Agricultural and Forest Meteorology 255 (2018) 81–91

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Surface-atmosphere exchange in a box: Space-time resolved storage and net vertical fluxes from tower-based eddy covariance

Ke Xu[a,*], Stefan Metzger[a,b], Ankur R. Desai[a]

[a] University of Wisconsin-Madison, Department of Atmospheric and Oceanic Sciences, Madison, WI, USA
[b] National Ecological Observatory Network, Battelle, 1685 38th Street, Boulder, CO 80301, USA

ARTICLE INFO

Keywords:
Eddy-covariance
Mass continuity equation
Storage flux
Location bias
Energy balance closure
Environmental response function

ABSTRACT

Systematic bias in eddy-covariance flux measurements are pervasive. These arise both from unmeasured terms such as advection, and sampling bias in representativeness of the footprint for both turbulent and storage fluxes. As a result, the majority of eddy-covariance towers suffer from unaccounted bias when comparing to gridded earth system models and fail to close the surface energy balance. We hypothesize that one cause for these two problems is a mismatch between mass and energy fluxes measured within a time-varying source area and the actual storage and net vertical flux over a presumed "control volume", a novel concept derived theoretically in Metzger (this issue). Here, we practically implement this theory to estimate the true net surface-atmosphere exchange (NSAE) over such control volume, thus resolving "storage flux" and "vertical advection" issues by applying the environmental response function (ERF) technique to a virtual control volume (VCV). In this method, flux observations are related at high spatio-temporal resolution to meteorological forcings and surface properties within the estimated flux footprint, and these relationships are utilized to map the control volume explicitly in 3-D over space and time. Volume integration then allows, for the first time, retrieval of the NSAE. When ERF was applied to eddy covariance and profile observations in July and August 2014 from the AmeriFlux Park Falls WLEF tower in Wisconsin, USA, heat emission integrated over the target domain increased substantially over the tower observations by $+18.2\ \mathrm{Wm^{-2}}$ (+20.6%). Storage flux contributes up to 30% of NSAE at hourly timescale. The systematic uncertainty of ERF-VCV method applied for vertical flux and storage flux is within 15% and 20%, respectively. This systematic uncertainty is effectively corrected in projections. Volume-controlled NSAE provides improvements for mapping unbiased surface-atmosphere exchange for model-data comparison, assimilation and model building at model grid scale. These advances also present a promising direction for reconciling energy balance non-closure.

1. Introduction

The eddy-covariance technique has been used worldwide to monitor the surface atmosphere exchange for decades (Baldocchi et al., 2001; Bonan et al., 2011). Theoretically, it provides reliable observations for comparison with remote sensing data and earth system models (ESMs) to constrain model performance and to improve our understanding of the impact of global climate change (Dietze et al., 2014; Richardson et al., 2012; Schaefer et al., 2012). However, despite its widespread use, the eddy covariance technique faces a number of challenges especially when attempting to monitor surface-atmosphere exchanges in conditions that deviate from assumptions used to derive the method.

Among those challenges, location bias (Desai et al., 2008; Chen et al., 2011; Schmid, 1997; Stoy et al., 2013) and energy imbalance (Leuning et al., 2012; Foken et al., 2011; Foken, 2008) are two major ones. Location bias refers to the surface influencing the eddy covariance flux varying temporally and usually within the areas of 10^{-1}–10^{1} $\mathrm{km^2}$. This area is far smaller than typical ecosystem scales and the resolution of most earth system models (10^{2}–10^{4} $\mathrm{km^2}$), leading to a mismatch of scale and potentially representativeness in model-data comparison. Furthermore, almost all flux towers suffer from an imbalanced energy budget: the sum of measured sensible and latent heat flux is consistently less than available net radiation minus ground heat flux, which may also indicate bias in other trace gas fluxes such as CO_2 (Foken, 2008).

Here, we argue that one root cause for these challenges are the inherent assumptions of the eddy covariance technique allowing the measured vertical turbulent flux (Fig. 1b) to represent the net surface-atmosphere exchange (NSAE) across a volume (Fig. 1a). This volume is

DOI of original article: http://dx.doi.org/10.1016/j.agrformet.2017.08.037
* Corresponding author at: Dept of Atmospheric and Oceanic Sciences, UW-Madison, 1225 W Dayton St, Madison, WI 53706 USA.
E-mail address: ke.xu@wisc.edu (K. Xu).

http://dx.doi.org/10.1016/j.agrformet.2017.10.011
Received 26 December 2016; Received in revised form 14 September 2017; Accepted 9 October 2017
Available online 16 October 2017
0168-1923/

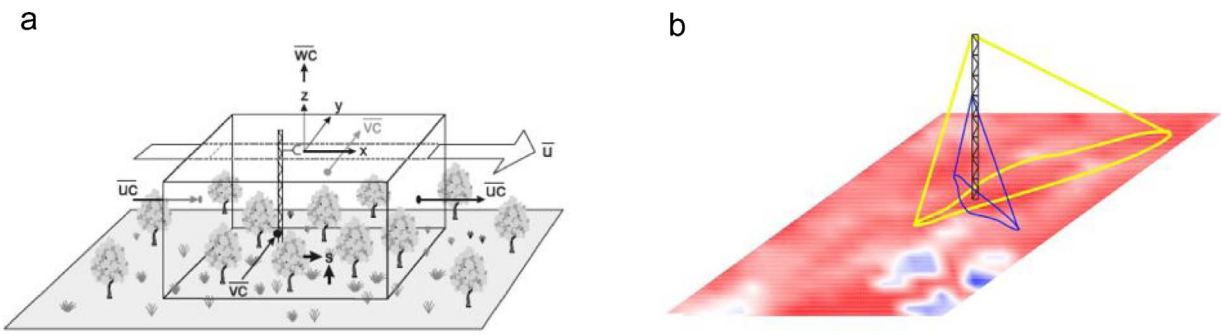

Fig. 1. Conceptual plot of eddy covariance measurements in ideal condition (a) after Finnigan (2004) and in reality (b). Ideally, tower measurements can represent the net surface atmosphere exchange across the virtual control volume around it (a). However, in reality tower measurements can only represent a cone-shape that is confined by the footprint area over a heterogeneous surface.

confined vertically from the surface to the measurement height, across a horizontal target domain that is large in comparison to the flux footprint. The vertical turbulent flux measured at the top of the tower is assumed to be spatially representative across the entire target domain. However, in reality, the corresponding flux measurement footprint typically represents a small fraction (order 1%) of the target ecosystem or ESM grid cell.

Further, the location of this fraction changes with time. Many sites have differing frequency of wind directions and magnitudes at daytime versus nighttime (Xu et al., 2017). This results in vertical turbulent flux observations at daily scales being a convolution of the diurnal cycle with a related systematic change in flux footprint. In non-homogenous conditions, which are common, this temporally varying and small flux footprints directly leads to location bias in eddy covariance measurements (Xiao et al., 2014; Desai et al., 2015).

Second, long-term storage flux is often assumed to be either zero, filtered out using steady state criteria, or vertical profile measurements are assumed to be horizontally representative of the storage flux across the virtual box as a whole. However, in reality, the storage flux is not necessarily negligible for typical eddy flux averaging intervals, in particular for tall towers over complex surfaces. Lack of consideration of this term has been shown to contribute to the observed energy imbalance (Leuning et al., 2012). Moreover, each vertical profile observation used to calculate storage flux has its own footprint, resulting in an influence area that increases with measurement height. This can be another source of location bias in flux tower observations (Schmid 1997; Raupach et al., 1988).

Third, advection terms are typically assumed to cancel each other or to be smaller than the turbulent flux. However, local advection can contain spurious low-frequency mesoscale contributions to the NSAE, which has been argued as another source for energy imbalance (Finnigan, 2008; Finnigan et al., 2003; Kanda et al., 2004; Foken 2008; Eder et al., 2015; Sakai et al., 2001). Low-frequency flux contributions can be caused by turbulent organized structures (TOS; Kanda et al., 2004; Finnigan et al., 2003) and/or secondary circulations associated with surface heterogeneity (Schlegel et al., 2014; Eder et al., 2015). In the presence of these structures, strong convection of warmer/wetter air occurs in spatially confined updraft zones, countered by a slight subsidence of cooler/dryer air across the majority of the target area. The corresponding timescales exceed typical eddy covariance averaging periods, thus these structures manifest themselves as advection rather than as turbulent contribution to the net surface-atmosphere exchange. In addition, this flux contribution can appear as a component of horizontal advection, vertical advection and storage, or combinations thereof, depending on the location of updraft and downdraft zones with respect to the measurement location. Due to the skewed distribution of updraft and downdraft zones, heat fluxes will thus tend to be underestimated when only considering the turbulent vertical flux in the standard eddy covariance approach (Mauder et al., 2008).

To mitigate the first assumption (representativeness) on vertical turbulent flux, two main upscaling approaches, process-based and data-driven approaches, have been utilized. However, each is subject to specific limitations. Purely process-based scaling (Wang et al., 2006; Desai et al., 2008; Xiao et al., 2011) relies on prescribed mechanistic relationships, oftentimes based on laboratory calibrations and far-reaching assumptions. Purely data-driven scaling (Xiao et al., 2014, 2008; Hutjes et al., 2010) minimizes the number of assumptions employed by inferring relationships among observations directly from the available data, but are limited in model robustness and predictive performance. More recently, Metzger et al. (2013) and Xu et al. (2017) developed the environmental response function (ERF) approach that combines the process-based and data-driven approaches. The underlying principle of ERF is to relate high-frequency (minute-to-minute) fluxes over fast-varying footprints with appropriate spatial or temporal drivers, e.g. meteorological and surface ecological forcings. The extracted relationships are then used for spatio-temporal mapping over a large domain.

To address the second and third assumption, Metzger (this issue) developed ERF further, and derived the ERF virtual control volume (ERF-VCV) framework. In essence, ERF-VCV attempts to apply the ERF technique to all terms in the continuity equation, including storage flux, advection and turbulent flux, to estimate the control volume net surface atmosphere exchange (NSAE). The two main assumptions of ERF-VCV theory are: i) the eddy covariance observations are dominated by surface flux, not entrainment, and ii) all relevant atmospheric and surface dynamics/state-space combinations are sampled by the eddy covariance platform.

Here we ask, based on an example case using one single tower during July and August 2014 over a heterogeneous environment of AmeriFlux Park Falls WLEF very tall tower in North Wisconsin, USA:

1. Can assumptions made in ERF-VCV theory be fulfilled in a real-world setting and enable the method to retrieve VCV-estimated storage and vertical heat fluxes?
2. How do ERF-VCV produced heat fluxes compare in magnitude and pattern to the tower-measured turbulent heat fluxes and what does it imply for location bias?
3. How does ERF-VCV enable addressing advective heat fluxes?
4. Do the virtual control volume integrated heat fluxes provide insight into how energy balance closure can be further addressed?

To answer these questions, we first introduce the climate, biophysical properties in the study area and footprint composition of the WLEF tower data (Sect. 2). The methodology of ERF-VCV and associated uncertainty algorithms are described in Sect. 2.2. We present the extracted relationships, scaled storage and vertical flux grids and the associated uncertainty budget in Sect. 3. The previous four key questions are discussed in Sect. 4, and conclusions are provided in Sect. 5.

2. Materials and methods

2.1. Study area and data acquisition

The 447-m tall WLEF television tower (45.9°N, 90.3°W) is located in the Mississippi River Basin, within the Park Falls Ranger District of the Chequamegon-Nicolet National Forest, Wisconsin, USA (Xu et al., 2017). The surrounding landscape is a mix of wetlands and upland forests. The tower footprint climatology samples a landscape that is representative of much of the Upper Midwest U.S. forested region (Desai et al., 2008, 2015). The surrounding forest canopy has approximately 70% deciduous and 30% coniferous trees, and a mean canopy height of 20 m. The whole region was heavily logged around the beginning of the 20th century. Soils are sandy loam and are mostly glacial outwash deposits. The site has an interior continental climate with cold winters and warm summers.

Observations used in this study include tower-measured meteorological variables, storage, turbulent, and vertical advection fluxes, as well as remote sensing products. Tower-based 10 Hz observations were chosen from 6th July to 31st August 2014 for WLEF at both 30 m and 122 m levels. The flux footprint along-wind distance was 0.1–0.8 km and 2–5 km for 30 and 122 m height measurement at 90% cumulative level (Fig. 2), respectively. For 30 m and 122 m levels, fast response wind speed and air temperature were derived from a sonic anemometer (Applied Technogies., Inc. Seattle, USA, ATI Type K). Dry mole fraction of water vapor were measured by a closed-path infrared gas analyzer (LI-COR, Inc. Lincoln, USA, LI-6262) at both levels. Reference air temperature and relative humidity were also measured (Vaisala, Inc. Louisville, USA, HMP45C). Additional measurements at the surface included the barometric air pressure (Vaisala, Inc. Louisville, USA, PT101B).

Land surface temperature (LST) and enhanced vegetation index (EVI) were chosen as biophysical surface drivers. These two drivers were acquired from Moderate Resolution Imaging Spectroradiometer (MODIS) data products. 250 m 16-day interval MOD13Q1 (V005) EVI and 1000 m 8-day daytime MYD11A2 (V005) LST. Atmospheric boundary layer (ABL) height, z_i, was obtained by linear interpolation into one-minute interval from the North American Regional Reanalysis (NARR) 3-hourly data produced by National Oceanic and Atmospheric Administration (NOAA).

2.2. Environmental response function − virtual control volume (ERF-VCV) approach

The ERF flux scaling procedure for tower eddy covariance measurements is based on Metzger et al. (2013) and Xu et al. (2017). The underlying principle of ERF is to relate high-frequency (minute-to-minute) fluxes via fast-varying footprint estimates with appropriate spatial or temporal drivers, e.g. biophysical surface and meteorological forcings. The extracted relationships are then utilized for spatio-temporal mapping over a domain that exceeds the typical footprint extent. The ERF method uses the footprint variation to re-assemble the NSAE. In order to project flux into one area, the exact same area doesn't necessarily have to be measured as long as its properties are within the state space of the training dataset used by the machine learning.

Unlike the work of Xu et al. (2017), where ERF was applied solely to vertical turbulent flux for a single month, here ERF flux scaling is realized not only for vertical turbulent flux ($\overline{w'c'}$), but also for low-frequency turbulent contributions to the vertical advection term ($\overline{wc}$) and storage flux ($\frac{\partial \bar{c}}{\partial t}$) observations over a longer-time period (2 months) to retrieve flux across the whole virtual control volume (VCV, Metzger, this issue). This procedure for the first-time permits addressing these terms in the mass continuity equation that are neglected by the standard eddy covariance methodology.

We first determined appropriate temporal scales for vertical transport, and calculated high-rate (one-minute) flux responses using wavelet discretization for tower-based vertical transport permitting inclusion of transporting scale up to three hours (Sect. 2.2.1). Next, the storage flux was determined using measurements at multiple vertical levels (Sect. 2.2.2). Lastly, ERFs were extracted using machine learning (Fig. 4) and used for projection for both vertical and storage flux (Fig. 5, Fig. 6) in Sect. 2.2.3. Our routines were developed in GNU R version 3.1 (R Development Core Team, 2014), and code and examples are being developed for a public repository (Metzger et al., 2017).

2.2.1. Wavelet discretized vertical exchange

Building on Metzger et al. (2013) and Xu et al. (2017), several preprocessing steps were performed: de-spiking after Brock (1986) (Table 1, Row 1) and Starkenburg et al. (2015), planar fit rotation (Wilczak et al., 2001), fixed lag correction using maximum correlation, and point-by-point conversion of sonic temperature to air temperature based on Schotanus et al. (1983). Only hours with more than 80% available raw data were used for further analysis (Table 1, Row 2).

Since flux footprint varies rapidly even within one hour (Xu et al., 2017), we use high-frequency (sub-hourly) flux responses to avoid footprint blending. At the same time, the spectral range of aggregated transporting scale from 10 Hz up to three hours permits combining high-frequency vertical turbulent flux and vertical advection (up to three hours) together into a single "vertical flux" term (Finnigan et al., 2003; Steinfeld et al., 2007; Mauder et al., 2008; Mahrt 1998). Wavelet decomposition (Torrence and Compo, 1998) can satisfy these two requirements to achieve high temporal resolution while including transporting scales up to several hours (Charuchittipan et al., 2014). However, wavelet decomposition assumes data are circulate, this

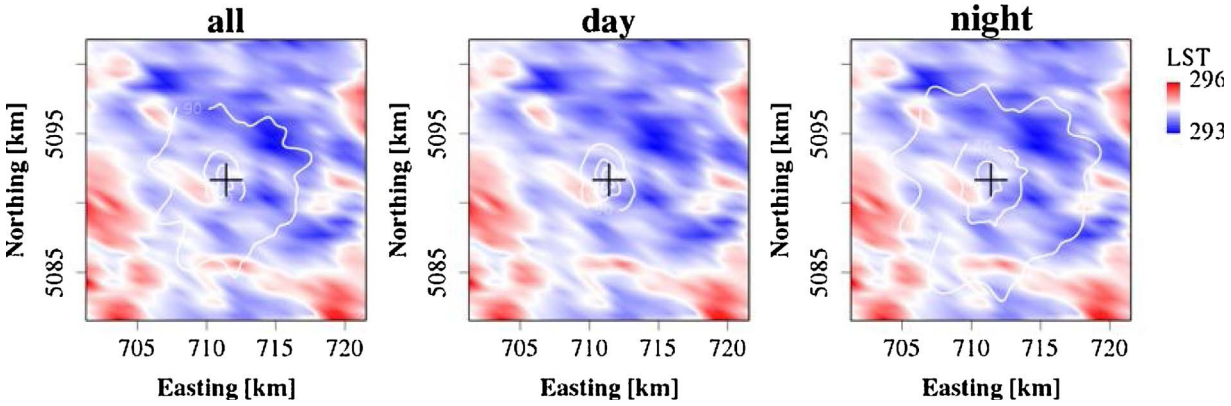

Fig. 2. Footprint climatology (30%, 60% and 90%, white contour lines) for 122 m level measurements superimposed over average MODIS land surface temperature (LST) within 20 × 20 km^2 target domain surrounding the tower, which is indicated with the central crosshairs. (a) entire study time period; (b) daytime (9:00–17:00 CST), and (c) nighttime (17:00–9:00 CST). In this study, the footprint had a diurnal cycle and LST over the footprint area was lower than over the whole domain.

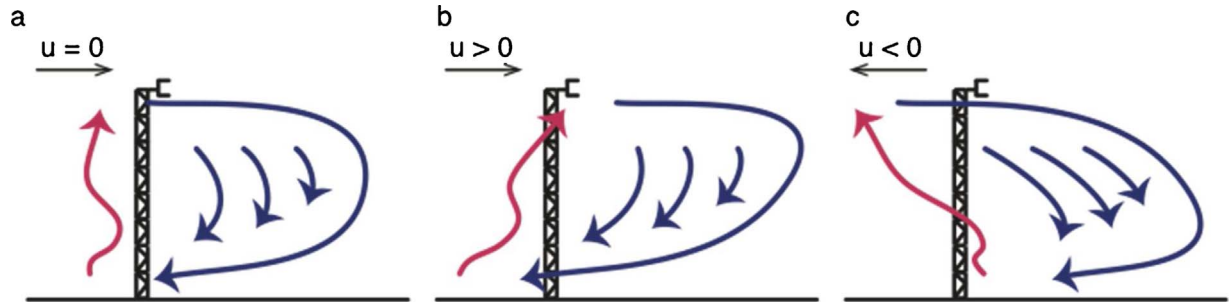

Fig. 3. Conceptual figure of a tower measuring (a) no mesoscale circulation, (b) the updraft branch and (c) the downdraft branch of mesoscale circulation. Red arrows are updrafts and blue arrows are downdrafts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

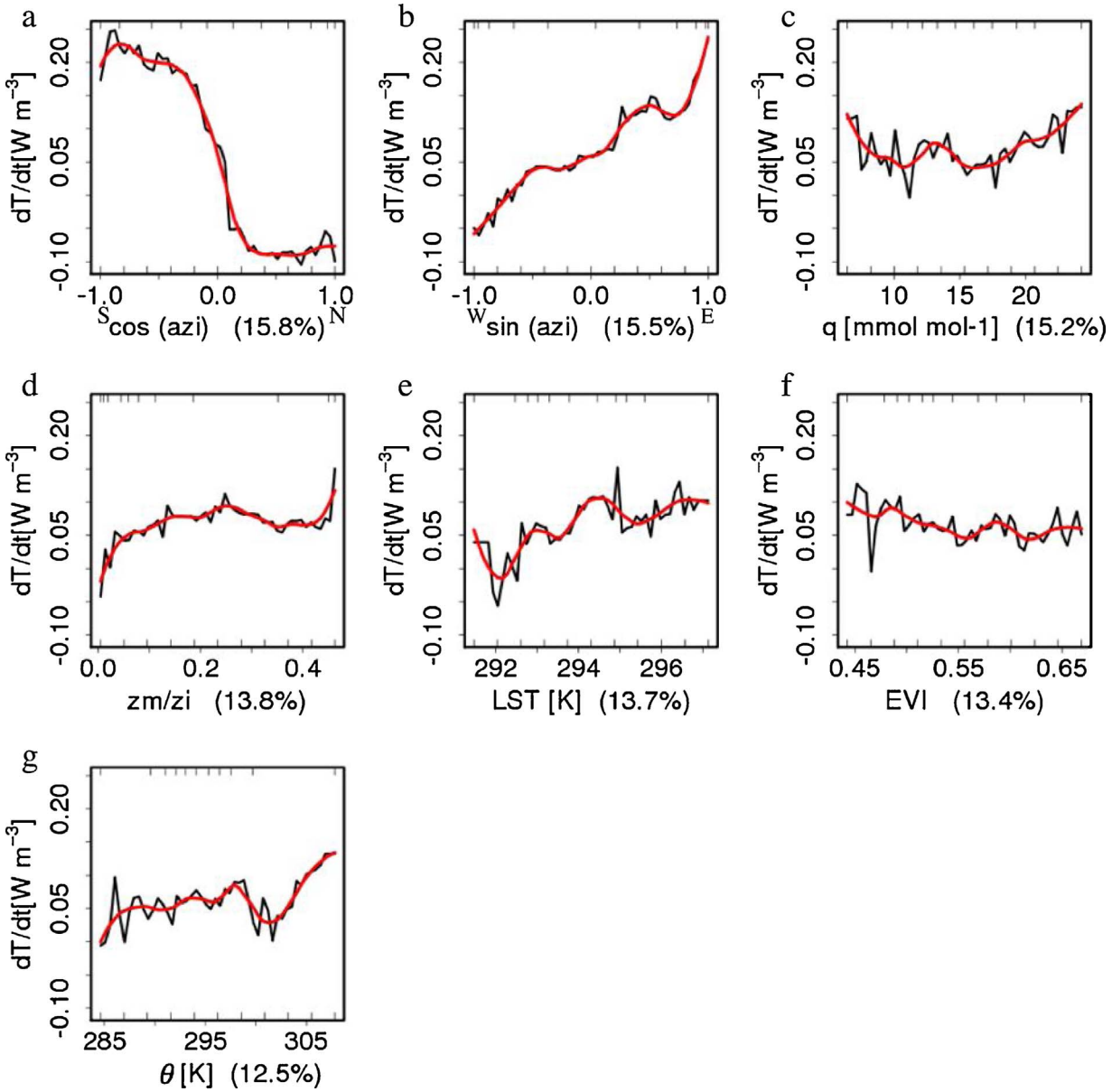

Fig. 4. Equidistant response plots of temperature (T) time-rate of change, from which heat storage is derived. In decreasing order of explained variation in the flux signal (partial R^2 in braces), the drivers are cosine of the azimuth angle (cos(azi)), sine of the azimuth angle (sin(azi)), water vapor mixing ratio (q), relative height within boundary layer (z_m/z_i), land surface temperature (LST), enhanced vegetation indix (EVI), and potential temperature (θ). The black lines are the fitted integrated response over the range of one individual driver. Smoothed representations of the fitted function (locally weighted polynomial regression) are in red bold lines. The equidistant response plots use uniformly distributed percentiles within the range of training data (inward tickmarks on the lower x-axis). Inward tickmarks on the upper abscissa represent training data percentiles. (W), (E) under sin(azi) subplot and (N), (S) under cos (azi) subplot indicate western, eastern, northern and southern direction of the solar azimuth, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

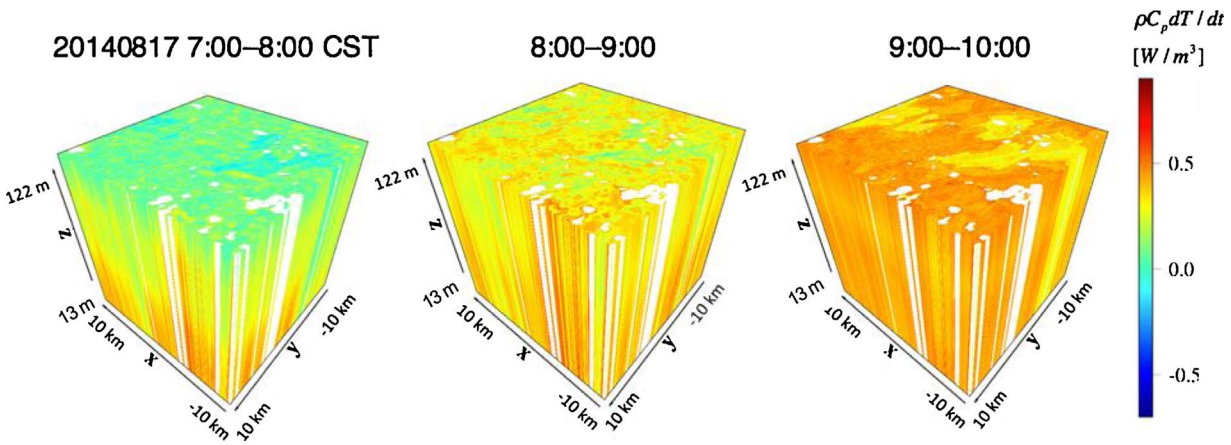

Fig. 5. Volume projection of heat storage flux from the displacement height to the 122 m measurement height, across the tower-centered 20 × 20 km^2 target region. The color is the temperature time-rate-of-change in Wm^{-3}. The volume projections show a developing convective boundary layer from 2014 Aug 17 7:00 a.m.–10:00 p.m. CST. White spaces are time-space locations that cannot be projected as they exceed the range of the training data.

assumption leads to larger uncertainty at the begin and end times of the dataset compared to its center (Torrence and Compo, 1998; Metzger et al., 2013). Therefore, the vertical flux over transporting scale of up to three hours were calculated using 15 h of raw data in order to avoid this edge effect. Here, we integrated over wavelet transport scales of up to three hours, with flux results discretized over a five-minute window that for each observation moves one-minute forward in time. Within the entire observation period, 55,160 one-minute vertical flux observations were obtained and qualified for machine learning.

2.2.2. Storage flux

We also determined the averaging time for storage flux. Following Finnigan (2006), we should avoid storage flux estimates influenced by single or a small number of eddies influencing the observations. Instead, we should use a period long enough to capture an adequate ensemble of these eddies. Here we consider the integral time scale of the turbulent time series between these eddies, and multiply it by one order of magnitude as the basis of storage flux computation. Considering the maximum integral space scale (98.4 m) and the average mean wind speeds of 2.2 m s^{-1} and 0.5 m s^{-1} at 122 m and 30 m, respectively, we determine 15 min as storage flux averaging time. For example, the storage flux estimates for timestamp 08:00:00 is calculated as the time average of measurements from 8:00:00 to 8:14:59 minus the average from 7:45:00 to 7:59:59. Within the entire observation period, 3304 storage flux observations are obtained and qualified for machine learning.

2.2.3. Environmental response function extraction and projection

The operator underlying the extraction of environmental relationships is the flux footprint model which links the flux responses to surface biophysical drivers. ERF builds the relationships among the observed fluxes, meteorology and footprint weighted surface properties to unveil the whole, time-varying flux field. In addition, by including transporting scales of three hours, one tower may sample updraft branch and downdraft branch in a mesoscale circulation, which principally enables ERF to reproduce large eddy circulations over the virtual box (Fig. 3).

Storage flux is the time-rate-of-change of a state variable such as temperature or mixing ratio of water vapor, as opposed to the state variable itself. Thus, we need to also consult a flux footprint model instead of concentration footprint model. Therefore, storage flux environmental relationships were also generated with observed storage exchange (response) and flux footprint-weighted biophysical surface properties and meteorological forcings (drivers). Flux footprint matrixes were calculated (Metzger et al., 2012; Kljun et al., 2004). The footprint model is valid for certain ranges of input parameters. Instead

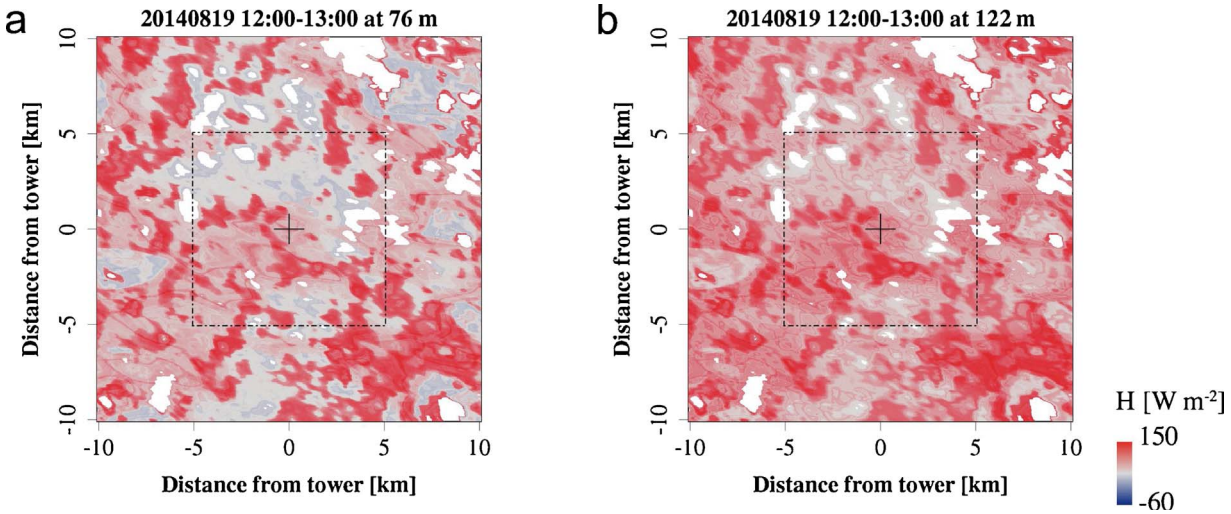

Fig. 6. Projected vertical flux grids of sensible heat for August 19th, 2014, 12:00-13:00 CST for the tower-centered 20 × 20 km^2 target region at (a) 76 m and (b) 122 m level. Subplot (a) may indicate suitable conditions for surface heterogeneity-induced mesoscale circulations. Subplot (b) shows the flux grids at a higher elevation with less spatially distinct features due to blending.

Table 1
Data filtering steps and the remaining percentage of data after each filtering step.

Processing step	Data filtering standard	Data remaining%
De-spiking	median filter de-spiking after Brock (1986)	99.5%
Data quality	3 h periods with missing data $\gg >$ 20% were discarded	60%
Unreliable nighttime data removal	Measurements at 122 m were discarded when sensible heat measured at 30 m $\ll -10$ W m^{-2}	50.4%
Spike removal	Lowest and highest 1% of flux values were removed	49.9%
Integral Turbulence Characteristics (ITC) test	Vertical velocity ITC $\ll$ 250% were selected	49.8%

of discarding footprints when the input parameters that exceed the allowable range, we instead set parameters that exceeds the input range to be minimum or maximum, i.e. roughness length less than $1e^{-5}$ is set to be $1e^{-5}$, and larger than 1 is set to be 1, vertical wind speed variation 0.23–1.23, and u* less than 0.2 is set to be 0.2, only for the purpose of footprint matrix calculation. Only for the purpose of footprint modeling, 11.7% of the data were constrained to match the valid range of footprint model inputs. Both storage and vertical fluxes are filtered for (i) unreliable nighttime data; (ii) spikes; and (iii) connectivity with the surface using (Table 1) the integral turbulence characteristic (ITC) test after Foken and Nappo (2008), as shown in Table 1, Row 3–5, separately. It should be mentioned that nighttime data considered as unreliable for the purpose of this study may still contain valid information for dealing with different phenomena.

In terms of driver selection, solar azimuth angle *azi*, relative measurement height within boundary layer, temperature and water vapor gradients between surface and atmosphere were selected according to Xu et al., 2017. To explain the diurnal cycle and solar radiation, we derived and included cos(azi) and sin(azi). Considering the relative measurement height in the ABL, z_m/z_i, not only permits combining eddy covariance measurements from multiple heights z_m, but also to explicitly account for vertical flux divergence during ERF projection. This property allows us to project to all vertical levels above displacement height. Lastly, the vertical gradients of temperature and water vapor were explained using air potential temperature (θ in K) and mole fraction of water vapor in dry air (q in $mmolmol^{-1}$), as well as LST and EVI as corresponding land surface drivers. LST and EVI matrixes were downscaled from MODIS data products, bi-linearly to 100 m in space, and linearly to one-hour in time. Surface properties are more responsible for spatial variability of the response in machine learning, while meteorological drivers, e.g. air temperature, are more responsible for the diurnal cycle and temporal variability. Considering the coarse temporal resolution (8/16 day) of surface properties, we chose to use simple linear interpolation in time for LST and EVI matrixes.

Building on Metzger et al. (2013) and Xu et al. (2017), we used boosted regression trees (BRT) as machine learning technique, which is based on categorization and regression. 55,160 vertical fluxes and 3304 storage fluxes served as the training dataset in machine learning to produce fluxes over grids and volumes, respectively. Ten cross-validations were operated at 10 nodes before the model with the best agreement between the fitted fluxes and the training dataset was selected. In the end, 70,000 split points were used for vertical flux regression and 3040 split points for storage flux regression.

The extracted ERFs were summarized in equidistant response-sensitivity plots (Cacuci, 2003), which show the driver-response relationships stored in the ERF (Fig. 4). In each plot, the ERF was evaluated with random combinations of drivers drawn from uniform distributions, and the response was aggregated for each driver individually. The contribution of each driver to the explained variance was determined from the reduction of R^2 when removing one driver at a time. The resulting ERF acted as a transfer function and was applied to project vertical flux to each 100 m grid cell across the 20×20 km^2 target domain at 122 m, and to project storage flux to the whole virtual box at five levels, 30 m, 53 m, 76 m, 99 m and 122 m, at hourly interval. During projection, for the meteorological drivers we used the median value during the time interval, assuming that the atmospheric state above the target area was spatially homogeneous. In particular, above the blending height, this assumption is weak compared to invoking a homogeneous land surface in the standard eddy covariance technique.

When summarized over the study period, due to the uneven distribution of qualified observations and projections (more observations and projections during daytime than nighttime), the monthly averaged observations and projections were calculated as the mean of the monthly-mean diurnal cycle.

2.2.4. Uncertainty budget

To evaluate the significance of the presented approach, we estimated the uncertainty budget for both storage and vertical fluxes. We used stratified cross-validation to evaluate how well ERF-VCV performs when projecting to areas the tower footprint had never covered during the training period. Following Xu et al. (2017), we divided the target area into four quadrants: northeastern, southeastern, southwestern and northwestern. On this basis, four incomplete training datasets were created, each of which omitting all data from one quadrant by wind direction. For each incomplete training dataset, (i) the ERF was trained with data from three quadrants; (ii) the resulting ERF along with the state variables from the omitted quadrant were used for projection; (iii) The resulting projection was compared to the observation. In this cross-validation, all uncertainty sources through the ERF-VCV method, i.e. input state variables, footprint modeling, and machine learning, were included in this uncertainty quantification.

We use median and median absolute deviation (MAD) for quantifying systematic and random uncertainty, respectively (Croux and Rousseeuw, 1992; Rousseeuw and Verboven, 2002). The resulting combined uncertainty estimates correspond to a single projected grid cell in the for virtual box.

3. Results

3.1. ERF-VCV projected storage flux

Fig. 4 shows ERFs extracted from BRT. To note, the absolute values shown on the y-axis of Fig. 4 do not imply the actual projected flux, as the responses were not projected with the actual driver combination, but random combinations of uniformly distributed samples within the range of the drivers. However, these equidistant plots are most useful for revealing the relationships among driver and response stored in the ERF. Temperature time-rate of change shows a strong diurnal cycle, negative during daytime and positive during nighttime. The pulse at sin (azi) = 0.4 corresponding to 7–9 a.m. local time is the sign of increase in storage flux when convective eddies are overcoming nighttime stable stratification during initiation of the convective boundary layer (CBL). Temperature time-rate of change is larger in shallower ABLs (larger zm/zi), and also has a negative relationship with LST and θ, while it has a negative relationship with EVI. The ERF-VCV predictions fit very well the observed fluxes (-1% for vertical H, 0% for vertical LE, -4% for dT/dt, and -8% for dq/dt), but in all cases the 99% confidence intervals include the unity slope.

On Aug 17, 2014, the lower part of the convective layer developed as follows: From 7:00–8:00 a.m., an initiation of buoyancy on the

surface due to solar forcing was detected (Fig. 5). Warm bubbles built up and began to release from the surface. During 8:00–9:00 a.m., critical buoyancy was reached. As a result, heated air detached from the surface in discrete events, leading to the creation of a spatial pattern with distinct zones that exhibit varying degrees of heating. From 9:00 a.m.–10:00 p.m., the whole volume was continued to be heated.

3.2. ERF-VCV projected vertical flux

Fig. 6 shows an example of the domain-scaled vertical transport of H flux over the predefined 20 km × 20 km target area at 86 m and the measurement height, 122 m. The white cells in the figure are areas with state-space combinations of drivers for which no extracted response relationship exists. Over the whole experiment, the coverage was extended from the original 1% (average footprint area relative to 20 km x 20 km) to 92% ± 3%, and 94% ± 3% for 20 × 20 km^2 target domain for H, LE respectively, where the tolerance here is one standard deviation.

In Fig. 6, for Aug 19, 2014 noontime, the sensible heat flux at 76 m was positive over warmer surfaces and negative over cooler surfaces. This may indicate suitable conditions for surface heterogeneity-induced mesoscale circulations. In addition, vertical flux divergence and blending can be quantified explicitly in space: Buoyant eddies were emitted from individual surface patches giving the impression of a clear spatial separation at lower levels, e.g. 76 m. The flux at 122 m level was spatially less distinct as a result of blending (Fig. 6b).

3.3. Volume controlled net surface atmosphere exchange

When superimposing vertically integrated projected storage flux with the vertical flux grids at 122 m, ERF-VCV NSAE of H and LE were 32.3 Wm^{-2} and 74.4 Wm^{-2}, 33.8% and 15.5% larger compared to tower observed turbulent flux, 24.1 $W m^{-2}$ and 64.4 Wm^{-2} for the study time period (Table 2). Here, vertical flux refers to the sum of vertical turbulent flux and vertical advection flux with transporting scale up to three hours. The sum of NSAE of sensible and latent heat flux was 106.7 Wm^{-2}, 20.6% greater than the turbulent observation, 88.5 Wm^{-2}. The standard deviation of the diurnal cycle over the study period was 120.9 Wm^{-2} and 228.9 Wm^{-2} for ERF-VCV-projected H and LE, and 149.5 Wm^{-2}, 256.6 Wm^{-2} for observed turbulent H and LE, because spatial averaging ERF-VCV projections were less scattered compared to direct observations. The fitted linear relationship shows that the ERF-VCV projections had a larger daytime-nighttime amplitude (Fig. 7). Temporal standard deviation of the spatially aggregated flux first decreased and then stabilized at about 10–15 km spatial scale (Fig. 8). Incorporation of the ERF-VCV approach also reflects the estimated energy fluxes increased until converging around the same 10–15 km spatial scale.

Storage flux is non-negligible even at the monthly scale, and exhibits a sizeable diurnal cycle at hourly temporal resolution for the virtual box of 122 m height, e.g. 42 Wm^{-2} amplitude (~30% of vertical transport flux) for storage flux of H (Fig. 9).

The median systematic and random uncertainty terms (median absolute deviation, in parenthesis) per single projected cell are 10% (162%), −19% (185%), −15% (232%), and −19% (178%) for vertical flux of H, storage flux of H, vertical flux of LE, and storage flux of LE, respectively. Here, we see that the ERF-VCV approach tends to overestimate vertical H by 10%, and to underestimate vertical LE by 15%. The ERF-VCV approach underestimated the storage fluxes of H and LE by 19%. On this basis, an overall uncertainty of ≪ 15% for H and LE vertical fluxes and ≪ 20% for storage flux is determined. To note, this systematic uncertainty is raw output from the machine projection, and subsequently corrected across the target-area via site-specific linear regression. In result, the + 20.6% or + 18.2 Wm^{-2} difference for the sum of H and LE between ERF-VCV-controlled NSAE projection and turbulent flux tower observation is shown. In addition, ensemble random uncertainty becomes very small when aggregating flux grid cells over the whole target domain: the ensemble random uncertainty for hourly projection over the target domain is confined to within 1% for all fluxes.

Table 2
Comparison between aggregated tower observed turbulent flux and volume-controlled NSAE over the study period. Numbers in parenthesis are standard deviation of the diurnal cycle.

	H (Wm^{-2})	LE (Wm^{-2})	H + LE (Wm^{-2})
Tower observed turbulent flux	24.1(± 149.5)	64.4(± 256.6)	88.5
Volume-controlled NSAE	32.3(± 120.9)	74.4(± 228.9)	106.7

4. Discussion

4.1. Evaluation of the assumptions made in the ERF-VCV theory

A number of assumptions influence the reliability of the ERF-VCV approach (Metzger, this issue). The first assumption is that the tower sufficiently samples both updrafts and downdrafts. We believe this is satisfied in the current case study because the near-zero averaged vertical wind speed ($\overline{w}$ = 0.01 ms^{-1}) over the entire study period. and non-zero hourly vertial wind speeds. These two characteristics imply that both updrafts and downdrafts were sampled by the WLEF tower during the study period.

The second assumption is that observations are dominated by surface fluxes rather than entrainment from the top of the boundary layer, which ensures the relationships reflect information content from the chosen surface drivers. This assumption is fulfilled by filtering flux responses for connectivity with the surface through turbulent mixing tests, primarily the ITC test, which tends to omit data affected by strong negative vertical velocity. The ITC test does not omit situations when surface heterogeneity-induced downdraft occurs in a growing boundary layer, since the net velocity is still upward. Thus, the downward branch of a mesoscale circulation is still sampled by the tower and used as training data in ERF-VCV.

Another implicit assumption is that flux footprint can attribute the main contribution of vertical flux to appropriate surface area even with transporting scale up to three hours. In other words, we have to assume that the majority contribution is from the last touchdown with the surface instead of entrainment (Flesch et al., 1996; Metzger, this issue). Since the majority contribution of the vertical flux is from vertical turbulent flux term (Finnigan et al., 2003; Finnigan, 2004), we use a flux footprint model instead of concentration footprint model. Further, the extracted relationships are built between flux and flux footprint-weighted land surface drivers as well as meteorological drivers. Although flux footprint may fail to attribute low frequency flux contribution to the appropriate surface drivers, the response relationship can still be propagated via the corresponding combination of meteorological drivers. At this study site, the extracted ERFs and flux projections appear sound. This suggests that over the ensemble of thousands of observations the simple footprint parameterization is sufficient to accurately and precisely relate a large fraction of the flux responses to their surface drivers.

Though these assumptions do require additional testing, our results support the idea that ERF-VCV projections enable explicitly identifying and quantifying vertical turbulent flux and storage flux. As shown in Fig. 6, the method was able to detect the initiation of buoyancy from solar forcing, development of critical buoyancy leading to a well-mixed volume, and continue mixing driving by mesoscale circulations generated by surface heterogeneity. These results are remarkably similar to theoretical expectations of boundary-layer mixing.

The ERF-VCV projections suggest that the ERF-VCV technique is a

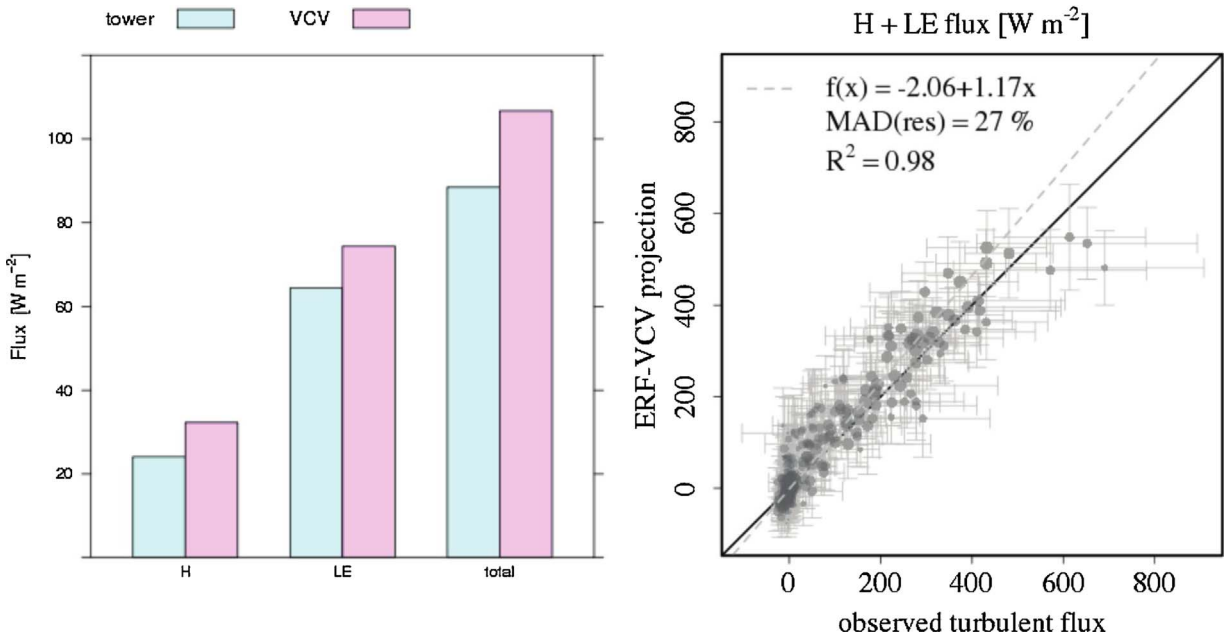

Fig. 7. a. Tower observed turbulent flux (light blue) and VCV-estimated net surface atmosphere exchange (violet) for H, LE and projected net surface atmosphere exchanges (NSAE) integrated over the study period. ERF-VCV estimated NSAE sum of sensible and latent heat flux is 20.6% greater than tower observation when aggregated over the study period. b. Scatterplot for the observed vertical turbulent heat flux measurements for the sum of H and LE and ERF-VCV projected heat flux. Each point represents a one hour averaging period. Uncertainties are one standard deviation of the random uncertainties (error bars in x and y direction). Due to spatial averaging, ERF-VCV projections are less scattered. The fitted linear relationship shows that the ERF-VCV projections have a larger daytime-nighttime amplitude. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

practical and theoretically sound approach to retrieve storage and vertical transport flux in a target box with a maximum (correctable on site-level) systematic uncertainty of less than 20%. The estimation of these terms and further evaluation of them allow us to directly examine several unmet assumptions of eddy covariance in real-world settings which we discuss below: i) location bias of vertical turbulent flux; b) difference of profile to whole domain storage flux; and c) the effect of neglecting vertical advection.

For future development, ERF-VCV requires development of a transporting-scale footprint model to identify and quantify large transporting scale eddies. The transporting scale should not be scaled beyond the spectral gap to avoid ambiguous interpretation of energy transport at a different scale, e.g. synoptic system. Storage flux with inclusion of transporting scale up to three hours could be obtained using wavelet decomposition to match the transporting scale of vertical flux. The overlapping information content within different drivers (multicollinearity) can lead to the extraction of process relationships by the machine-learning that could not be reconciled with physical theory.

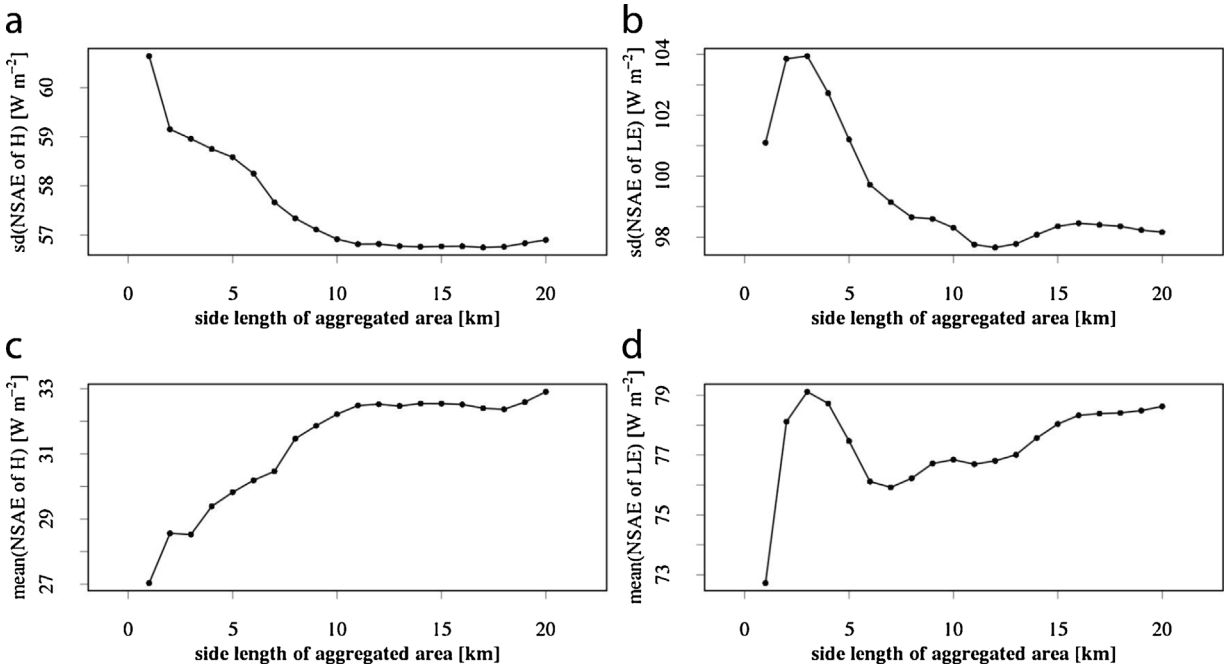

Fig. 8. Temporal standard deviation (top) and mean (bottom) of hourly area-aggregated net surface flux as a function of the control volume side length. The left and right panels show net sensible and latent heat fluxes, respectively. Temporal standard deviation decreased and stabilized at spatial transporting scale of 10–15 km.

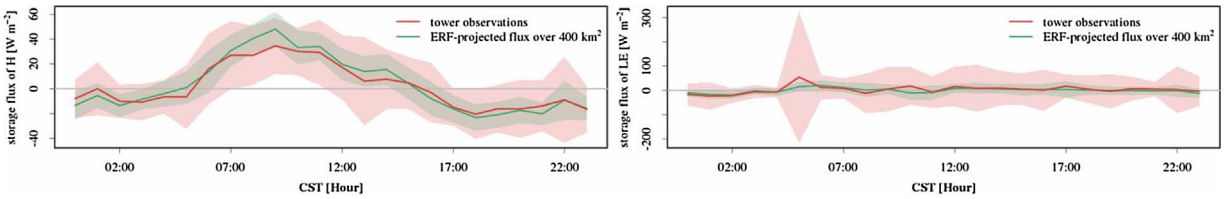

Fig. 9. Diurnal cycle of ERF-VCV projected storage flux (green) and tower observed storage flux (red) for H and LE generated from all data for the study period. Shading area is the temporal spatial variation for the study area and time period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

It will be interesting to work on reconciling additional constraints and multicollinearity in machine-learning. Lastly, with sub-canopy eddy covariance measurements, ERF-VCV should be applicable to capture the features of eddy scalar and heat transport in sub-canopy.

4.2. The implication on location bias in eddy covariance measurements

Combining reliable, precise measurements with time-frequency decomposition and flux footprint modeling allows us to transform our observations into a unified Eulerian coordinate representation, through which we can finally estimate terms that used to be inaccessible. Our study reflects that storage flux has very strong diurnal amplitude (42 W m^{-2}). Therefore, storage flux should not be ignored in observation and model-data comparison, at hourly temporal resolution, especially in the case of tall towers. However, in many cases storage flux is usually assumed to rapidly converge to zero and thus is not measured at many sites (Foken 2008).

ERF-VCV estimation has a larger vertical H than tower turbulent observations during nighttime, likely owing to a higher LST across the target domain than the footprint area (Fig. 2). Similarly, a smaller vertical LE during daytime is observed in association with a slightly lower EVI across the domain. In addition, ERF-VCV NSAE shows larger magnitude both in daytime and nighttime despite substantial scatter in the tower observations. This reflects the influence of the diurnal cycle on the transient footprint. Further, in theory, ERF-VCV-projected vertical flux is expected to be larger than tower turbulent observation because projected vertical flux includes the energy contribution from mesoscale eddies (Foken et al., 2006). The inclusion of mesoscale eddies tends to increase emitted heat flux, since spatially confined thus under-sampled updrafts are net warmer compared to spatially expansive thus over-sampled downdrafts. The resulted difference between the tower observation and ERF-VCV estimation reflects that WLEF tower cannot consistently represent the mean VCV estimated surface-atmosphere flux, implying location bias is pervasive in this case and potentially reflective also of other flux tower sites in heterogeneous terrain.

4.3. The role of horizontal advection and the implication on spatial transporting scale

Theoretically, in a mesoscale circulation we expect vertical advection to be compensated by horizontal advection and storage flux (Mahrt 1998; Foken et al., 2003; Mauder et al., 2008). Hence, studies considering only vertical advection instead of advection from both vertical and horizontal directions have incurred criticism. For example, Finnigan et al. (2003) and Mauder et al. (2008) argue that only considering the vertical advection term can produce unrealistic large convection and rather noisy results on an observation-by-observation basis.

Here, we argue that the spatial aggregation of projected fluxes over heterogeneous terrain improves the resilience of the ERF-VCV projection to horizontal advection. Especially when the spatial scale of aggregation is large, the net low-frequency vertical flux due to updrafts and downdrafts trends towards its ensemble mean. As shown in Fig. 8, the temporal standard deviation of the spatially aggregated flux first decreased and then stabilized at about 10–15 km horizontal scale. This indicates that ERF-VCV also projects the localized vertical advection resulting from the observation-by-observation compensation of horizontal advection. However, when integrated over a critical landscape scale, positive and negative compensatory fluxes cancel out each other, and only the net low-frequency flux remains. This justifies the choice of neglecting horizontal advection when aggregating over a horizontal scale of order 10 km as compared to point-by-point observations. Our finding here is validated in theory (Metzger, this issue) and generally agrees with previous studies suggesting that $w = 0$ over a large spatial domain is a much weaker assumption compared to $w = 0$ for an individual measurement location (Mahrt 1998).

In addition, the indicated surface patch size of 15 km implies that it is the minimum reliable patch size in the comparison of WLEF eddy covariance flux observations with model and remote sensing data, akin to a critical landscape scale. Our approach thus provides a suitable bridge for the spatial gap between WLEF tower-measured fluxes and both, remote sensing products and ESM outputs.

4.4. Implications of ERF-VCV on the energy balance closure problem

As discussed in Sect. 4.1, surface heterogeneity-induced updrafts and downdrafts were sampled by the tower at different times during the study period (Fig. 3). While changing wind direction and varying source areas complicate direct interpretation of the flux observations, they provide an opportunity for including secondary circulations in ERF-VCV.

At any given time a single tower is unable capture both updrafts and downdrafts. Due to the skewed spatial distribution of updrafts and downdrafts, a tower observes sparser but stronger updrafts only occasionally, if at all (Fig. 3b). More frequently, a tower observes abundant but weaker downdrafts (Fig. 3c), or cannot capture any low frequency energy transport at all (Fig. 3a). For example Kanda et al. (2004) and Mahrt (1998) link this spatial patterning and corresponding conditional sampling characteristics to a buoyancy-related underestimation in particular of the sensible heat flux.

Different from the standard tower eddy covariance technique, ERF-VCV provides a potential approach for improving energy balance closure through spatially and temporally explicit flux projections: the projections include the likely spatial distribution of both low frequency updrafts and downdrafts. Once aggregated to the critical landscape spatial scale the compensatory fluxes cancel out. The resulting net flux trends positively to its ensemble mean (Fig. 8c,d), which exceeded the standard tower flux observations over the study period. This warrants investigating further impacts on the energy balance closure problem, once spatially explicit net radiation and ground heat flux are available, e.g. from intensive in-situ observations.

5. Conclusions

Using AmeriFlux Park Falls WLEF tall tower in North Wisconsin, USA during July and August 2014, ERF-VCV proved useful for retrieving the volume-controlled net surface atmosphere exchange (NSAE). This retrieval is achieved by resolving the storage flux, vertical turbulent, and vertical advection fluxes, which are not easily measured.

ERF-VCV can improve the tower observation from a footprint-variable representation to a fixed-coordinate representation. This aids reducing the location bias typically incurred from single-location vertical turbulent flux and single-profile storage flux measurements, as well as the influence of vertical advection. Particularly, in this study, storage flux did not converge to zero as often assumed and had significant diurnal cycle, and should thus not be ignored when comparing eddy covariance measurements with earth system models (ESMs) at hourly or finer temporal resolution. Low frequency flux contributions were detected in this study, and inclusion in the ERF-VCV landscape-scale exchange of sensible and latent heat led to a 20.6% increase over the tower observations. In addition, the derived flux can spatiotemporally resolve mesoscale circulation that contribute to this source of energy imbalance. Further, we show that spatial aggregated fluxes over heterogeneous land cover resilient to horizontal advection. Lastly, in our case we find that ERF-VCV-estimated NSAE always increased the turbulent heat fluxes and thus provides a promising research direction for improving energy balance closure.

