Results of a panel discussion about the energy balance closure correction for trace gases. Bull Am Meteorol Soc 92(4):ES13–ES18. https://doi.org/10.1175/2011bams3130.1

Kanda M, Inagaki A, Letzel MO, Raasch S, Watanabe T (2004) LES study of the energy imbalance problem with Eddy covariance fluxes. Boundary-Layer Meteorol 110(3):381–404. https://doi.org/10.1023/B:BOUN.0000007225.45548.7a

Kenny WT, Bohrer G, Morin TH, Vogel CS, Matheny AM, Desai AR (2017) A numerical case study of the implications of secondary circulations to the interpretation of eddy-covariance measurements over small lakes. Boundary-Layer Meteorol 165(2):311–332. https://doi.org/10.1007/s10546-017-0268-8

Kljun N, Calanca P, Rotach M, Schmid H (2004) A simple parameterisation for flux footprint predictions. Boundary-Layer Meteorol 112(3):503–523. https://doi.org/10.1023/B:BOUN.0000030653.71031.96

Maronga B, Raasch S (2013) Large-eddy simulations of surface heterogeneity effects on the convective boundary layer during the LITFASS-2003 Experiment. Boundary-Layer Meteorol 146(1):17–44. https://doi.org/10.1007/s10546-012-9748-z

Maronga B, Gryschka M, Heinze R, Hoffmann F, Kanani-Sühring F, Keck M, Ketelsen K, Letzel MO, Sühring, M, Raasch S (2015) The parallelized large-eddy simulation model (PALM) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. Geosci Model Dev 8:2515–2551. https://doi.org/10.5194/gmd-8-2515-2015

Mauder M, Desjardins RL, Pattey E, Gao Z, van Haarlem R (2008) Measurement of the sensible eddy heat flux based on spatial averaging of continuous ground-based observations. Boundary-Layer Meteorol 128(1):151–172. https://doi.org/10.1007/s10546-008-9279-9

Metzger S (2017) Surface-atmosphere exchange in a box: Making the control volume a suitable representation for in situ observations. Agric For Meteorol (Massman special issue)

Metzger S, Junkermann W, Mauder M, Butterbach-Bahl K, Widemann BTY, Neidl F, Foken T (2013) Spatially explicit regionalization of airborne flux measurements using environmental response functions. Biogeosciences 10(4):2193–2217. https://doi.org/10.5194/bg-10-2193-2013

Metzger S, Durden D, Sturtevant C, Luo H, Pingintha-Durden N, Sachs T et al (2017) eddy4R 0.2.0 a DevOps model for community-extensible processing and analysis of eddy-covariance data based on R, Git, Docker, and HDF5. Geosci Model Dev 10(9):3189–3206. https://doi.org/10.5194/gmd-10-3189-2017

Misztal P, Karl T, Weber R, Jonsson H, Guenther AB, Goldstein AH (2014) Airborne flux measurements of biogenic isoprene over California. Atmos Chem Phys 14:10631–10647

Moeng C-H, Wyngaard JC (1984) Statistics of conservative scalars in the convective boundary layer. J Atmos Sci 41(21):3161–3169. https://doi.org/10.1175/1520-0469(1984)041%3c3161:socsit%3e2.0.co;2

Novick KA, Biederman JA, Desai AR, Litvak ME, Moore DJP, Scott RL, Torn MS (2018) The AmeriFlux network: a coalition of the willing. Agric For Meteorol 249(Supplement C):444–456. https://doi.org/10.1016/j.agrformet.2017.10.009

Patton EG, Sullivan PP, Moeng C-H (2005) The influence of idealized heterogeneity on wet and dry planetary boundary layers coupled to the land surface. J Atmos Sci 62(7):2078–2097

R Development Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Raasch S, Schröter M (2001) PALM–a large-eddy simulation model performing on massively parallel computers. Meteorol Z 10:363–372

Richardson AD, Anderson RS, Arain MA, Barr AG, Bohrer G, Chen G, Chen JM, Ciais P, Davis KJ, Desai AR, Dietze MC (2012) Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. Glob Change Biol 18(2):566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Running SW, Baldocchi DD, Turner DP, Gower ST, Bakwin PS, Hubbard KA (1999) A global terrestrial monitoring ntework integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sens Environ 70:108–127