Substantial improvements are still possible and needed for these kinds of scaling and rectification methods. First, high intensity in-situ observations or large eddy simulations (LES) model results can be used to comprehensively evaluate and verify assumptions of the ERF-VCV process. Second, with spatially explicit net radiation and ground heat flux, the potential for improving the energy balance closure problem can be studied. Third, multiple flux tower or airborne eddy covariance data lend themselves to investigate how the horizontal advection terms not addressed in this study could be explicitly resolved.

ERF-VCV is applicable to retrieve volume-controlled NSAE across a target area. This practically permits rectifying footprint bias for eddy covariance flux in model data comparison, and provides the potential direction to improve energy budget closure in eddy covariance technique.

Acknowledgements

This work was supported by the National Ecological Observatory Network contract #1010-0196-000 to U. Wisconsin, DOE Office of Science Ameriflux Network Management Project subaward to ChEAS Core Site Cluster, and NSF Advances in Biological Informatics awards #1062204 and #1457897. The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by Battelle Ecology, Inc. This material is based 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.

References

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Evans, R., 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. 11, 2415–2434. http://dx.doi.org/10.1175/1520-0477(2001)082≪2415:FANTTS≫>2.3. (CO;2).

Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., Swenson, S.C., 2011. Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. J. Geophys. Res.: Biogeosci. 116 (G2). http://dx.doi.org/10.1029/2010jg001593. (2005–2012).

Brock, F.V., 1986. A nonlinear filter to remove impulse noise from meteorological data. J. Atmos. Oceanic Technol. 3 (1), 51–58. http://dx.doi.org/10.1175/1520-0426(1986)003≪0051:ANFTRI≫>2.0. (CO;2).

Cacuci, D.G., 2003. Sensitivity & Uncertainty Analysis, Volume 1: Theory. CRC Presshttp://dx.doi.org/10.1201/9780203498798.

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 (3), 303–327. http://dx.doi.org/10.1007/s10546-014-9922-6.

Chen, B., Coops, N.C., Fu, D., Margolis, H.A., Amiro, B.D., Barr, A.G., Flanagan, L.B., 2011. Assessing eddy-covariance flux tower location bias across the Fluxnet-Canada Research Network based on remote sensing and footprint modelling. Agricul. Forest Meteorol. 151 (1), 87–100. http://dx.doi.org/10.1016/j.agrformet.2010.09.005.

Croux, C., Rousseeuw, P.J., 1992. Time-efficient Algorithms for Two Highly Robust Estimators of Scale. Springerhttp://dx.doi.org/10.1007/978-3-662-26811-7_58.

Desai, A.R., Richardson, A.D., Moffat, A.M., Kattge, J., Hollinger, D.Y., Barr, A., Stauch, V.J., 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agricul. Forest Meteorol. 148 (6-7), 821–838. http://dx.doi.org/10.1016/j.agrformet.2007.11.012.

Desai, A.R., Xu, K., Tian, H., Weishampel, P., Thom, J., Baumann, D., Kolka, R., 2015. Landscape-level terrestrial methane flux observed from a very tall tower. Agric. Forest Meteorol. 201, 61–75. http://dx.doi.org/10.1016/j.agrformet.2014.10.017.

Dietze, M.C., Serbin, S.P., Davidson, C., Desai, A.R., Feng, X., Kelly, R., McHenry, K., 2014. A quantitative assessment of a terrestrial biosphere model's data needs across North American biomes. J. Geophys. Res.: Biogeosci. 3, 286–300. http://dx.doi.org/10.1002/2013JG002392.

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. Agricul. Forest Meteorol. 211 (-212), 115–127. http://dx.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 Meteorol. 107 (1), 1–48. http://dx.doi.org/10.1023/a:1021554900225.

Finnigan, J., 2004. A re-evaluation of long-term flux measurement techniques part II: coordinate systems. Boundary-Layer Meteorol. 113 (1), 1–41. http://dx.doi.org/10.1023/B:BOUN.0000037348.64252.45.

Finnigan, J., 2008. An introduction to flux measurements in difficult conditions. Ecol. Appl. 18 (6), 1340–1350. http://dx.doi.org/10.1890/07-2105.1.

Foken, T., Nappo, C.J., 2008. Micrometeorology. Springer.

Foken, T., Wimmer, F., Mauder, M., Thomas, C., Liebethal, C., 2006. Some aspects of the energy balance closure problem. Atmos. Chem. Phys. 6, 4395–4402.

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. Bull. Am. Meteorol. Soc. 92 (4), ES13–ES18. http://dx.doi.org/10.1175/2011BAMS3130.1.

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

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

Leuning, R., van Gorsel, E., Massman, W.J., Isaac, P.R., 2012. Reflections on the surface energy imbalance problem. Agricul. Forest Meteorol. 156, 65–74. http://dx.doi.org/10.1016/j.agrformet.2011.12.002.

Mahrt, L., 1998. Flux sampling errors for aircraft and towers. J. Atmos. Oceanic Technol. 15 (2), 416–429. http://dx.doi.org/10.1175/1520-0426(1998)015≪0416:fsefaa≫>2.0. (co;2).

Mauder, M., Desjardins, R.L., 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. http://dx.doi.org/10.1007/s10546-008-9279-9.

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

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H.P., Foken, T., 2012. Eddy-covariance flux measurements with a weight-shift microlight aircraft. Atmos. Meas. Tech. 5 (7), 1699–1717. http://dx.doi.org/10.5194/amt-5-1699-2012.

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

Metzger, S., Durden, D., Sturtevant, C., Luo, H., Pingintha-Durden, N., Sachs, T., Desai, A.R., 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 Develop. 10 (9), 3189–3206. http://dx.doi.org/10.5194/gmd-10-3189-2017.

Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Desai, A.R., 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. Global Change Biol. 18 (2), 566–584. http://dx.doi.org/10.1111/j. 1365-2486.2011.02562. x.

Rousseeuw, P.J., Verboven, S., 2002. Robust estimation in very small samples. Comput. Statis. Data Anal. 40 (4), 741–758. http://dx.doi.org/10.1016/S0167-9473(02)00078-6.

Sakai, R.K., Fitzjarrald, D.R., Moore, K.E., 2001. Importance of low-frequency contributions to eddy fluxes observed over rough surfaces. J. App. Meteorol. 40 (12), 2178–2192. http://dx.doi.org/10.1175/1520-0450(2001)040≪2178:iolfct≫>2.0. (co;2).

Schaefer, K., Schwalm, C.R., Williams, C., Arain, M.A., Barr, A., Chen, J.M., Hollinger, D.Y., 2012. A model-data comparison of gross primary productivity: results from the North American Carbon Program site synthesis. J. Geophys. Res.: Biogeosci. 117 (G3). http://dx.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. http://dx.doi.org/10.1007/s10546-014-9962-y.

Schmid, H.P., 1997. Experimental design for flux measurements: matching scales of observations and fluxes. Agricul. Forest Meteorol. 87 (2-3), 179–200. http://dx.doi.org/

10.1016/s0168-1923(97)00011-7.
Schotanus, P., Nieuwstadt, F., De Bruin, H., 1983. Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes. Boundary-Layer Meteorol. 26 (1), 81–93. http://dx.doi.org/10.1007/BF00164332.
Starkenburg, D., Fochesatto, G.J., Cristóbal, J., Prakash, A., Gens, R., Alfieri, J.G., Kane, D.L., 2015. Temperature regimes and turbulent heat fluxes across a heterogeneous canopy in an Alaskan boreal forest. J. Geophys. Res.: Atmosp. 120 (4), 1348–1360. http://dx.doi.org/10.1002/2014JD022338.
Steinfeld, G., Letzel, M.O., 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, P.C., Mauder, M., Foken, T., Marcolla, B., Boegh, E., Ibrom, A., Bernhofer, C., 2013. A data-driven analysis of energy balance closure across FLUXNET research sites: the role of landscape scale heterogeneity. Agricul. Forest Meteorol. 171, 137–152. http://dx.doi.org/10.1016/j.agrformet.2012.11.004.
Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78. http://dx.doi.org/10.1175/1520-0477(1998)079≪0061:APGTWA≫>2.0. (CO;2).
Wilczak, J.M., Oncley, S.P., Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. Boundary-Layer Meteorol. 99 (1), 127–150. http://dx.doi.org/10.1023/A:1018966204465.
Xu, K., Metzger, S., Desai, A.R., 2017. Upscaling tower-observed turbulent exchange at fine spatio-temporalresolution using environmental response functions. Agricul. Forest Meteorol. 232, 10–22. http://dx.doi.org/10.1016/j.agrformet.2016.07.019.