Sakai RK, Fitzjarrald DR, Moore KE (2001) Importance of low-frequency contributions to eddy fluxes observed over rough surfaces. J Appl Meteorol 40(12):2178–2192. https://doi.org/10.1175/1520-0450(2001)040%3c2178:iolfct%3e2.0.co;2

Schaefer K, Schwalm CR, Williams C, Arain MA, Barr A, Chen JM, Davis Kenneth J, Dimitrov D, Hilton TW, Hollinger DY (2012) A model-data comparison of gross primary productivity: results from the North American Carbon Program site synthesis. J Geophys Res Biogeosci. https://doi.org/10.1029/2012jg001960

Schlegel F, Stiller J, Bienert A, Maas H-G, Queck R, Bernhofer C (2014) Large-eddy simulation study of the effects on flow of a heterogeneous forest at sub-tree resolution. Boundary-Layer Meteorol 154(1):27–56. https://doi.org/10.1007/s10546-014-9962-y

Schwalm CR, Williams CA, Schaefer K, Anderson R, Arain MA, Baker I, Barr A, Black TA, Chen G, Chen JM (2010) A model-data intercomparison of CO_2 exchange across North America: results from the North American Carbon Program site synthesis. J Geophys Res Biogeosci. https://doi.org/10.1029/2009jg001229

Steinfeld G, Letzel MO, Raasch S, Kanda M, Inagaki A (2007) Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study. Boundary-Layer Meteorol 123(1):77–98

Stoy PC, Mauder M, Foken T, Marcolla B, Boegh E, Ibrom A, Arain MA, Arneth A, Aurela M, Bernhofer C (2013) A data-driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity. Agric For Meteorol 171:137–152. https://doi.org/10.1016/j.agrformet.2012.11.004

Sühring M, Metzger S, Xu K, Durden D, Desai AR (2019) Tradeoffs in flux disaggregation: a large-eddy simulation study. Boundary-Layer Meteorol 170(1):69–93. https://doi.org/10.1007/s10546-018-0387-x

Vaughan AR, Lee JD, Misztal PK, Metzger S, Shaw MD, Lewis AC, Purvis RM, Carslaw DC, Goldstein AH, Hewitt CN, Davison B (2016) Spatially resolved flux measurements of NO_X from London suggest significantly higher emissions than predicted by inventories. Faraday Discuss 189:455–472

Wolfe GM, Kawa SR, Hanisco TF, Hannun RA, Newman PA, Swanson A, Bailey S, Barrick J, Thornhill KL, Diskin G, DiGangi J (2018) The NASA Carbon Airborne Flux Experiment (CARAFE): instrumentation and methodology. Atmos Meas Tech 11:1757–1776. https://doi.org/10.5194/amt-11-1757-2018

Wyngaard JC, Brost RA (1984) Top-down and bottom-up diffusion of a scalar in the convective boundary layer. J Atmos Sci 41(1):102–112. https://doi.org/10.1175/1520-0469(1984)041%3c0102:tdabud%3e2.0.co;2

Xu K, Metzger S, Desai AR (2017a) Upscaling tower-observed turbulent exchange at fine spatio-temporalresolution using environmental response functions. Agric For Meteorol 232:10–22. https://doi.org/10.1016/j.agrformet.2016.07.019

Xu K, Metzger S, Desai AR (2017b) Surface-atmosphere exchange in a box: space-time resolved storage and net vertical fluxes from tower-based eddy covariance. Agric For Meteorol. https://doi.org/10.1016/j.agrformet.2017.10.011

Affiliations

Ke Xu[2,3] · Matthias Sühring[4] · Stefan Metzger[1,2] · David Durden[1] · Ankur R. Desai[2]

Stefan Metzger
smetzger@battelleecology.org
https://w3id.org/saegroup

1 National Ecological Observatory Network Program, Battelle, 1685 38th Street, Boulder, CO 80301, USA

2 Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, 1225 W Dayton St, AOSS 1549, Madison, WI 53706, USA

3 Climate and Space Sciences and Engineering, University of Michigan-Ann Arbor, 2455 Hayward Street, Ann Arbor, MI 48109, USA

4 Institut für Meteorologie und Klimatologie, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany