Does increased spatial replication above heterogeneous agroforestry improve the representativeness of eddy covariance measurements?

José Ángel Callejas-Rodelas[1], Alexander Knohl[1,2], Ivan Mammarella[3], Timo Vesala[3,4], Olli Peltola[5], and Christian Markwitz[1]

[1]University of Göttingen, Bioclimatology, Göttingen, Germany
[2]Centre for Biodiversity and Land Use, University of Göttingen, Göttingen, Germany
[3]Institute for Atmosphere and Earth System Research (INAR)/Physics, Faculty of Science, University of Helsinki
[4]Institute for Atmosphere and Earth System Research (INAR)/Forest Science, Faculty of Agriculture and Forestry, University of Helsinki
[5]Natural Resources Institute Finland (LUKE), Latokartanonkaari 9, Helsinki, 00790, Finland

Correspondence: José Ángel Callejas-Rodelas (joseangel.callejasrodelas@uni-goettingen.de)

Abstract. Spatial heterogeneity in terrestrial ecosystems compromises the accuracy of eddy covariance measurements. Examples of heterogeneous ecosystems are temperate agroforestry systems, that have been poorly studied by eddy covariance. Agroforestry systems get an increasing attention due to their potential environmental benefits, e.g. a higher carbon sequestration, improved microclimate and erosion reduction compared to open cropland agricultural systems. Lower-cost eddy covariance setups might offer an opportunity to better capture spatial heterogeneity by allowing for more spatial replicates of flux towers. The aim of this study was to quantify the spatial variability of carbon dioxide (*FC*), latent heat (*LE*) and sensible heat (*H*) fluxes above a heterogeneous agroforestry system in northern Germany using a distributed network of three lower-cost eddy covariance setups across the agroforestry systems. Fluxes from the three towers in the agroforestry were further compared to fluxes from an adjacent open cropland site. The campaign took place from March 2023 until September 2024. The results indicated that the spatial variability of fluxes was largest for *FC*, attributed to the effect of different crops (rapeseed, corn and barley) within the flux footprints contributing to the measured fluxes. Differences between fluxes across towers were enhanced after harvest events. However, the temporal variability due to the seasonality and diurnal cycles during the campaign was larger than the spatial variability across the three towers. When comparing fluxes between the agroforestry and the open cropland systems, weekly sums of carbon and evapotranspiration fluxes followed similar seasonality, with peak values of -50 g C m^{-2} $week^{-1}$ and 40 mm $week^{-1}$ during the growing season, respectively. The variation of the magnitude depended on the phenology of the different crops. The effect size, which is an indicator of the representativeness of the fluxes across the distributed network of three eddy covariance towers against only one, showed in conjunction with the other results that the spatial heterogeneity across the agroforestry was better captured by the network of three stations. This supports previous findings that spatial heterogeneity should be taken into account in eddy covariance studies, and that lower-cost setups may offer the opportunity to bridge this gap and improve the accuracy of eddy covariance measurements above heterogeneous ecosystems.

1 Introduction

The eddy covariance (EC) technique is the central approach to measuring the exchange of energy, trace gases and momentum between terrestrial ecosystems and the atmosphere (Baldocchi, 2014). The EC technique has been established as a standard method within the scientific community when rapid-response instruments, capable of measuring wind speed, temperature, and gas concentrations over the major frequency ranges of the turbulent energy spectrum became commercially available (Aubinet et al., 2012; Wohlfahrt et al., 2009). These instruments provided the capability to measure the exchange of energy and matter between the land surface and the atmosphere, driven by eddies of diverse sizes and frequencies (Kaimal and Finnigan, 1994).

At a majority of flux sites, a single EC station is installed (Hill et al., 2017) and measurements are made based on the ergodic hypothesis. The ergodic hypothesis states that covariances (fluxes) calculated over the time domain are equivalent to covariances calculated over the spatial domain (Higgins et al., 2013). The measured turbulent fluxes and carbon and water balances, when integrated over a defined time interval, are representative of the tower footprint area corresponding to the averaging interval (Vesala et al., 2008). This is true for homogeneous sites where the spatial representativeness of fluxes within the ecosystem of interest is guaranteed with a high degree of confidence (Hurlbert, 1984). However, these conditions of homogeneity are often not met in many ecologically and socioeconomically interesting sites, such as mixed forests, wetlands, urban forest interfaces or small-scale farmlands (Finnigan et al., 2003; Hill et al., 2017).

Agroforestry (AF) systems are an example of heterogeneous agroecosystems. They combine trees and crops on the same agricultural land in order to benefit from the presence of trees on the land (Veldkamp et al., 2023; Kay et al., 2019). These systems offer several benefits, including the potential to prevent wind erosion over crops (van Ramshorst et al., 2022; Böhm et al., 2014), improve soil fertility (Kanzler et al., 2021), or reduce water loss through evaporation in crops (Kanzler et al., 2019). Short Rotation Alley Cropping systems, a type of agroforestry, represent an alternative land use practice with the potential to increase carbon sequestration and improve water use efficiency (WUE) in comparison to conventional open cropland (OC) agriculture (Markwitz et al., 2020; Veldkamp et al., 2023). These AF systems consist of alternating rows of trees and crops. The trees employed in these systems are typically fast-growing species, such as poplar (*Populus*) or willow (*Salix*), and are harvested in cycles of 5-6 years for biomass production. Crops are cultivated in an annual rotation.

In general, heterogeneity poses a challenge for EC measurements and, in a broader context, for any type of measurement across the atmospheric boundary layer (Bou-Zeid et al., 2020). Heterogeneity in surface properties induces horizontal advection, secondary mesoscale circulations and non-equilibrium turbulence processes, which occur near and downstream of changes in the surface properties (Bou-Zeid et al., 2020). As shown by previous studies over heterogeneous sites, such as pine forest (Katul et al., 1999; Oren et al., 2006) or managed grassland (Peltola et al., 2015), spatial heterogeneity induced relevant spatial variability in the EC measured fluxes. According to the classification of Bou-Zeid et al. (2020), the heterogeneity of these AF systems can be classified as unstructured heterogeneity (Fig. 1 therein), because the site consists of a certain number of interleaved trees and crop strips, but it is small enough that the AF site might be affected by other elements in the surrounding landscape. Upon changes in surface properties (like roughness or moisture), the mean wind field and the turbulence adjust to

the new surface, with more complex effects on the flow when multiple changes in the surface properties co-occur, as it is the case at the AF (Bou-Zeid et al., 2020).

The location of the EC station within a land use system has been demonstrated to potentially introduce a bias in the measured fluxes (Chen et al., 2011), indicating that a single EC station may not be sufficient to properly account for the spatial variability of fluxes induced by landscape heterogeneity (Katul et al., 1999). The high cost and labor intensity of deploying an EC station are the main reasons for the lack of spatial replicates of EC measurements in many studies (Hill et al., 2017). The infrared gas analyzer (IRGA), the crucial component to measure trace gases, typically accounts for a large proportion of the total installation costs associated with an EC station. Lower-cost EC (LC-EC) setups represent a potential solution to the spatial replication problem of EC measurements, as several EC stations could be deployed for the cost of a single conventional station. LC-EC employ a more economical infrared gas analyser and a sonic anemometer, though these instruments necessitate more rigorous post-processing corrections. Notably, previous studies have demonstrated that LC-EC setup can yield comparable results to those of conventional EC (CON-EC) setups. Hill et al. (2017) compared a custom-built LC-EC setup for CO_2 and H_2O measurements with a CON-EC, with very good agreement in CO_2 and H_2O fluxes. In addition, a different LC-EC setup for H_2O flux measurements was compared with a conventional setup (Markwitz and Siebicke, 2019), resulting in good agreement in H_2O fluxes. Furthermore, another version of the LC-EC setup deployed in Hill et al. (2017) was extensively validated in the studies of Callejas-Rodelas et al. (2024) and van Ramshorst et al. (2024), with very good agreement in CO_2 fluxes and good agreement in H_2O fluxes.

The LC-EC setups can allow for a higher degree of spatial replication of EC and support conventional EC setups. In addition, they provide a powerful tool for the verification of carbon and water balances in the agricultural and forestry sectors in developing carbon credit markets (Trouwloon et al., 2023) or for an improved water management. However, the increased uncertainty associated with these setups must be taken into account when calculating balances of energy, carbon, or other variables, and when comparing different land uses. One of the main differences between LC-EC and CON-EC setups is the spectral response of the sensors. The LC-EC setups used in the Callejas-Rodelas et al. (2024), Cunliffe et al. (2022), Hill et al. (2017) and van Ramshorst et al. (2024) studies were characterized by a slower frequency response in CO_2 and H_2O measurements, which induces a higher spectral attenuation in the high-frequency range of the turbulent energy spectrum, compared to CON-EC. The higher attenuation introduces a greater degree of uncertainty when applying spectral corrections, as observed by Ibrom et al. (2007) and Mammarella et al. (2009), among others.

The impact of landscape heterogeneity within an AF system on turbulence, latent heat flux (*LE*), sensible heat flux (*H*) and carbon dioxide flux (*FC*) remains to be examined. Markwitz and Siebicke (2019) and Markwitz et al. (2020) conducted evapotranspiration (*ET*) measurements across multiple AF and OC systems in Northern Germany; however, their measurements were not replicated within a single site. In contrast, in the study of Cunliffe et al. (2022) a total of eight LC-EC setups were deployed in different locations across a landscape of ecological interest (Cunliffe et al., 2022). The objective of this study was to capture the heterogeneity of *FC* and *ET* across a semiarid ecosystem, with low magnitude of both *FC* and *ET*. Replicated EC measurements in heterogeneous agroforestry systems are so far lacking.

In the present study, a network of three LC-EC setups was deployed, analogous to those utilized in the studies of Callejas-Rodelas et al. (2024), Cunliffe et al. (2022) and van Ramshorst et al. (2024), above an AF site, and one additional LC-EC setup above an adjacent OC site in northern Germany. To the best of our knowledge, this was the first time a distributed network of EC towers has been installed above a temperate agroforestry system. With one and a half years of concomitant flux data from the four EC setups, the objective was to quantify the spatial and temporal variability of *FC* and *LE*, as well as the statistical effect of the increased spatial replication of EC measurements above a heterogeneous site. According to Hill et al. (2017), it is possible to estimate the sampling variability and total uncertainty for an ecosystem with independent spatial replication of EC measurements. This allows for the estimation of the effect size (see Section 2). The present study tested the hypothesis that the increased uncertainty inherent to the use of slower-frequency response sensors in EC measurements can be counteracted by the improvement of spatial replication of EC, which increases its statistical robustness. The objectives of this study were threefold: (i) to quantify the spatial and temporal variability of turbulent fluxes and parameters above AF; (ii) to calculate the effect size of the experimental site at the daily scale, following Hill et al. (2017); and (iii) to compare the ecological functioning of the AF to the OC in terms of *FC* and *ET* balances.

2 Methods

2.1 Site description

The measurements were conducted from 1 March 2023 to 19 September 2024 at an agroforestry system located in Wendhausen (Lehre), Lower Saxony, Germany (52.63° N, 10.63° E). Elevation above sea level is 80 m. The field is divided into two distinct systems: an AF system (17.3 ha) in the north and a OC system (8.5 ha) in the south (see Fig. 1). The crops cultivated within both systems kept a similar distribution from west to east. In 2023, rapeseed was cultivated at the western side, barley at the eastern side, and corn at the center (Fig. 1a). In 2024, rapeseed was cultivated at the eastern side, barley at the center, and corn at the western side (Fig. 1b). The management of the crops was similar at both AF and OC sites and crops were fertilized. The mean long-term annual precipitation is 617 mm, and the mean annual air temperature is 9.9 °C, for the reference period 1981-2010 at Braunschweig airport ((DWD, 2024)). The soil at both AF and OC sites was classified as a Clay Cambisol, with an organic carbon (SOC) content of 5.8 kg C m^{-2} at the OC and and 6.75 kg C m^{-2} at the AF. Additionally, the soil bulk density was determined as 1.0 g cm^{-3} at both AF and OC (Veldkamp et al., 2023). Soil characteristic were last measured in 2019.

The harvest of rapeseed, barley and corn in the 2023 campaign season occurred on 13 July, 22 August and 26 and September. The harvest of rapeseed, barley and corn in the campaign of 2024 took place on 15 July, 5 August and 13 September, respectively. In 2024, rapeseed did not grow well and a mulch cut was carried out, therefore the eastern part of the field was covered by a combination of grasses, bare soil and mulch. Canopy height was estimated from pictures taken during field visits. The maximum height attained by the crops at the peak of their development stage was around 1.5 m for rapeseed, 2.5 m for corn and 1.3 m for barley. The trees present at the AF system are fast-growing poplar (*Populus nigra* × *Populus maximowiczii*)

and are typically harvested every 4 to 5 years. The most recent tree harvest occurred in 2019. Trees grew from around 4.0 m till 5.5 m on average across the measurement period.

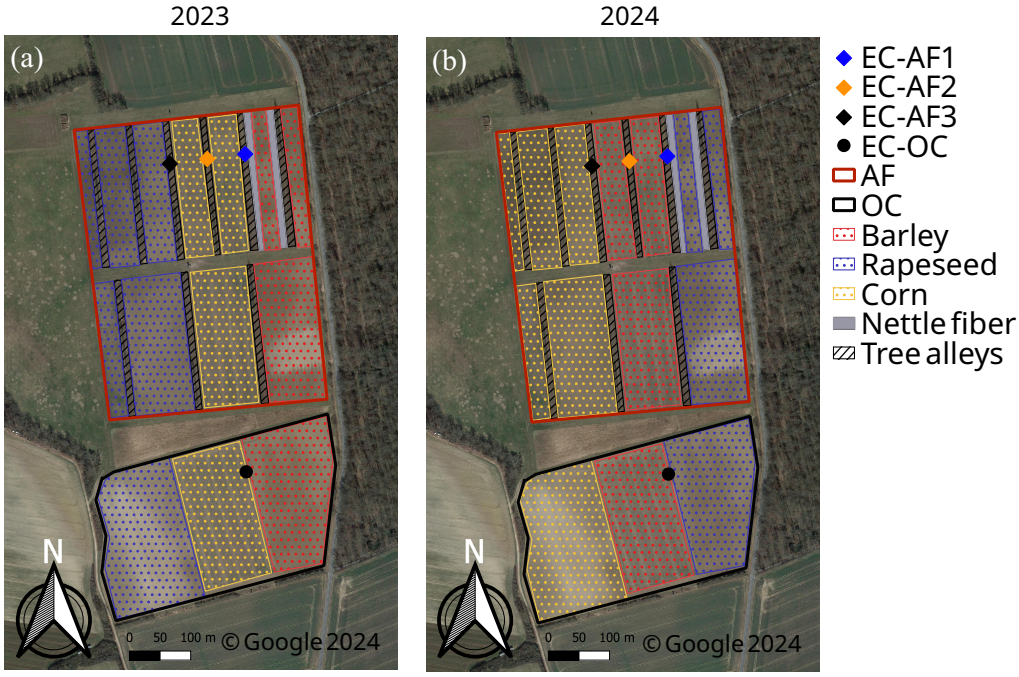

Figure 1. Satellite view and land cover classification of the experimental site for 2023 (a) and 2024 (b), together with the location of the EC stations (blue diamond for EC-AF1, orange diamond for EC-AF2, black diamond for EC-AF3 and black circle for EC-OC). The area bordered red corresponds to the AF system and the area bordered blue to the OC system. Figure created with QGIS v. 3.22, aerial map by Google Satellite Maps. © Google 2024.

2.2 Experimental setup

Measurements were made at four EC stations, one located at the OC site and three located at the AF site (Fig. 1). The stations are designated as OC, AF1, AF2 and AF3. Each station was equipped with a complete set of meteorological sensors and a LC-EC setup (see Table 1 in Callejas-Rodelas et al., 2024). The measured meteorological variables were air temperature (*TA*), relative humidity (*RH*), atmospheric pressure (*PA*), precipitation (*P*), global radiation (*SW_IN*), outgoing shortwave (*SW_OUT*) and longwave (*LW_OUT*) radiation, and net radiation (*NETRAD*). The EC measurement heights were 10 m above ground for AF1, AF2 and AF3, and 3.5 m for OC. Only one photosynthetic active radiation (*PPFD_IN*) sensor was installed at AF1, and two barometers for atmospheric pressure measurements were installed at AF1 and AF2. All the stations were equipped with two soil heat flux plates to measure soil heat flux (*G*) at 5 cm depth. Only one soil heat flux plate was installed at AF3. Radiation sensors were placed in a beam facing south at 9.5 m height at AF1, AF2 and AF3 and at 3 m height at OC. *TA* and *RH* measurements were taken at 2 m height at all stations. *P* was measured at 1 (AF1, OC) or 1.5 (AF2, AF3) m height. Meteorological data were recorded on CR1000X dataloggers (Campbell Scientific Inc., Logan, UT, USA).

The LC-EC setups consisted of a three-dimensional sonic anemometer for wind measurements (uSonic3-Omni, METEK GmbH, Elmshorn, Germany) and a gas analyzer enclosure. The enclosure consisted of an IRGA for CO_2 molar density measurements (GMP343, Vaisala Oyj, Helsinki, Finland), and a *RH* capacitance cell for *RH* measurements (HIH-4000, Honeywell International Inc., Charlotte, North Carolina, USA) and was installed at the bottom of the tower. Air was drawn through a 9 m tube at the AF stations and 2.5 m tube at the OC station. Two temperature sensors were installed, one inside the IRGA measuring cell and one inside the enclosure; and two pressure sensors, one to measure differential pressure inside the enclosure and another to measure absolute pressure inside the IRGA measuring cell. Measurements from all components were recorded at 2 Hz frequency on a CR6 datalogger (Campbell Scientific Inc., Logan, UT, USA). A more detailed description of the setup can be found in Callejas-Rodelas et al. (2024).

The GMP343 sensors were calibrated in February 2023 and February 2024. Frequent inspections were performed to clean the tubing, replace filters, measure flow rate, and clean the lens of the GMP343. The nominal flow rate was 5.0 L$\cdot$ min^{-1} at all AF stations, with some drops due to filter clogging.

During the study period, there were generally large percentages of missing data. Missing data were either short gaps (a few 30-minute periods or a few hours) caused by data filtering during the quality control after flux processing (see Section 2.3.3), or longer gaps (hours to a few days) due to power outages during the winter, mostly at night, at all stations. Due to other technical problems, there were few larger gaps at some stations, in particular a gap of three months from mid-July to early October 2023 at AF3, for *FC* and *LE*.

Although generally recommended in EC studies (Aubinet et al., 2012), no storage terms were considered in the calculation of *FC* and *LE* because no concentration profiles were installed at the stations.

2.3 Flux computation

2.3.1 Pre-processing

Data processing prior to flux calculation included (i) the calculation of CO_2 dry mole fraction measurements from the CO_2 molar density provided by default by the instrument, using some sensor-specific parameters and the observed values of pressure and relative humidity in the measurement system (Callejas-Rodelas et al., 2024); and (ii) the calculation of the H_2O dry mole fraction from relative humidity, temperature and pressure measurements inside the measurement cell using the derivation of Markwitz and Siebicke (2019). More details on the pre-processing steps are given in Callejas-Rodelas et al. (2024) and van Ramshorst et al. (2024).

2.3.2 Flux processing

H, *LE*, *FC* and momentum flux were calculated using the EddyUH software (Mammarella et al., 2016) in its Matlab version (MATLAB®R2023a, The Mathworks, Inc., Natick, MA, USA). Raw data were de-spiked using limits for absolute differences between consecutive values. Detrending was performed by block averaging. Wind coordinates were binned into eight sectors of 45° each and rotated according to the planar fit correction procedure of Wilczak et al. (2001), following the default recom-

mendation by ICOS (Sabbatini et al., 2018). Time-lag optimization was performed through cross-covariance maximization, using predefined windows of 2 to 10 s for CO_2 and 2 to 20 s for H_2O (Callejas-Rodelas et al., 2024). Low-frequency losses were corrected after Rannik and Vesala (1999) and high-frequency losses, after Mammarella et al. (2009). The latter is based on determining the time response of CO_2 and H_2O separately, calculated from the measured co-spectra. In the case of CO_2 the time response determined by the experimental method was similar to the nominal time response of 1.36 s calculated in Hill et al. (2017) for the GMP343. This time response was used for all flux calculations for all the four towers. In the case of H_2O the time response was estimated by a exponential fit as a function of relative humidity. Data quality was flagged from 1 to 9 following Foken et al. (2005).

2.3.3 Filtering and gap filling

Fluxes were filtered using data with quality flags < 7 to avoid periods with poorly developed turbulence (Foken et al., 2005). Outliers were removed using a running median absolute deviation (MAD) filter, based on the approach by Mauder et al. (2013), with a window of two weeks. The *q* parameter in Eq. (1) of the paper by Mauder et al. (2013) was set as 7.5. The MAD filter was iterated three times over each time series. Hard upper and lower limits were applied afterwards to remove additional outliers not detected by the MAD filter. Values outside the ranges from -100 W m^{-2} to 700 W m^{-2} for *H*, from -20 W m^{-2} to 700 W m^{-2} for *LE*, and from -50 µmol m^{-2} s^{-1} to 50 µmol m^{-2} s^{-1} for *FC*, were discarded. Additional hard limits were applied specifically to winter (November to February) and transition periods (March and October) separately. The aim was to avoid outliers that went through the previous filters which might bias the application of the gap-filling algorithms. For *LE* and *H*, these limits were of 50 W m^{-2} during winter, and 100 W m^{-2} in March and October. For the *FC*, these limits were (in absolute values) $\pm$ 10 µmol m^{-2} s^{-1} during winter, and $\pm$ 15 µmol m^{-2} s^{-1} in March and October. Finally, a friction velocity (*USTAR*, m s^{-1}) filter was applied to remove periods with non-existent or weak turbulence. The filter of *USTAR* was applied using REddyProc (Wutzler et al., 2018), which removed values based on a *USTAR* threshold calculated as the maximum of the seasonally derived *USTAR* values. These seasonal values were calculated based on Papale et al. (2006). The average *USTAR* thresholds for the stations were 0.21, 0.21, 0.18 and 0.16 m·s^{-1} for AF1, AF2, AF3 and OC, respectively.

Before filtering, the total available data accounted for 63.4 % (AF1), 80.0 % (AF2), 76.2 % (AF3) and 61.5 % (OC) for *FC* and *LE*, respectively, and 85.7 % (AF1), 86.0 % (AF2), 83.1 % (AF3) and 75.9 % (OC) for *H*, respectively, relative to the duration of the entire measurement campaign. These gaps occurred due to instrumental or power failure. After filtering, the available data accounted for 39.3 % (AF1), 49.2 % (AF2), 35.7 % (AF3) and 33.8 % (OC), for *FC*; 42.0 % (AF1), 53.6 % (AF2), 36.4 % (AF3) and 38.7 % (OC) for LE; and 61.5 % (AF1), 61.4 % (AF2), 56.7 % (AF3) and 52.8 % (OC) for *H*.

Meteorological data were gap-filled at the 30-minute time scale to provide complete time series of the predictor variables for flux gap-filling. The procedure differed slightly for the different variables of interest. Short gaps of up to one hour were filled using linear interpolation, except for *P*. Missing data at the AF1 station, when available at the OC station, were filled using linear regression models with the OC data as predictors, and vice versa. Missing data at AF2 and AF3 that were available at AF1 were filled using a similar procedure, with AF1 as the reference. Finally, *P*, *TA*, RH, vapor pressure deficit (*VPD*), *SW_IN*, wind speed (*WS*) and wind direction (*WD*) were filled at the stations using ERA5-Land re-analysis data (Muñoz-Sabater et al.,

2021) as predictors, following the approach implemented in Vuichard and Papale (2015). Linear reduced major axis regression models were derived from the ERA5-Land data and the station data, using the library *pylr2* in Python. The coefficients (slope and intercept) from the linear models were then used to calculate the missing values. *PPFD_IN* was filled based on global radiation (*SW_IN*), by multiplying *SW_IN* by the average ratio between *PPFD_IN* and *SW_IN* for the available periods at the site. *P* was filled by multiplying the ERA5-Land data by the ratio between the station data and the re-analysis data, as in Vuichard and Papale (2015). Any inaccuracies resulting from this replacement did not introduce additional bias in the gap-filled flux time series because precipitation was not used for gap-filling. A quality flag was developed for the meteorological data: 0 indicates measured data; 1, interpolated data; 2, data filled using a nearby station as a reference; 3, data filled using ERA5-Land as a reference.

Gaps in the flux time series were filled using a double-step procedure, analogous to the approach applied in Winck et al. (2023). Short gaps were filled using the Marginal Distribution Sampling method (Reichstein et al., 2005) with the online version of the REddyProc package (Wutzler et al., 2018). Short gaps were considered by taking the filled data with quality flags of 0 (original measured data) or 1 (highly reliable filled data). Subsequently, the remaining gaps (flagged with 2 or 3 in REddyProc) were filled using a machine learning (ML) tool based on the Extreme-Gradient-Boosting (XGBoost) algorithm (Chen and Guestrin, 2016). The code was adapted from Vekuri et al. (2023) to include *H*, *LE* and *FC*. The predictor variables of the model were the previously filled *TA*, *VPD*, *SW_IN*, *WS* and *WD*. The inclusion of *WD* followed the recommendation of Richardson et al. (2006) to account for site heterogeneity, as different land covers depending on wind sectors can contribute to flux variability. A quality flag was developed for the flux variables: 0 for measured data, 1 for data filled with REddyProc, and 2 for data filled with XGBoost. There were two very long gaps, one for AF3 during summer 2023 (mid-July until beginning of October) and another for AF1 during winter 2023/24 (beginning of December 2023 until beginning of March 2024), besides gaps of few days duration. Such long gaps would introduce significant uncertainty into any gap-filling method, so the analysis only considered measured and gap-filled data for gaps not exceeding two weeks duration.

The evaluation of the gap-filled fluxes with XGBoost was performed by splitting the initial dataset into 80 % training data and 20 % test data. The root mean squared error (RMSE) between modeled and measured data, for the test dataset, was taken as the error in the individual 30-minute flux value (Table 1).

Table 1. Root mean squared error (RMSE) of modeled and measured data, for *FC*, *LE* and *H*, for the four stations used in this study.

	AF1	AF2	AF3	OC
FC (μmol $m^{-2}s^{-1}$)	3.1	3.2	3.4	3.2
LE (W m^{-2})	24.2	25.0	20.2	26.5
H (W m^{-2})	14.7	13.5	13.2	14.4

2.3.4 Footprint calculation.

A footprint climatology was calculated for all stations, for five different periods considered in the study: (i) growing season 2023: from March to 13 July 2023, with the latter being the harvest date of rapeseed; (ii) harvest period 2023: from 13 July to 22 September 2023, with the latter being the harvest date of corn; (iii) winter 2023/24: from 22 September 2023 to 1 March 2024; (iv) growing season 2024, from 1 March to 15 July 2024, with the latter being the harvest date of the rapeseed; and (v) harvest period 2024, from 15 July to 19 September 2024. The footprint climatology was calculated using the Python version of the model by Kljun et al. (2015).

The input data for the footprint model included non gap-filled wind data (*WS*, m s^{-1}, and WD, °), roughness length (z_0, m), *USTAR*, Obukhov length (L, m), the standard deviation of lateral wind speed (*V_SIGMA*, m s^{-1}), boundary layer height (*BLH*, obtained from ERA5, Hersbach et al. 2023), measurement height (z_m, m) and displacement height (d_h, m). Daytime and nighttime values were used for the footprint modeling. z_0 and d_h were estimated from the aerodynamic canopy height (h_a, m), which was calculated under near-neutral conditions (stability parameter $|(z\text{-}d)/L| \leq 0.1$) using the procedure described by Chu et al. (2018). Complete time series of h_a were estimated by calculating the running mean of h_a for eight different wind sectors of 45° each, using a running mean of 100 30-min intervals. This procedure is described in more detail in van Ramshorst et al. (in prep.). This procedure allowed for a more comprehensive representation of the effects of a varying canopy roughness and is therefore more precise than using a single value to represent the average canopy height for the entire site at each time step. d_h and z_0 were calculated as 0.6 and 0.1 times the aerodynamic canopy height, respectively, following Chu et al. (2018). The mean values of d_h were 3.1 m at the AF and 0.6 m at the OC, while the mean values of z_0 were 0.5 at the AF and 0.1 m at the OC. A thorough discussion about the footprint model uncertainties can be found in Section 4.4.

2.4 Spatial and temporal variability of fluxes and turbulence parameters and effect size

To disentangle spatial and temporal variability of fluxes and turbulence parameters across the site, the data were classified in two ways. First, the data were aggregated into wind sectors of 45° each, similar to the sectors used for the planar fit division (see section 2.3), and separated into five time periods as described in the previous paragraph. Second, the data were grouped into one-week periods throughout the entire measurement campaign without division into wind sectors. For each classification, coefficients of spatial variation (CVs) were calculated and the variance was partitioned into temporal and spatial components. In this analysis used only measured data filtered according to the previously described criteria, not gap-filled data.

The *CVs* were defined as follows

$$CV_x = \overline{\left[\frac{< [x(t) - X(t)]^2 >^{\frac{1}{2}}}{X(t)} \right]} \quad (1)$$

based on Katul et al. (1999) and Oren et al. (2006). *X* is the spatial average of variable *x* across the three towers in the AF for the respective averaging time interval. Angular brackets (<>) denote the spatial averaging operator and the overbar denotes temporal average across all the individual time steps *t*. This formula was applied to *H*, *LE* and *FC*, and to the standard deviation of the vertical wind velocity (*W_SIGMA*, m s^{-1}), *USTAR* and *WS*. The coefficients of variation are dimensionless, normalized

by the spatial average of variable *x*, such that they can be compared between different variables. Lower limits were set for some of the variables, in order to avoid biasing the coefficients of variation by some very low fluxes in the denominator of Equation 1. These limits were 10 W m^{-2} for *H* and *LE*, $\pm$ 2 µmol m^{-2} s^{-1} for *FC*, and 0.5 m s^{-1} for *WS*.

The partitioning of the variance into temporal and spatial components was done as presented in Peltola et al. (2015) (Eq. 2 therein) based on Sun et al. (2010):

$$\sigma_{tot}^2 = \frac{m(n-1)}{m \cdot n - 1}\bar{\sigma_s^2} + \frac{n(m-1)}{m \cdot n - 1}\sigma_t^2(\xi) = \sigma_s + \sigma_t \quad (2)$$

with *m* the number of temporal data points, *n* the number of measurement locations, $\bar{\sigma_s^2}$ the time average of the spatial variance, and $\sigma_t^2(\xi)$ the temporal variance of the time series of spatial averages ξ. Consequently, the first term on the right hand side of the equation is equivalent to the spatial variance (σ_s), which includes as well the instrumental variance, while the second term is equivalent to the temporal variance (σ_t) (Peltola et al., 2015).

Furthermore, the effect size (*d*) was calculated in order to assess the statistical robustness of our distributed network, in accordance with the hypothesis of Hill et al. (2017) that the enhanced error observed in LC-EC setups can be counteracted by an improved statistical representativeness of the measurements, provided that the effect size is sufficiently large. In our case, with the three-tower network we calculated *d* across the three towers inside the AF and between the AF and the OC. *d* was calculated, following Hill et al. (2017), as

$$d = |\frac{f_1 - f_2}{\sigma}| \quad (3)$$

where f_1 is the flux from ecosystem 1, f_2 is the flux from ecosystem 2 and σ is the pooled standard deviation of data from both ecosystems. We used daily cumulative sums of gap-filled *FC* and *LE*. The value f_1 in Eq. 3 refers to the daily cumulative sums of *FC* (g C m^{-2}) or *LE* (W m^{-2}) at the AF, as an average across the three stations, while f_2 corresponds to the daily cumulative sum of *FC* or *LE* for AF1 or for OC, depending on the case under study. We calculated *d* for two different cases: (i) to test whether fluxes over AF (averaged across the three towers) differed significantly from fluxes over OC, to compare both ecosystems; and (ii) to test whether fluxes over AF differed significantly from those of the reference tower AF1, in order to compare the increase in statistical robustness of the distributed network to the hypothetical case in which only one station was installed at the AF. AF1 was selected as the reference tower because it was the longest-running tower on site, having been in operation since 2016. σ was calculated as in Hill et al. (2017)

$$\sigma = \sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}} \quad (4)$$

where σ_1 and σ_2 are the standard deviations of both datasets being compared, and n_1 and n_2 are the number of data points in each of the datasets. σ_1 and σ_2 were calculated as the error of the daily cumulative sums, from the individual 30-min error in the fluxes (see next section). Afterwards Eq. 4 was applied to get the error for the ensemble of stations being compared.

2.5 Uncertainty of the LC-EC setups

The uncertainty in *FC* and *LE* was considered by assigning an error to each 30-min flux value. This error was then propagated when aggregating data to daily cumulative sums for the effect size calculations. The error was considered differently for measured and gap-filled data. For measured data, the error in the 30-min *FC* and *LE* was obtained from the inter-comparison of LC-EC and conventional EC setups in the studies of Callejas-Rodelas et al. (2024) and van Ramshorst et al. (2024). The error was taken as the worst-case RMSE of all the comparisons between LC-EC and conventional EC setups, separately for *FC* and *LE*. The values were 3.1 µmol m^{-2} s^{-1} and 44.1 W m^{-2}, respectively, for *FC* and *LE*. This error was considered as a systematic deviation from the conventional EC setup, and not a random error.

For gap-filled data, the error was addressed differently for the two gap-filling steps. For the data filled with REddyProc, the error was defined as the standard deviation of the data points used for gap-filling (Wutzler et al., 2018), provided as an output from the REddyProc processing. In contrast, for the data filled with XGBoost, the individual error in the fluxes was assigned as the RMSE of the modelled data (Table 1). The uncertainty in a cumulative sum was then calculated using error propagation from the single 30-minute uncertainties to the daily sums.

3 Results

3.1 Meteorological conditions

SW_IN followed a seasonal cycle. The maximum magnitude was observed at the end of June 2023, with daily means above 300 W m^{-2}, followed by a decrease in radiation intensity. Minimum values close to 0 W m^{-2} were reached in winter, and then the intensity increased again until reaching similar maximum values in June 2024 (Fig. 2a). Total monthly values of *P* were large, especially from June to December 2023, and in July 2024, reaching up to 125 mm (Fig. 2d). There were some very dry months, with *P* sums lower than 20 mm, especially from March to June in 2024. Compared to the climatological averages (Table 2), all seasons during the measurement period were more rainy than the period 1981-2010, especially during summer and autumn of 2023, when the recorded precipitation was more than three times the reference value (272 mm vs. a reference value of 65 mm for summer 2023, and 218 mm vs. a reference value of 52 mm for autumn 2023). Spring 2024 was the only season slightly drier than the climatological reference, with a record of 30 mm of rain instead of 49 mm.

TA followed a seasonal cycle, with the lowest values in winter (daily means between 0 and 10 °C, with occasional lower values) and the highest values in July and August of both 2023 and 2024 (daily means around 20 °C). *TA* was slightly larger at the OC tower than at the other three AF towers during most of the campaign, with enhanced differences in summer and very small differences in winter. The mean *TA* during the campaign was 12.86 °C at the OC, while it was 12.49 °C at the AF. The three AF stations showed very similar *TA*. *TA* was higher in all seasons compared to the climatological averages (Table 2), except in spring 2023 in which both values were similar (9.1 °C). Summer 2023 and summer 2024 were slightly warmer (18.7 and 18.64 °C, respectively) than the reference value (17.4 °C). Autumn 2023, winter 2023/24 and spring 2024 were

Table 2. Measured and reference climatological averages of *TA* and *P* by seasons. Measured seasonal values were calculated as averages across all four stations at the site. Reference values were taken as the seasonal 1981-2010 climatological average from the German Weather Service (https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/, last accessed 25-09-2024) from the nearby station at Braunschweig airport (ID 662).

Season	Measured *TA* (°C)	Measured *P* (mm)	*TA* reference (°C)	*P* reference (mm)
Spring 2023	9.1	102.5	9.1	48.7
Summer 2023	18.7	272.3	17.4	65.0
Autumn 2023	11.9	218.5	9.8	52.0
Winter 2023/24	4.3	198.0	1.7	46.7
Spring 2024	11.8	30.1	9.1	48.7
Summer 2024	18.6	165.8	17.4	65.0

clearly warmer than the climatological averages, with 11.9, 4.3 and 11.8 °C vs. the reference values of 9.8, 1.7 and 9.1 °C, respectively. The absolute difference between measured and historical data was largest in winter.

VPD values also showed a marked seasonality (Fig. 2c). Values were very low in winter, between 0 and 0.2 kPa, and increased towards summer in both 2023 and 2024, reaching daily means between 1 and 1.5 kPa, while in autumn of 2023 *VPD* was lower with values of around 0.5 kPa. Comparing the four stations, the OC site experienced a larger *VPD* from July to October 2023, while during the rest of the campaign no significant differences were observed across the stations. The mean *VPD* was 0.41 kPa at the OC and 0.4 kPa at the AF as an average of the three stations. The differences between the three AF stations were very small.

3.2 Footprint climatology

All footprints exhibited larger contributions from the western side of the towers in all periods (growing season 2023, harvest period 2023, winter 2023/24, growing season 2024 and harvest period 2024), corresponding to the dominant wind direction at the site (Fig. 3). For all periods under consideration and for both 50 and 80 % footprint areas, the footprint of the OC tower was smaller than for the three AF towers, due to the lower measurement height. At the AF, footprints decreased from 2023 (Fig. 3a and b) to 2024 (Fig. 3d and e), likely due to the increase in canopy height of the trees. In the case of the OC, footprints were similar during the growing season of 2023 compared to the growing season of 2024 (Fig. 3a and d), and smaller during the harvest period of 2023 compared to the harvest period of 2024 (Fig. 3b and e). The 50 % footprint climatology contribution was concentrated in a small area around the stations, covering only the two crop fields at both sides of the stations, plus one or two tree rows in the case of the AF. There were small variations from season to season and a partial overlap between towers AF1 and AF2, and towers AF2 and AF3.

The 80 % footprint climatology contribution was larger, covering a larger portion of both AF and OC sites and therefore a surface with a larger heterogeneity due to the presence of more diverse crops and/or trees. The three stations at the AF exhibited partially overlapping footprints for the 80 % footprint climatology, with different sizes and degrees of similarity depending on

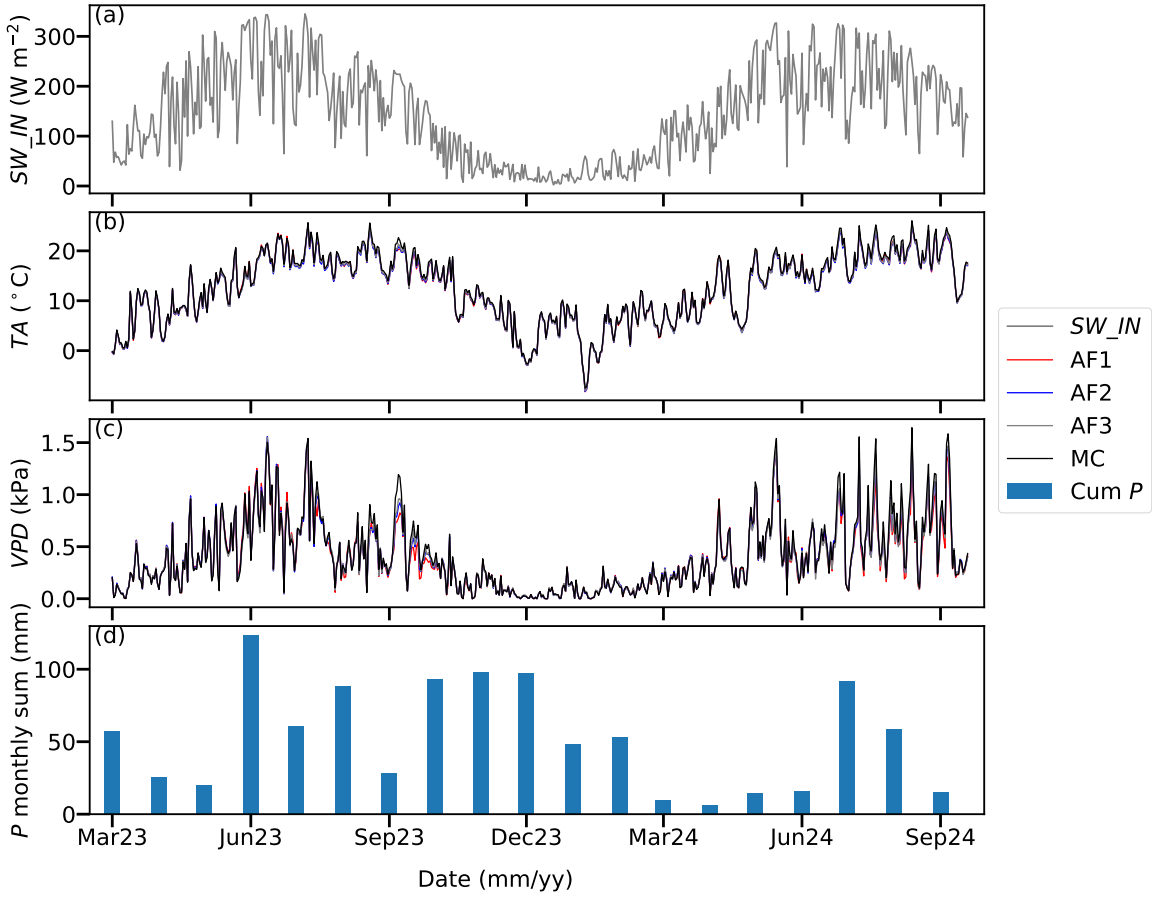

Figure 2. Time series of daily mean meteorological parameter and the cumulative sum of precipitation across the measurement campaign: (a) global radiation (*SW_IN*), (b) air temperature (*VPD*), (c) vapor pressure deficit (*VPD*) and (d) monthly sums of precipitation (*P*). *SW_IN* and *P* were considered as common to all the stations, because the size of the site is small enough to assume homogeneity in these parameters, whereas *TA* and *VPD* were plotted separately for all four stations. Data were filtered for outliers using lower and upper limits, gap-filled as detailed in Section 2.3.3, and then aggregated to daily values by taking the daily mean for *SW_IN*, *TA* and *VPD* and the daily sum for *P*.

the evaluated period. The most intense overlap occurred during the growing season of 2023 (Fig. 3a). The 80 % footprint of the three towers covered approximately four tree rows and four crop rows each. The three towers at the AF presented different footprint sizes, with the largest areas being covered by AF3, followed by AF2 and finally by AF1. This rank of magnitude was the same in all seasons. The footprint from the OC tower covered both the western and eastern fields around the tower, but the contribution was larger from the western part in all seasons. For all stations, there were some contributions to the 80 % footprints from the areas beyond the AF or the OC fields, especially remarkable in the case of AF3. However, the contributions of the areas outside the AF were expected to be negligible regarding the interpretation of the results.

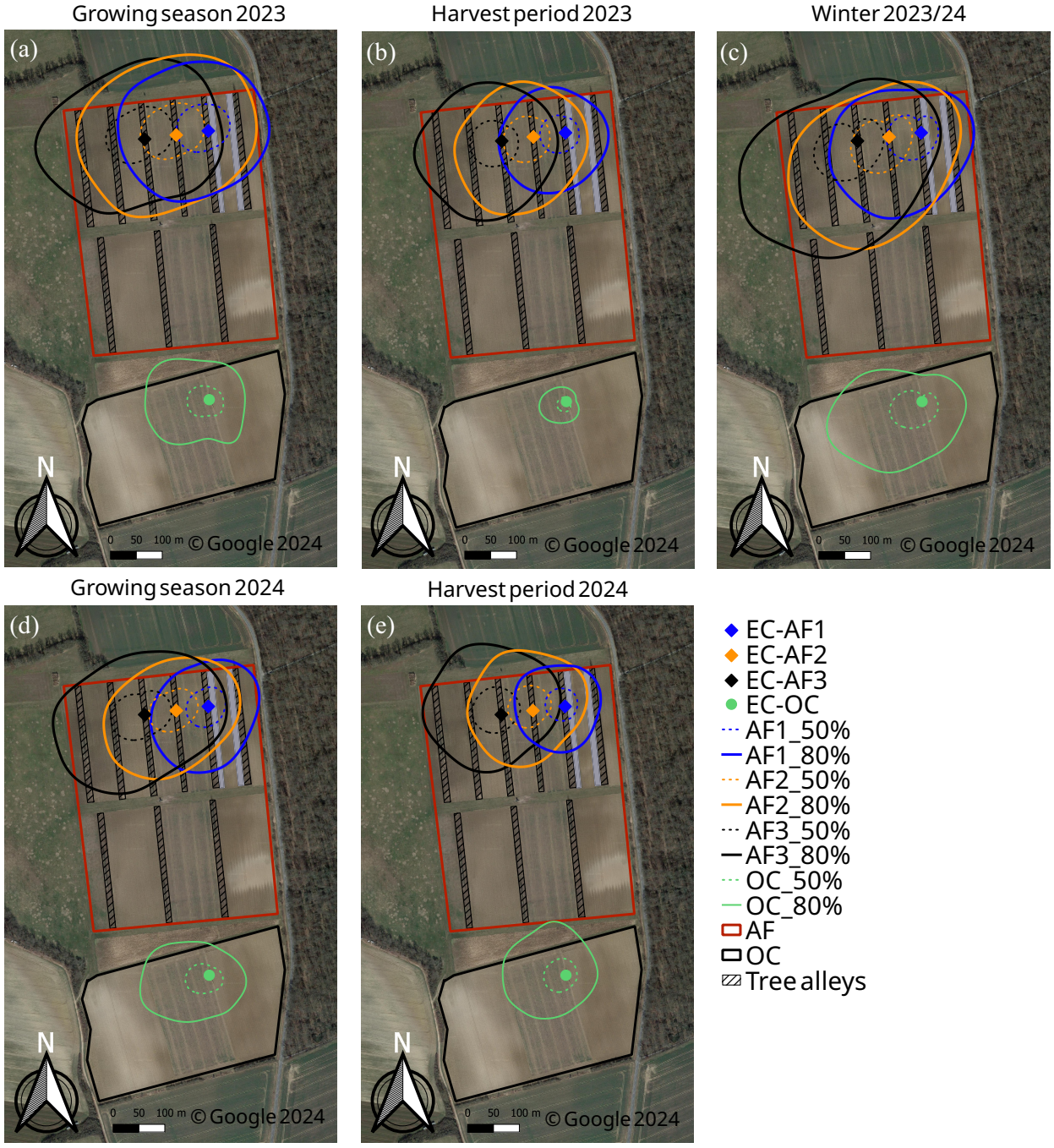

Figure 3. Footprint climatologies, calculated from the model of Kljun et al. (2015) as detailed in section 2.3.4, for the three towers at the AF (AF1, blue; AF2, orange; AF3, black) and the tower at the OC (green), divided into five different periods: growing season 2023 (a), harvest period 2023 (b), winter period 2023/24 (c), growing season 2024 (d) and harvest period 2024 (e). The lines plotted in the map represent the 80 % (solid line) and 50 % (dashed line) contributing areas to the footprint. The station locations are marked with diamonds for the AF stations and a circle for the OC station. Figure created with QGIS v. 3.22, aerial map by Google Satellite Maps. © Google 2024.

The analysis of the differences in land cover measured by the different stations revealed seasonal variations. Because all the AF stations covered some of the tree rows, specifically 3 or 4 in the case of AF1 and 4 to 6 in the case of AF2 and AF3, the description of the differences will focus on the different crops covered by the 80 % footprints. During the growing season of 2023 (Fig. 3a), the three stations at the AF covered all crops, whereby AF3 only covered a small portion of the barley field and the nettle fiber. During the harvest period in 2023 (Fig. 3b), AF2 covered all crops, including harvested rapeseed, while AF1 covered corn, barley (harvested at the end of August 2023) and nettle fiber; and AF3 covered rapeseed (harvested) and corn. In winter 2023/24 (Fig. 3c), all towers covered most of the crop fields, but these were mostly bare soil at this stage. During the growing season of 2024 (Fig. 3e), AF1 covered nettle fiber,rapeseed and barley; AF2 covered all crops; AF3 covered corn, barley and only a small portion of rapeseed and nettle fiber. Finally, during the harvest period of 2024, AF1 covered nettle fiber, rapeseed (already harvested) and barley (harvested three weeks after the beginning of this period); AF2 covered all crops; AF3 covered corn and barley. In all seasons, the OC tower covered mostly the western field (corn in 2023 and barley in 2024) and partially the eastern field (barley in 2023 and rapeseed in 2024).

3.3 Weekly sums of carbon and evapotranspiration

The weekly cumulative sums of *FC* (Fig. 4a) exhibited a marked seasonal behavior and similar variability across the four towers. The seasonal cycle was characterized by carbon uptake (negative values) during the growing season and carbon loss (positive values) during winter. The differences were smaller across the three AF towers, with AF1 and AF2 exhibiting more similar behavior. During the 2023 growing season, there was a strong uptake of around -30/-40 g C m^{-2} per week at all stations, from April to September 2023. This was interrupted by a short, three-week dry period at the end of May and the beginning of June of 2023 (DWD, 2024), during which the AF site turned into a weak carbon source (as measured by AF2) or to a weak carbon sink (as measured by AF1 and AF3). AF3 showed stronger uptake until mid-June. After that, OC showed the strongest uptake (-40 to -60 g C m^{-2} per week) for the rest of the growing season. After the rapeseed harvest on 13 July 2023, the weekly sums decreased in magnitude but remained substantial at AF1, AF2 and OC (AF3 was missing during this period). Around the barley harvest on 22 August 2023 the sums decreased notably. From October 2023 to March 2024, the values were positive and comparable across all stations, indicating a carbon release from the ecosystems. During the 2024 growing season, carbon uptake diminished compared to the 2023 growing season. The strongest uptake of around -25 g C m^{-2} per week occurred in July 2024. AF2 and OC showed the strongest uptake in June and July. However, after the rapeseed harvest on 15 July, the uptake decreased, and AF2 and OC changed to a carbon source. Meanwhile, AF1 and AF3 still showed negative values. After the barley harvest on 5 August 2024, the uptake at AF1 and AF3 decreased further, with AF1 changing to a carbon source. AF3 exhibited a CO_2 sequestration behavior until the end of the measurement period.

The weekly cumulative sums of *ET* (Fig. 4b) also exhibited a strong seasonality and similar variability across all stations. During the 2023 growing season, the weekly *ET* sums increased from April (around 10 mm per week) until the maximum values were reached in July, with a magnitude of 30 mm at AF2, AF3 and OC, and 40 mm at AF1. Afterwards there was a progressive reduction in *ET*, especially enhanced after the rapeseed harvest on 13 July 2023 and the corn harvest on 26 September 2023. AF1 showed the highest values until October 2023. After that, all stations showed low values, of around 5 mm per week,

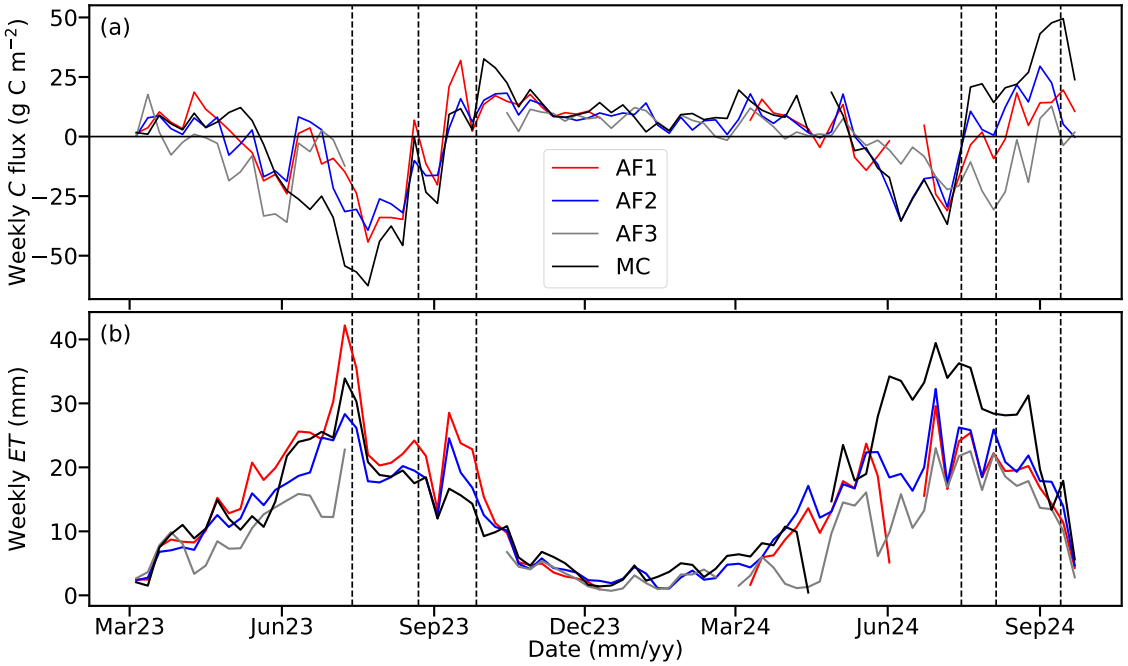

Figure 4. Weekly sums of the net ecosystem carbon exchange as a carbon (*C*) flux (a) and evapotranspiration (b, *ET*) measured at the four stations, across the measurement campaign. Sums were calculated from the gap-filled time series. Missing values correspond to gaps longer than 2 weeks, which were not considered in the analysis. The horizontal line in sub-plot (a) highlights the zero line, separating the uptake (negative fluxes) from the emission (positive fluxes). Vertical dashed lines represent, from left to right, the harvest dates of rapeseed (13 July 2023), barley (22 August 2023) and corn (26 September 2023) in 2023; and rapeseed (15 July 2024), barley (5 August 2024) and corn (13 September 2024). Due to the requirement of taking only gap-filled data for gaps up to two weeks of duration, there were some missing weeks for all stations and two very long gaps, in summer 2023 for AF3 and in winter 2023/24 for AF1.

coinciding with the winter period, until March 2024. During the 2024 growing season, *ET* increased progressively at all the stations until reaching the maximum values of 30 and 40 mm. This increase was interrupted only by a reduction in *ET* in June, more pronounced at the AF towers. After the peak of the growing season, *ET* reduced especially after the rapeseed harvest on 15 July 2024 and the barley on 5 August 2024. The highest values during the 2024 growing season and harvest period were found for the OC. AF3 exhibited lower values at the beginning of the growing season, but the three towers at the AF showed a good agreement from July on.

3.4 Coefficients of variation, spatial and temporal variance

3.4.1 Classification in wind direction bins

The *CVs* calculated at the half-hourly scale (Eq. 1) were the largest for *FC* in the eastern and southeastern wind sectors (60-180°) and all the evaluated periods, followed by the *CVs* of *LE* and *H* (Fig. 5). The largest values of the *CVs* of *FC* were

reached during the 2023 growing season, up to 8.4. The magnitude of the *CVs* of *FC* was comparable to the magnitude of the *CVs* of *LE* and *H* in the other wind sectors and periods, with values between 0.25 and 0.4. Notably, the *CVs* of *FC* were larger during the harvest period of 2024 than during the 2024 growing season. The *CVs* of *WS*, *USTAR* and *W_SIGMA* were low compared to the *CVs* of *FC*, *LE* and *H*. The lowest variability across wind sectors in all periods was found for *W_SIGMA*, followed by *USTAR* and *WS*, with *CV* values below 0.15 in most of the cases.

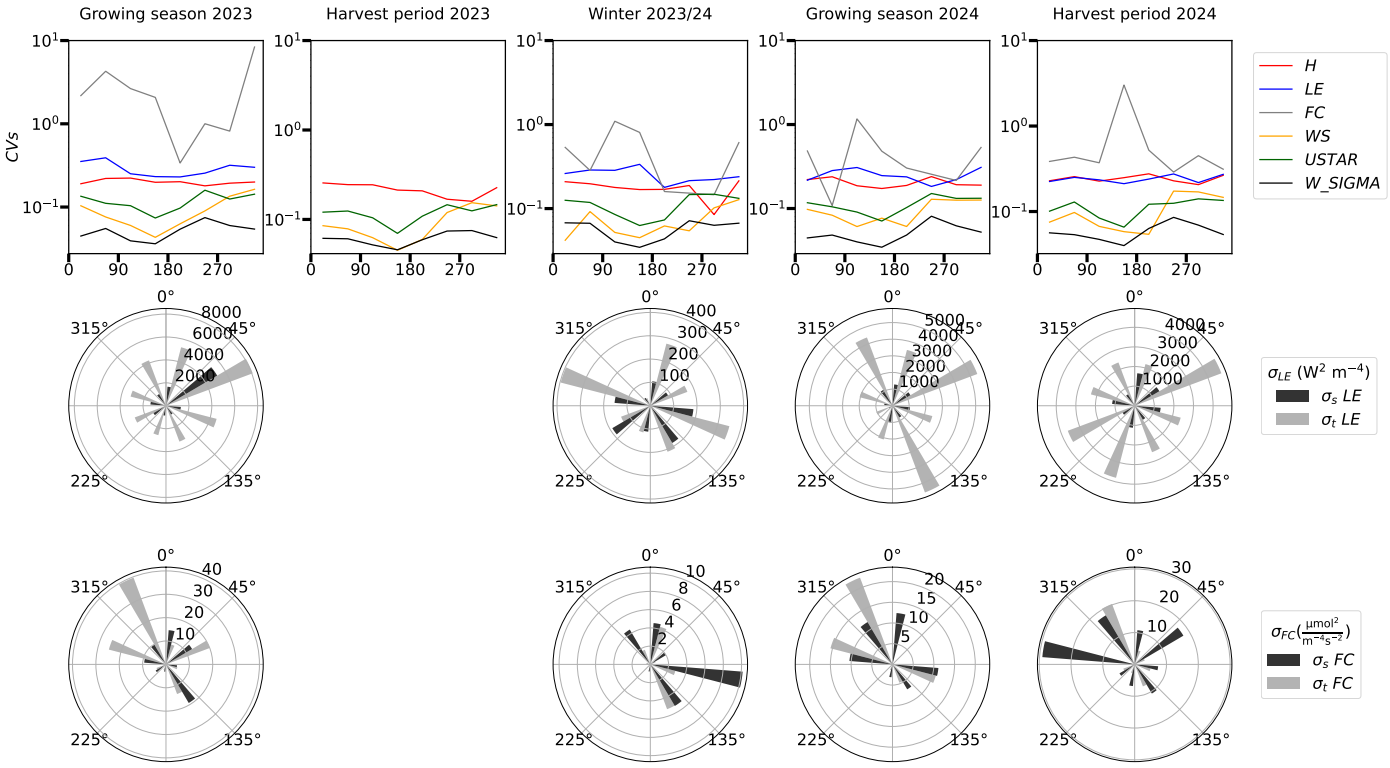

Figure 5. (Top row) Coefficients of variation (CVs), calculated after Oren et al. (2006), for *FC*, *LE* and *H*, *WS*, *USTAR*, and *W_SIGMA*; (mid row) spatial (σ_s *LE*) and temporal (σ_t *LE*) variance for *LE*; (bottom row) spatial (σ_s *FC*) and temporal (σ_t *FC*) variance for *FC*. Data were grouped in all cases by wind direction bins of 30° each and separated into the five analysis periods (growing season 2023, harvest period 2023, winter 2023/24, growing season 2024 and harvest period 2024) detailed in Section 2.3.4. Due to the two very long gaps in AF1 and AF3 (see Fig. 4), plus some shorter gaps, there were no data corresponding to the harvest period in 2023 for *FC* or *LE*, therefore the sectorial plots for the variance partition are missing. Note that in the first row, due to the large magnitude of some of the *CVs* of *FC*, the variability in the lines corresponding to the other variables is more difficult to visualize. Note that the y-axis is in logarithmic scale in the CV plots, to facilitate visualization. Note also that the scale is different in the circular plots, depending on the magnitude of what is represented in each season. No gap-filled data were used to create this plot.

Both for *FC* and *LE*, both variance values were larger during the growing season and the harvest period in both years than during winter, due to the larger magnitude of fluxes. Due to the scope of this analysis, it is important to remark in which wind

sectors σ_s was larger than σ_t. Looking first at *LE* (Fig. 5, mid row) σ_s was larger than σ_t only in the sectors 225-270° and 315-360° during the winter 2023/24. For all other wind sectors and periods, σ_s was lower than σ_t.

Regarding *FC* (Fig. 5, bottom row), the picture was different compared to *LE*, with a higher relevance of the spatial component of the variance. During the 2023 growing season σ_s was larger than σ_t in the northeastern sector (0-45°) and the southern half (90-270°). During winter 2023/24, σ_s was larger than σ_t in all wind sectors. During the 2024 growing season, σ_s was larger than σ_t in the eastern and southern sectors (0-270°). Finally, during the 2024 harvest period, σ_s was larger than σ_t in all sectors except in the Northwest (315-360°).

3.4.2 Classification in weekly intervals

The weekly *CVs* across the measurement campaign were largest for *FC*, with a large difference to the other evaluated variables (Fig. 6a). The difference was especially remarkable during winter and from March to May in both years 2023 and 2024. During most weeks, the *CVs* of *FC* ranged between 0.2 and 4.0, but reached high values of around 30 in some specific times of the growing season in both years and during winter. The *CVs* of *FC* were much larger than the *CVs* of *LE* and *H*, while in the summer months (after June) and the harvest period in both 2023 and 2024 the *CVs* of *FC* and *LE* were similar, with values between 0.2 and 0.5, closely followed by the *CVs* of *H*. Throughout the entire campaign, the *CVs* of *USTAR*, and *W_SIGMA* were much lower than for *H*, *LE* and *FC*, similar as shown in Figure 5, with values below 0.2 across the entire period. However, the *CVs* of *WS* were similar to those of *H* during the growing season and the 2023 harvest period. After summer 2023 the *CVs* of *WS* reduced their magnitude. The *CVs* of *USTAR*, and *W_SIGMA* were the lowest and did not change much during the campaign. In general, the harvest events did not clearly affect the variation of *CVs* for all variables.

With regards to partitioning the variance into its temporal and spatial components, σ_t was higher than σ_s for both *LE* and *FC* (Fig. 6b and 6c) during the summer months in both year. During winter and the months of March and April, both variance components were of similar magnitude for *LE* and *FC*. The highest variance (for both components) was observed during the end of the growing season in both years and during the harvest period in 2024, while the lowest occurred in winter time. The effect of harvest events in 2024 was shown by a reduction in the difference between σ_t and σ_s compared to previous summer months and a reduction in the variance magnitude (Fig. 6b).

3.5 Effect size and statistical representativeness of the three-towers network

The effect size (*d*) values were larger in the case of the comparison of *LE* sums than for the comparison of *FC* sums (Fig. 7). The values calculated using only the random error as the error in the measured data (Fig. 7a) were larger than the values calculated inserting random plus systematic error as the error in the measured data (Fig. 7b). This is a direct consequence of the inclusion of a larger denominator in Equation 3.

In regard to effect size, *d* values were lower in 2023 than in 2024, for *FC* and *LE* and in both error cases being considered. For *FC*, the values of $d_{\text{AF-MC},FC}$ were larger than the values of $d_{\text{AF},FC}$ in both years, and increased at the end of the growing season and during the harvest period in 2024. In the case of *LE*, the values of $d_{\text{AF-MC},LE}$ were lower than the values of $d_{\text{AF},LE}$ in 2023, but larger in 2024. The largest values of *d* were attained during July, August and September of 2024 for *LE* (magnitudes

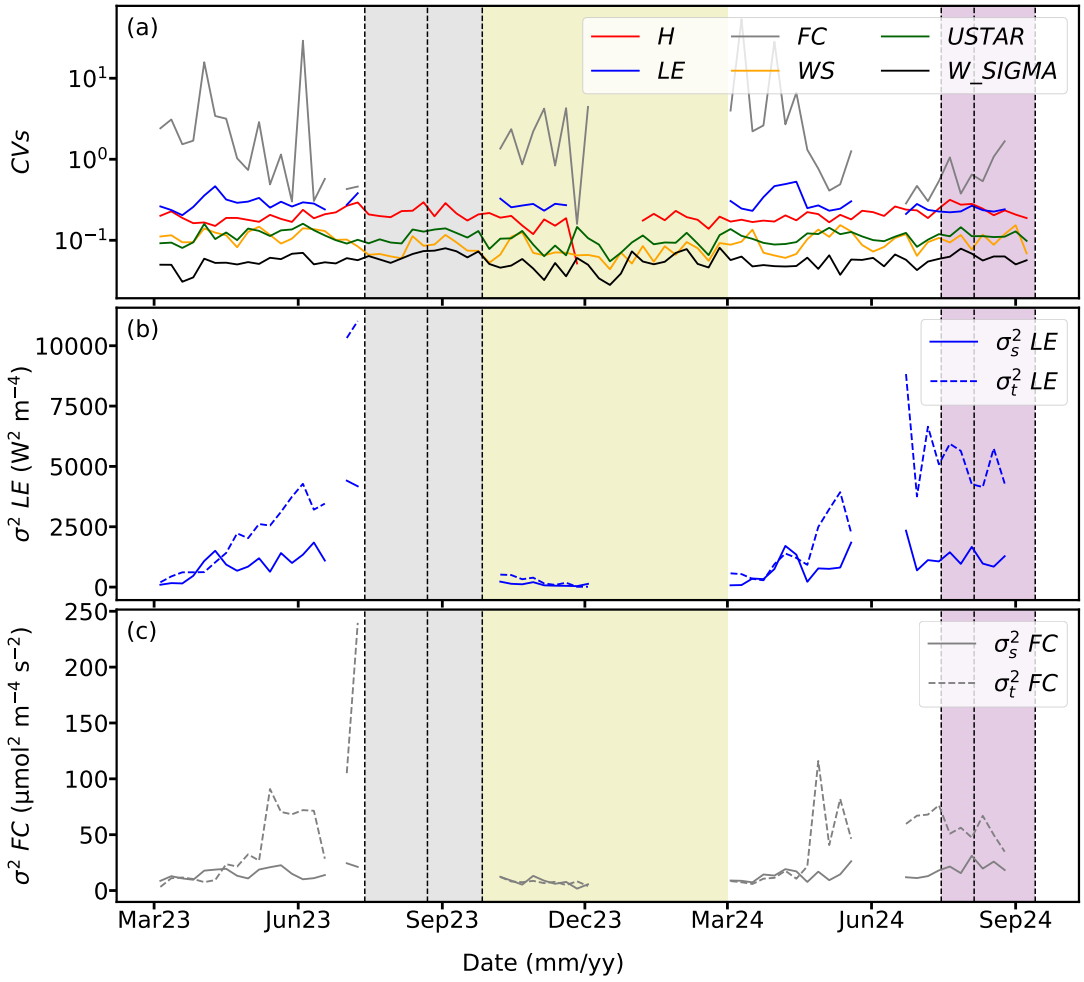

Figure 6. (a) Coefficients of variation (CVs), calculated after Oren et al. (2006), for *FC*, *LE* and *H*, $\bar{u}$, *USTAR*, and *W_SIGMA* (logarithmic scale); (b) spatial (σ_s *LE*) and temporal (σ_t *LE*) variance for *LE*; (c) spatial (σ_s *FC*) and temporal (σ_t *FC*) variance for *FC*. The plotted values are weekly means calculated at 30-min temporal resolution from the flux time series.Vertical dashed lines represent, from left to right, the harvest dates of the crops in 2023, for rapeseed (13 July 2023), barley (22 August 2023) and corn (26 September 2023); and in 2024, for rapeseed (15 July 2024), barley (5 August 2024) and corn (13 September 2024). Dashed areas correspond to the 2023 harvest period (grey), the winter period (yellow) and the 2024 harvest period (purple), for a better comparison with Figure 5. Due to the two very long gaps in AF1 and AF3 (see Fig. 4), plus some shorter gaps, there were no data corresponding to the harvest period in 2023 for *FC* or *LE* and only few weeks of data in the winter period. Note the logarithmic scale in panel (a), introduced due to the large magnitude of some of the *CVs* of *FC* for visualization purposes. No gap-filled data were used to create this plot.

up to 28), while in the case of *FC* values were largest at the end of the growing season in 2023 (magnitudes up to 12). The values of *d* for *LE* were larger than for *FC* in all periods except for the end of the growing season of 2023, in the case of

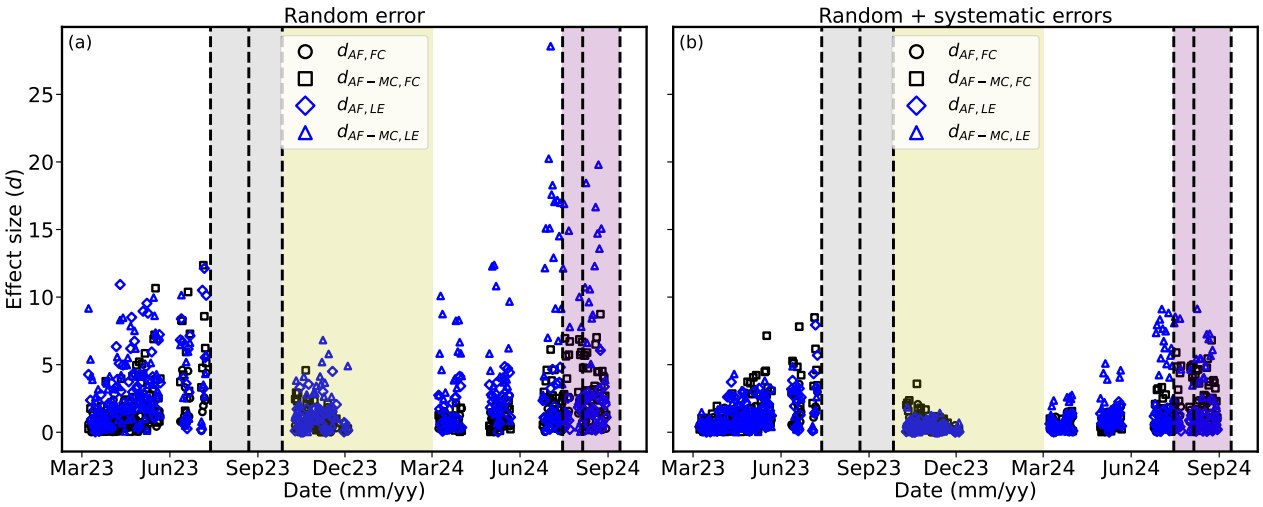

Figure 7. Time series of the effect size (d) for *FC* and *LE*, using as the error in the measured data the random error (a) or the sum of random and systematic error (b). d was calculated according to Eq. 3, based on the daily sums of *FC* and *LE*. Time series of *FC* and *LE* had been filtered and gap-filled as described in section 2.3.3, and gaps with a duration over two weeks were excluded from the analysis. Black circles represent the comparison between AF1 and the average of the three stations at the AF (AF1, AF2 and AF3) for the *FC*. Black squares represent the comparison between the average of the three stations at the AF (AF1, AF2 and AF3) and the OC station for *FC*. Blue diamonds represent the comparison between AF1 and the average of the three stations at the AF (AF1, AF2 and AF3) for *LE*. Blue triangles represent the comparison between the average of the three stations at the AF (AF1, AF2 and AF3) and the OC station for *LE*. Vertical dashed lines represent, from left to right, the harvest dates of the crops in 2023, for rapeseed (13 July 2023), barley (22 August 2023) and corn (26 September 2023); and in 2024, for rapeseed (15 July 2024), barley (5 August 2024) and corn (13 September 2024). Dashed areas correspond to the 2023 harvest period (grey), the winter period (yellow) and the 2024 harvest period (purple), as in Figure 6.

considering only random error (Fig. 7a). In the case of considering random and systematic errors (Fig. 7b), d values were larger for *FC* in 2023 and for *LE* in 2024.

4 Discussion

4.1 Spatial and temporal variability of *FC* and *LE* above the AF system

Using three distributed EC stations over the same AF system, a small spatial variability in meteorological parameters was found, but the spatial variability in CO_2 and energy fluxes was larger. Several rows of trees perpendicular to the main wind direction may potentially influence microclimatic conditions across the AF, compared to open croplands (Kanzler et al., 2019), but this AF site (17.3 ha) is smaller than the median farm size (29.4 ha) in Lower Saxony (Jänicke et al., 2022) and the meteorological variables were measured at the AF stations located within the tree strips. These two factors can explain the low variability in meteorological parameters. Therefore, the observed variability in *FC* and *LE* should not be attributed to the

meteorological drivers, but rather to differences in the footprint areas of the three stations. The footprint climatology of the stations partially overlapped (Fig. 3), but the most intense flux contributions originated from a small area around the towers. Differences in crop development and management practices could explain most of the variability of the observed fluxes across the three towers throughout the campaign, because of the different crops sown between the tree strips (spatial variability) and the different crop distribution from 2023 to 2024 (temporal variability) (Fig. 1).

The higher spatial variability in turbulent fluxes compared to other turbulence and wind parameters (Fig. 5), especially for *FC* and *LE*, was found also in the studies of Katul et al. (1999) and Oren et al. (2006). This can be explained by the complex nature of sources and sinks for CO_2 and H_2O fluxes (Katul et al., 1999) and the effects of landscape heterogeneity (Bou-Zeid et al., 2020). The explanation for the spatial variability in the fluxes is the land cover attribution thanks to the footprint modeling, however other effects of the heterogeneity were not studied.

The larger *CVs* of *FC* at the eastern wind sectors (Fig. 5) during all evaluated periods relate directly to differences in footprint climatology, because the footprints differed most at the eastern side of the three AF stations, especially for the 50 % footprint climatology (Section 3.2). The harvest events in 2024 did not seem to have a big impact on the *CVs* (Figs. 5 and 6a), but they slightly reduced the variance magnitude (Fig. 6b and c).

The larger temporal variance, compared to spatial variance, for both *FC* and *LE*, could be explained by the seasonal and diurnal flux variability, which was more relevant than spatial variability (see Sections 3.4.1 and 3.4.2). Nevertheless, σ_t was similar to σ_s in winter for both *LE* and *FC*, which can be attributed to the dormant state of the ecosystem, leading to small diurnal variations and consequently small temporal variations. In summer 2024, for *LE*, σ_s was similar to σ_t, due to the lower area overlap caused by smaller footprints (Fig. 3) compared to the 2023 growing season. This was also due to the absence of a fully developed crop in the eastern part of the field because of the poor rapeseed growth during this season. This resulted in weaker *LE*, especially at tower AF1, and led to a lower spatial variation.

In comparison to similar approaches in the literature, Peltola et al. (2015) found a paired temporal and spatial variability in CH_4 fluxes measured at three different heights on a tall EC tower and two additional EC stations over an agricultural landscape. Hollinger et al. (2004) measured fluxes using two towers with non-overlapping footprints in a forest and found that the temporal variability was larger; however, the spatial disagreement in *FC* was not negligible, despite the apparent homogeneity of the studied ecosystem. Rannik et al. (2006) also compared *FC* measured from two nearby towers over the same ecosystem, with partially overlapping footprints, and found relevant systematic errors in the daytime fluxes attributed to the variability in the turbulent flow field caused by the complexity of the terrain. These systematic differences were important for attributing long-term uncertainties in ecosystem carbon uptake, as would be the case in the complex AF site of the present study. Davis et al. (2010) investigated heterogeneity in *FC* above an arable land and demonstrated the significant impact of spatial heterogeneity on annual carbon balances. Furthermore, Soegaard (2003) quantified the annual carbon budget of an agricultural landscape by combining footprint-weighted fluxes and spatial variability in different crops, demonstrating the large potential of spatial heterogeneity to bias annual flux estimates. In the present study, the influence of different land covers around the towers was detectable for both *FC* and *LE*, except during the winter period. However, the differences were smaller than expected for crops with clearly different seasonality. As other effects of heterogeneity on flux measurements cannot be captured with this setup,

a first explanation could be the partially overlapping footprints and the buffering effect caused by the presence of the trees. As trees were assumed to behave similarly across the AF, their similar CO_2 and water fluxes attenuated the potentially larger differences in turbulent fluxes that would be expected among the crops without trees.

The observed variations in the weekly cumulative sums of *FC* and *ET* across the campaign (Fig. 4) can be attributed to the developmental and management differences among the crops cultivated around the stations, provided that the trees were growing similarly across the entire AF site. These differences can be directly connected to the previously explained behavior of the *CVs* and partitioning of the variance. Spatially replicated experiments demonstrated the potential to more accurately estimate the uncertainty in turbulent fluxes, e.g. by using non overlapping paired towers as in Hollinger and Richardson (2005), but this could not be applied in the present study due to the overlapping footprints. Conversely, the deployment of three towers provided a more comprehensive dataset compared to the single tower approach. However, the choice of the towers location in the present study might not have been optimal (Chen et al., 2011), since footprints were partially overlapping (Fig. 1 and 3). This was due to logistic constraints that precluded the selection of any other location within the AF site, such as in the southernmost part of the field. On the other hand, the purpose of the study was to investigate small-scale variability in the highly heterogeneous AF, a goal that was generally accomplished.

Specifically, the earlier development of rapeseed in 2023 led to an initial carbon uptake at AF3, because the main footprint covered rapeseed (Fig. 3a). This matched the larger *CVs* of *FC* in the eastern side of the field (Fig. 5), and during March and April 2023 (Fig. 6a). However, the earlier growth of rapeseed did not increase *ET* in AF3 (Fig. 4b), leading to comparable *CVs* of *LE* for all wind sectors (Fig. 5). This is because rapeseed can maintain a relatively large carbon uptake while using limited water resources (Najibnia et al., 2014). The subsequent development of corn and barley led to similar weekly uptakes of carbon at AF1 and AF2, but a larger *ET* at AF1, leading to a decrease in the *CVs* of *FC* and a modest increase in the *CVs* of *LE*. Besides the partially overlapping footprint (Fig. 3), another reason is a different water use efficiency among barley and corn, being lower for barley and therefore explaining similar carbon uptake as corn at a higher *ET* (see e.g. Pohanková et al., 2018). After the short drought in May-June, which affected all three stations by reducing both carbon uptake and *ET*, weekly carbon uptakes of AF1 and AF2 and weekly *ET* sums were larger than for AF3 until the harvest period. This can be attributed to corn and barley being less present in the footprint area of AF3 (Fig. 3a). Corn and barley exhibited a more intense physiological activity, immersed in the growing season, while rapeseed was likely at its maturity stage.

The rapeseed harvest in 2023 had a negligible effect on the carbon uptake of AF1 and AF2 much, but seemed to have an effect on *ET*, which reduced for both stations. This can be attributed to a period of several precipitation events, low *TA* and *VPD* (Fig. 2) which reduced both physiological activity and atmospheric water demand. The barley and corn harvests reduced the carbon uptake and *ET*. Especially the corn harvest had a large impact because it was the main crop in the footprints of AF1 and AF2 (Fig. 3b). After the harvest period, the slightly larger difference between the three stations may be an effect of the larger gap-filling uncertainty due to the longer gaps and agrees with an enhanced spatial variance compared to the temporal variance (Fig. 6b and c).

In 2024, the very dry spring (Table 2) did not affect weekly sums of *ET*, but reduced the magnitude of weekly sums of *FC* compared to 2023. In 2024, there was no earlier development of the rapeseed as it occurred in 2023, due to the very wet

winter conditions. The variability in *ET* was larger than in 2023 due to less overlapping footprints and due to the difference in rapeseed growth (Fig. 3d). The larger carbon uptake at AF2 as well as larger *ET* (Fig. 4) during the whole growing season of 2024 can be explained by the influence of barley and partially corn, while AF1 detected only part of the barley field and the non-well developed rapeseed (Fig. 3c). Carbon uptake and *ET* were smaller at AF3 because corn developed later, but reached similar values as AF1 once corn started to grow. After the rapeseed harvest, AF1 and AF2 reduced both their carbon uptake and *ET* release, with AF2 turning into a carbon source. This was explained not by the footprint of AF2 in the rapeseed field (Fig. 3e), but rather by the mature barley and the strong ecosystem respiration under wet conditions. Carbon uptake and *ET* release at AF3, on the other hand, did not detect the effect of the rapeseed harvest, because AF3 was not measuring the corresponding portion of the field (Fig. 3e). AF3 kept a large weekly carbon uptake and similar *ET* due to the presence of the corn in its footprint area (Fig. 3). Afterwards, the barley harvest reduced the uptake of AF1, turning it onto a carbon source, and of AF3, as well as *ET* due to the footprint covered by both stations (Fig. 3e). Carbon uptake progressively reduced until it eventually turned to emissions around the corn harvest, which was the main crop in the footprint area of AF3.

4.2 Differences in *FC* and *ET* between AF and OC systems

The AF site had typically lower air temperature and higher *RH* than the OC (Fig. 2), because the trees at the AF act as a buffer to keep cooler air temperatures and cooler soil resulting in a larger RH. This is pointed out in a review by Quandt et al. (2023). The authors stated that during drought events and under drier and warmer climatic conditions, as projected in future climate scenarios, trees might potentially help in sustaining cooler temperatures and the air more humid.

Carbon uptake and *ET* release were enhanced at the AF at the beginning of the 2023 growing season, because of the earlier development of the trees and the rapeseed, both present in the footprint of all three AF stations (Fig. 1 and 3a), while the OC station was measuring mostly corn (Fig. 3a). Corn is a crop with a later development compared to barley or rapeseed (Lokupitiya et al., 2009; Soegaard, 2003), but is typically very productive (Hollinger et al., 2005; Lokupitiya et al., 2016). Therefore, carbon uptake was larger at the OC during most of the 2023 growing season after corn started to grow, later than rapeseed and barley. A similar *ET* between OC and AF2 (Fig. 4b) indicated a larger water use efficiency at the OC. In our study, the short dry period in May/June 2023 took place before corn reached its peak growth stage, while rapeseed and barley were in a more advanced stage and were more affected by the dry conditions. In general, the whole campaign took place during very wet conditions. This might have increased the ecosystem respiration because it led to more soil organic matter decomposition driven by larger litter amounts at the AF. This, together with a larger respiration from the trees, can explain why AF2, even though it was surrounded by corn, did not take up as much carbon as the other towers in the AF system.

During the 2023 harvest period, the footprint of the OC station was limited to corn, not rapeseed (Fig. 3b). Corn continued to grow in July and August of 2023 at the OC, which explains why at the OC a very large carbon uptake and *ET* release was observed, while AF2 and AF1 showed reduced fluxes. In winter, the ecosystems were dormant which explains the small differences between AF and OC. However, fluxes were very small in magnitude and it was difficult to observe differences between sites.

During the 2024 growing season, carbon uptake at the OC was similar to the AF, but *ET* was larger at the OC, opposite to what occurred in 2023. This could be explained by barley grown in the main footprint area of the OC (Fig. 3d), as well as a portion of the rapeseed field, which did not grow well this year. Barley is a crop with less intense physiological activity and lower water use efficiency than corn (Pohanková et al., 2018). This explains the smaller differences to the AF stations in C uptake and a much larger *ET*. Also, the meteorological conditions were very wet in winter with a dry spring. During the harvest period in 2024 the carbon uptake and *ET* reduced more sharply at the OC than at the AF, after the rapeseed harvest, because of its partially contributing footprint (Fig. 3e). The reduction was more pronounced after the barley harvest, which contributed the most to the footprint covered by the station.

4.3 Effect size and spatial representativeness of the distributed network

The effect size *d* is a measure of the relative difference of two variables for two different populations (in this case two ecosystems or towers within an ecosystem) with respect to the pooled standard deviation of the two populations. The interpretation of the calculated values was done according to Figure 3 in the paper by Hill et al. (2017), where the number of EC replicates over an ecosystem or for comparing two ecosystems was estimated based on the desired statistical power (from 0 to 1) and the effect size value. The statistical power related to the confidence in the accuracy of the measurements, such that a value of 1 means we can be 100 % certain about the measured differences.

In the case of comparing the AF, similar values for both *LE* and *FC* were attained, mostly between 0 and 5. Values of 5 meant that with three towers a statistical power between 0.7 and 0.95 was achieved, however with values close to 0, the statistical power dropped dramatically so no confidence in the accuracy of the differences could be drawn. In the case of comparing AF-MC, *d* values were larger than for the comparison of the AF, which meant that a larger statistical power was achieved because the daily sums were larger than the pooled uncertainty. Values larger than 2 or 3 in many cases, reaching up to 15 or 20, meant a statistical power above 0.975, therefore a very large confidence in the daily sums. Furthermore, d_{LE} was larger than d_{FC}, meaning that the statistical confidence was larger for *LE*. When using random and systematic errors as the errors attributed to measured data (Fig. 7b), *d* values were much lower. This matches the interpretation of Hill et al. (2017): if the EC systems are too uncertain, the number of systems needed to achieve a large statistical power (above 0.9) increases exponentially. If the LC-EC setups used in this study would be a lot less accurate, e.g. with two times more systematic error compared to conventional EC, the effect size values would be too low so no certainty about the data could be ensured, unless the number of towers would increase according to counteract the loss of accuracy.

Several studies addressed the spatial representativeness of fluxes and footprint climatology. These studies focused on studying RE (Hollinger and Richardson, 2005), on separating ecosystem structure and sampling errors in the spatial variability of fluxes (Oren et al., 2006), on disentangling temporal and spatial variability of fluxes using a single tower approach and footprint modeling (Levy et al., 2020; Soegaard, 2003), on the representativeness of single point measurements at the pixel scale for regional to global scale models (Chasmer et al., 2009; Chen et al., 2009; Wang et al., 2016; Ran et al., 2016), or on studying the effect of diverse meteorological conditions in the footprint climatology and canopy structure (Abdaki et al., 2024). To the best of our knowledge, the study of Cunliffe et al. (2022) was the only one that deployed several LC-EC setups

similar to those used in our study and one additional conventional EC setup to quantify the impact of landscape heterogeneity on turbulent fluxes. They studied a dryland site with very low flux magnitudes, which is different from our site. They obtained a useful agreement between different LC-EC and conventional EC setups and attributed the differences between setups to the ecosystem heterogeneity, covered by different bushes and grass species. However, a less detailed analysis on the spatial and temporal variability of the fluxes was performed.

In the EC community, EC replicates are uncommon (Hill et al., 2017; Stoy et al., 2023). Therefore, the effect size of either means or sums of fluxes is typically not estimated. Hill et al. (2017), as the first paper showing the potential of LC-EC setups in increasing spatial replication in EC studies, estimated the effect size by comparing the average carbon sequestration and the standard deviation of the cumulative sums, for ideal and non-ideal FLUXNET sites (Baldocchi, 2014). In the present study, the effect size was calculated similarly, but based on daily sums and pooled standard deviations (errors) of the 30-min time series. The concept in Hill et al. (2017) was different since measurement errors tend to decrease relative to the aggregation period when cumulative sums are calculated (Moncrieff et al., 1996). Their calculated standard deviation was based on the uncertainty in the cumulative sums of the half-hourly carbon fluxes, rather than on time series with a higher temporal resolution (30 minutes). These time series are commonly characterized by higher variability and a potentially lower effect size.

In general, there is still an ongoing discussion on how much the landscape heterogeneity affect balances of carbon and H_2O measured by single EC towers. The LC-EC setups could help to bridge the gap of low spatial replication across such heterogeneous sites by allowing the installation of multiple setups due to their reduced cost. This could be complementary to other methodologies developed, to understand the effect of spatial heterogeneity on fluxes measured from single towers, such as in Levy et al. (2020) or Griebel et al. (2016), or measured with several conventional EC setups (Soegaard, 2003; Hollinger et al., 2004; Katul et al., 1999; Oren et al., 2006).

4.4 Heterogeneity as a challenge to EC measurements and footprint modeling

As mentioned in the introduction, the heterogeneity in the surface properties of a certain ecosystem induces horizontal advection, secondary mesoscale circulations and non-equilibrium turbulence processes (Bou-Zeid et al., 2020). Horizontal advection at different spatial scales can distort flux measurements (Cuxart et al., 2016). Furthermore, the dynamics of the roughness sublayer (RSL), defined as the atmospheric layer influenced by the roughness elements and located below the inertial sublayer (Katul et al., 1999), can be modified by the wind barrier of trees in the AF (van Ramshorst et al., 2022). Upon a change in the underlying surface, an internal equilibrium layer (IEL, Brutsaert 1998) and an internal boundary layer (IBL, Garratt 1990) develop. Multiple IELs and IBLs can develop if there are multiple transitions in the surface, such as at the AF (Bou-Zeid et al., 2020). At the AF, the major change in the surface is represented by the tree rows (Markwitz, 2021). These rows create persistent waves that enhance the differences in the turbulence-related parameters *WS*, *USTAR*, and *W_SIGMA*, though these changes are less pronounced than flux variations. Furthermore, the classical tests of stationarity and equilibrium may fail if the EC station is placed above the IEL (Mahrt and Bou-Zeid, 2020), due to a disequilibrium between the mean flow, turbulence and the new surface (Bou-Zeid et al., 2020). Additionally, the complex canopy structure at the AF could lead to significant carbon and energy storage, particularly at the crop-tree interfaces and within the dense tree rows. These storage terms may influence

advection in the horizontal and vertical directions (Mammarella et al., 2007; Aubinet et al., 2010; Feigenwinter et al., 2008). These effects may affect the turbulence and flux measurements, however they could not be quantified with the current setup.

The footprint size and the overlap between footprints decreased between 2023 and 2024 due to tree growth (Fig. 3). Combined with changes in crop development and meteorological conditions, this increased the spatial components of the variance for *FC* and *LE* (Fig. 5). While the three towers at the AF had partially overlapping 80 % footprint climatology areas (Fig. 3), the main footprint contributions concentrated in the immediate areas around each tower (Kljun et al., 2002). Therefore, most of the flux variability can be attributed to land cover differences around the stations. The three-tower network helped disentangle the effects of management activities (e.g. crop harvest) and providied insights into small-scale features caused by the alternating structure of the AF. The division of the data into wind direction bins, as done in e.g. Kutsch et al. (2005), to address spatial variability in fluxes and turbulence parameters, as well as the spatial and temporal components of the variance, complemented the information provided by the footprint maps.

The footprint model used in the present study (Kljun et al., 2015) allowed to understand, at a basic level, where the source/sink areas of CO_2 and H_2O were located. Nevertheless, the parametrization of the footprint model does not consider the effects of spatial heterogeneity on the basic parameters roughness length and *USTAR*, nor how canopy heterogeneity influences wind speed and turbulence dynamics within the AF (Markwitz, 2021). Due to the structure of the AF, the footprint model likely overestimated the footprint area by attributing sources and sinks to areas that do not actually contribute to the flux. Additionally, footprint estimates are sensitive to the vertical distribution of sources and sinks within the canopy and to the time that air parcels spend within it (Launiainen et al., 2007; Prabha et al., 2008). This is likely happening at this AF site, due to the structure of the tree rows. Implementing the aerodynamic canopy height after Chu et al. (2018) helped to partially account for the heterogeneity of the AF site in the footprint modeling, but this procedure was also limited. More accurate footprint estimates could be obtained by combining flow dynamics and spatial structure information using Large Eddy Simulations (Markwitz, 2021; van Ramshorst et al., 2022), with a more advanced footprint modeling as described in e.g. Göckede et al. (2006) to account for the spatial heterogeneity in roughness length and *USTAR*. Additionally, aggregating the footprint climatology based on weighted footprints, as in Chen et al. (2009), would allow for a more detailed characterization of the sources and sinks of carbon and water vapor.

Furthermore, the sensor location bias, defined as the uncertainty caused by measuring at only one point above a heterogeneous site, depends on the stability conditions (Chen et al., 2011). Under more unstable conditions, the footprint size decreases and the location bias of each tower increases. This justifies the use of multiple EC towers to sample an entire ecosystem more effectively. A more detailed study of stability regimes, footprint size and spatial variability of fluxes would inform about this feature. However, this was not performed in this study due to the limited data availability and to the difficulty in gap-filling turbulence parameters needed to classify stability regimes, such as Obukhov length. With longer time series and more complete turbulence and footprint information, some of the previously detailed shortness of this study could be addressed.

4.5 Errors in *FC*, *LE* and *H*

Errors affecting flux calculations are difficult to disentangle because they propagate throughout the entire processing routine, from raw data measurements to final flux corrections. Therefore, the uncertainty in the measured fluxes from the LC-EC setups was assigned based on the random error and the previous inter-comparison studies of Callejas-Rodelas et al. (2024) and van Ramshorst et al. (2024), as detailed in Section 2.5. This procedure is similar to the approach applied in Peltola et al. (2015), where they used a previous instrument cross-comparison campaign (Peltola et al., 2014) to assign instrumental uncertainty to the setups they deployed. However, the uncertainty in the use of LC-EC, defined in relation to conventional EC, was obtained during a specific campaign and under specific site conditions with the same footprint area; therefore, there might be a bias in the LC-EC error attribution. Additionally, as explained in Section 2.5, the uncertainty in the gap-filled fluxes was calculated by assigning individual errors to the 30-min fluxes, which can then be propagated when performing the daily cumulative sums. This was detailed as an initial method for easily evaluating and propagating errors through cumulative sums when comparing new EC setups to conventional EC setups and calculating carbon or *ET* balances using gap-filled data. Other approaches, as described in e.g. Richardson and Hollinger (2007), could potentially be applicable to this dataset as well.

Including all gap-filled data, with some very long gaps, particularly affecting AF3, would have made the analysis more uncertain (Lucas-Moffat et al., 2022). Therefore, only measured data and gaps shorter than two weeks were used, although this did not permit a more complete spatial heterogeneity study throughout the campaign. The optimal solution for this study was to use a combination of REddyProc for short gaps and the XGBoost model for long gaps, similar to what was done in Winck et al. (2023). This solution allowed to assign individual errors to each 30-min flux, as explained in Section 2.5. Additionally, applying more strict filtering criteria, such as a higher *USTAR*-threshold or a lower quality flag, would provide higher-quality data but would also increase uncertainty due to the filling of more gaps.

We used *TA*, *SW_IN* and *VPD* as predictors for gap-filling, which are generally recognized as the main drivers of CO_2 and H_2O fluxes (Vekuri et al., 2023; Wutzler et al., 2018). *WS* was used because of its influence on turbulence development and the spatial information carried by eddies, especially above a very rough surface, such as the AF. Wind direction was selected to account for the spatial heterogeneity across the different measurement locations of the towers (Richardson et al., 2006). Other meteorological variables were either less relevant for the analysis, such as atmospheric pressure, or more complex to gap-fill, such as net radiation.

5 Conclusions

This study presents for the first time 1.5 years of measurements from a distributed network of three eddy covariance towers above a temperate heterogeneous agroforestry system, as well as a comparison to an adjacent, open cropland agricultural system. Using three eddy covariance stations allowed to capture the spatial and temporal variability across the site, which particularly affected carbon flux. The main differences were attributed to the different developmental stages of the crops across seasons, with larger disturbances in carbon flux and latent heat flux after harvest events. Due to the high degree of spatial heterogeneity, a broader footprint coverage was necessary to capture small-scale differences at the agroforestry. Furthermore,

binning the data by wind direction sectors and weeks provided a detailed picture of the temporal and spatial components of the variance and the coefficients of spatial variation. This was important because the differences between stations were small enough to be masked by a less resolved analysis.

Secondly, this study incorporated a complex gap-filling procedure that complemented previously published recommendations on working with lower-cost eddy covariance data. The datasets gathered during the campaign and the processing scheme added value to the data collection of the project from previous years, from several agroforestry and open cropland sites. Future research will address the contrast between different agroforestry and open cropland sites in more detail, using more years of data under a broader range of meteorological conditions.

Finally, the footprint coverage required to capture the spatial heterogeneity across the agroforestry, and within the agroforestry and open cropland, was improved by using lower-cost eddy covariance setups. We satisfactorily proved the hypothesis that the degree of uncertainty introduced by using slower-response gas analyzers for CO_2 and H_2O was counteracted by the better representation of all processes occurring within the agroforestry system. Therefore, we recommend installing multiple eddy covariance setups, including lower-cost setups, whenever the degree of heterogeneity of an ecosystem is large. Future studies could benefit from comparing overlapping and non-overlapping measurements in terms of footprint.

Code and data availability. Data corresponding to this publication, as well as the codes to analyze results and prepare the figures for this publication are available at Zenodo, https://doi.org/10.5281/zenodo.14855288 (Callejas-Rodelas et al., 2025).

Author contributions. JACR performed the measurements, data analysis and manuscript writing. AK and CM wrote the project proposal, contributed to data analysis and manuscript editing. IM, TV and OP contributed to data analysis and manuscript editing.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Disclaimer. TEXT

Acknowledgements. We wish to acknowledge the funding agencies for providing the necessary funds to perform this research, as well as the technical support in the field work received by Marek Peksa, Frank Tiedemann, Edgar Tunsch, Dietmar Fellert, and student assistants (Bioclimatology group) from the University of Göttingen. We wish to acknowledge as well the support from the team of the Micrometeorology Group at the University of Helsinki and from the Natural Resources Institute Finland (LUKE) in Helsinki.

Financial support. This research was supported by the German Federal Ministry of Education and Research (BMBF, project BonaRes, Module A, SIGNAL 031B1063A). This project also received funding from the European Unions´ Horizon 2020 research and innovation program under Grant Agreement No. 862695 EJP SOIL, the German Academic Exchange Service (DAAD), and the Reinhard-Süring-Foundation (RSS), affiliated to the German Weather Society and ICOS-Finland by University of Helsinki. Olli Peltola acknowledges Research Council of Finland for funding (grant no. 354298).

References

Abdaki, M., Sanchez-Azofeifa, A., Vargas, R., Ludwig, R., and Hamann, H. F.: Spatial and Temporal Variation of Three Eddy-Covariance Flux Footprints in a Tropical Dry Forest, Agricultural and Forest Meteorology, 345, 109 863, https://doi.org/10.1016/j.agrformet.2023.109863, 2024.

Aubinet, M., Feigenwinter, C., Heinesch, B., Bernhofer, C., Canepa, E., Lindroth, A., Montagnani, L., Rebmann, C., Sedlak, P., and Van Gorsel, E.: Direct Advection Measurements Do Not Help to Solve the Night-Time CO_2 Closure Problem: Evidence from Three Different Forests, Agricultural and Forest Meteorology, 150, 655–664, https://doi.org/10.1016/j.agrformet.2010.01.016, 2010.

Aubinet, M., Feigenwinter, C., Heinesch, B., Laffineur, Q., Papale, D., Reichstein, M., Rinne, J., and Van Gorsel, E.: Nighttime Flux Correction, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by Aubinet, M., Vesala, T., and Papale, D., Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2351-1, 2012.

Baldocchi, D.: Measuring Fluxes of Trace Gases and Energy between Ecosystems and the Atmosphere – the State and Future of the Eddy Covariance Method, Global Change Biology, 20, 3600–3609, https://doi.org/10.1111/gcb.12649, 2014.

Böhm, C., Kanzler, M., and Freese, D.: Wind Speed Reductions as Influenced by Woody Hedgerows Grown for Biomass in Short Rotation Alley Cropping Systems in Germany, Agroforestry Systems, 88, 579–591, https://doi.org/10.1007/s10457-014-9700-y, 2014.

Bou-Zeid, E., Anderson, W., Katul, G. G., and Mahrt, L.: The Persistent Challenge of Surface Heterogeneity in Boundary-Layer Meteorology: A Review, Boundary-Layer Meteorology, 177, 227–245, https://doi.org/10.1007/s10546-020-00551-8, 2020.

Brutsaert, W.: Land-Surface Water Vapor and Sensible Heat Flux: Spatial Variability, Homogeneity, and Measurement Scales, Water Resources Research, 34, 2433–2442, https://doi.org/10.1029/98WR01340, 1998.

Callejas-Rodelas, J. Á., Knohl, A., van Ramshorst, J., Mammarella, I., and Markwitz, C.: Comparison between Lower-Cost and Conventional Eddy Covariance Setups for CO_2 and Evapotranspiration Measurements above Monocropping and Agroforestry Systems, Agricultural and Forest Meteorology, 354, 110 086, https://doi.org/10.1016/j.agrformet.2024.110086, 2024.

Callejas-Rodelas, J. Á., Knohl, A., Mammarella, I., Vesala, T., Peltola, O., and Markwitz, C.: Dataset of the Journal Article "Does Increased Spatial Replication above Heterogeneous Agroforestry Improve the Representativeness of Eddy Covariance Measurements?", https://doi.org/10.5281/ZENODO.14855287, 2025.

Chasmer, L., Barr, A., Hopkinson, C., McCaughey, H., Treitz, P., Black, A., and Shashkov, A.: Scaling and Assessment of GPP from MODIS Using a Combination of Airborne Lidar and Eddy Covariance Measurements over Jack Pine Forests, Remote Sensing of Environment, 113, 82–93, https://doi.org/10.1016/j.rse.2008.08.009, 2009.

Chen, B., Black, T. A., Coops, N. C., Hilker, T., (Tony) Trofymow, J. A., and Morgenstern, K.: Assessing Tower Flux Footprint Climatology and Scaling Between Remotely Sensed and Eddy Covariance Measurements, Boundary-Layer Meteorology, 130, 137–167, https://doi.org/10.1007/s10546-008-9339-1, 2009.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Barr, A. G., Black, T. A., Arain, M. A., Bourque, C. P.-A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Assessing Eddy-Covariance Flux Tower Location Bias across the Fluxnet-Canada Research Network Based on Remote Sensing and Footprint Modelling, Agricultural and Forest Meteorology, 151, 87–100, https://doi.org/10.1016/j.agrformet.2010.09.005, 2011.

Chen, T. and Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, ACM, San Francisco California USA, https://doi.org/10.1145/2939672.2939785, 2016.

Chu, H., Baldocchi, D. D., Poindexter, C., Abraha, M., Desai, A. R., Bohrer, G., Arain, M. A., Griffis, T., Blanken, P. D., O'Halloran, T. L., Thomas, R. Q., Zhang, Q., Burns, S. P., Frank, J. M., Christian, D., Brown, S., Black, T. A., Gough, C. M., Law, B. E., Lee, X., Chen, J., Reed, D. E., Massman, W. J., Clark, K., Hatfield, J., Prueger, J., Bracho, R., Baker, J. M., and Martin, T. A.: Temporal Dynamics of Aerodynamic Canopy Height Derived From Eddy Covariance Momentum Flux Data Across North American Flux Networks, Geophysical Research Letters, 45, 9275–9287, https://doi.org/10.1029/2018GL079306, 2018.

Cunliffe, A. M., Boschetti, F., Clement, R., Sitch, S., Anderson, K., Duman, T., Zhu, S., Schlumpf, M., Litvak, M. E., Brazier, R. E., and Hill, T. C.: Strong Correspondence in Evapotranspiration and Carbon Dioxide Fluxes Between Different Eddy Covariance Systems Enables Quantification of Landscape Heterogeneity in Dryland Fluxes, Journal of Geophysical Research: Biogeosciences, 127, https://doi.org/10.1029/2021JG006240, 2022.

Cuxart, J., Wrenger, B., Martínez-Villagrasa, D., Reuder, J., Jonassen, M. O., Jiménez, M. A., Lothon, M., Lohou, F., Hartogensis, O., Dünnermann, J., Conangla, L., and Garai, A.: Estimation of the Advection Effects Induced by Surface Heterogeneities in the Surface Energy Budget, Atmospheric Chemistry and Physics, 16, 9489–9504, https://doi.org/10.5194/acp-16-9489-2016, 2016.

Davis, P., Brown, J. C., Saunders, M., Lanigan, G., Wright, E., Fortune, T., Burke, J., Connolly, J., Jones, M., and Osborne, B.: Assessing the Effects of Agricultural Management Practices on Carbon Fluxes: Spatial Variation and the Need for Replicated Estimates of Net Ecosystem Exchange, Agricultural and Forest Meteorology, 150, 564–574, https://doi.org/10.1016/j.agrformet.2010.01.021, 2010.

DWD: Deutscher Wetterdienst Climatological Means, 2024.

Feigenwinter, C., Bernhofer, C., Eichelmann, U., Heinesch, B., Hertel, M., Janous, D., Kolle, O., Lagergren, F., Lindroth, A., Minerbi, S., Moderow, U., Mölder, M., Montagnani, L., Queck, R., Rebmann, C., Vestin, P., Yernaux, M., Zeri, M., Ziegler, W., and Aubinet, M.: Comparison of Horizontal and Vertical Advective CO_2 Fluxes at Three Forest Sites, Agricultural and Forest Meteorology, 148, 12–24, https://doi.org/10.1016/j.agrformet.2007.08.013, 2008.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., and Cleugh, H.: A Re-Evaluation of Long-Term Flux Measurement Techniques Part I: Averaging and Coordinate Rotation, Boundary-Layer Meteorology, 107, 1–48, https://doi.org/10.1023/A:1021554900225, 2003.

Foken, T., Göockede, M., Mauder, M., Mahrt, L., Amiro, B., and Munger, W.: Post-Field Data Quality Control, in: Handbook of Micrometeorology, edited by Lee, X., Massman, W., and Law, B., vol. 29, pp. 181–208, Kluwer Academic Publishers, Dordrecht, https://doi.org/10.1007/1-4020-2265-4_9, 2005.

Garratt, J. R.: The Internal Boundary Layer ? A Review, Boundary-Layer Meteorology, 50, 171–203, https://doi.org/10.1007/BF00120524, 1990.

Göckede, M., Markkanen, T., Hasager, C. B., and Foken, T.: Update of a Footprint-Based Approach for the Characterisation of Complex Measurement Sites, Boundary-Layer Meteorology, 118, 635–655, https://doi.org/10.1007/s10546-005-6435-3, 2006.

Griebel, A., Bennett, L. T., Metzen, D., Cleverly, J., Burba, G., and Arndt, S. K.: Effects of Inhomogeneities within the Flux Footprint on the Interpretation of Seasonal, Annual, and Interannual Ecosystem Carbon Exchange, Agricultural and Forest Meteorology, 221, 50–60, https://doi.org/10.1016/j.agrformet.2016.02.002, 2016.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 Hourly Data on Single Levels from 1959 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)., 2023.

Higgins, C. W., Katul, G. G., Froidevaux, M., Simeonov, V., and Parlange, M. B.: Are Atmospheric Surface Layer Flows Ergodic?, Geophysical Research Letters, 40, 3342–3346, https://doi.org/10.1002/grl.50642, 2013.

Hill, T., Chocholek, M., and Clement, R.: The Case for Increasing the Statistical Power of Eddy Covariance Ecosystem Studies: Why, Where and How?, Global Change Biology, 23, 2154–2165, https://doi.org/10.1111/gcb.13547, 2017.

Hollinger, D. Y. and Richardson, A. D.: Uncertainty in Eddy Covariance Measurements and Its Application to Physiological Models, Tree Physiology, 25, 873–885, https://doi.org/10.1093/treephys/25.7.873, 2005.

Hollinger, D. Y., Aber, J., Dail, B., Davidson, E. A., Goltz, S. M., Hughes, H., Leclerc, M. Y., Lee, J. T., Richardson, A. D., Rodrigues, C., Scott, N., Achuatavarier, D., and Walsh, J.: Spatial and Temporal Variability in Forest–Atmosphere CO_2 Exchange, Global Change Biology, 10, 1689–1706, https://doi.org/10.1111/j.1365-2486.2004.00847.x, 2004.

Hollinger, S. E., Bernacchi, C. J., and Meyers, T. P.: Carbon Budget of Mature No-till Ecosystem in North Central Region of the United States, Agricultural and Forest Meteorology, 130, 59–69, https://doi.org/10.1016/j.agrformet.2005.01.005, 2005.

Hurlbert, S. H.: Pseudoreplication and the Design of Ecological Field Experiments, Ecological Monographs, 54, 187–211, https://doi.org/10.2307/1942661, 1984.

Ibrom, A., Dellwik, E., Flyvbjerg, H., Jensen, N. O., and Pilegaard, K.: Strong Low-Pass Filtering Effects on Water Vapour Flux Measurements with Closed-Path Eddy Correlation Systems, Agricultural and Forest Meteorology, 147, 140–156, https://doi.org/10.1016/j.agrformet.2007.07.007, 2007.

Jänicke, C., Goddard, A., Stein, S., Steinmann, H.-H., Lakes, T., Nendel, C., and Müller, D.: Field-Level Land-Use Data Reveal Heterogeneous Crop Sequences with Distinct Regional Differences in Germany, European Journal of Agronomy, 141, 126 632, https://doi.org/10.1016/j.eja.2022.126632, 2022.

Kaimal, J. C. and Finnigan, J. J.: Atmospheric Boundary Layer Flows: Their Structure and Measurement, Oxford University Press, https://doi.org/10.1093/oso/9780195062397.001.0001, 1994.

Kanzler, M., Böhm, C., Mirck, J., Schmitt, D., and Veste, M.: Microclimate Effects on Evaporation and Winter Wheat (Triticum Aestivum L.) Yield within a Temperate Agroforestry System, Agroforestry Systems, 93, 1821–1841, https://doi.org/10.1007/s10457-018-0289-4, 2019.

Kanzler, M., Böhm, C., and Freese, D.: The Development of Soil Organic Carbon under Young Black Locust (Robinia Pseudoacacia L.) Trees at a Post-Mining Landscape in Eastern Germany, New Forests, 52, 47–68, https://doi.org/10.1007/s11056-020-09779-1, 2021.

Katul, G., Hsieh, C.-I., Bowling, D., Clark, K., Shurpali, N., Turnipseed, A., Albertson, J., Tu, K., Hollinger, D., Evans, B., Offerle, B., Anderson, D., Ellsworth, D., Vogel, C., and Oren, R.: Spatial Variability of Turbulent Fluxes in the Roughness Sublayer of an Even-Aged Pine Forest, Boundary-Layer Meteorology, 93, 1–28, https://doi.org/10.1023/A:1002079602069, 1999.

Kay, S., Rega, C., Moreno, G., Den Herder, M., Palma, J. H., Borek, R., Crous-Duran, J., Freese, D., Giannitsopoulos, M., Graves, A., Jäger, M., Lamersdorf, N., Memedemin, D., Mosquera-Losada, R., Pantera, A., Paracchini, M. L., Paris, P., Roces-Díaz, J. V., Rolo, V., Rosati, A., Sandor, M., Smith, J., Szerencsits, E., Varga, A., Viaud, V., Wawer, R., Burgess, P. J., and Herzog, F.: Agroforestry Creates Carbon Sinks Whilst Enhancing the Environment in Agricultural Landscapes in Europe, Land Use Policy, 83, 581–593, https://doi.org/10.1016/j.landusepol.2019.02.025, 2019.

Kljun, N., Rotach, M., and Schmid, H.: A Three-Dimensional Backward Lagrangian Footprint Model For A Wide Range Of Boundary-Layer Stratifications, Boundary-Layer Meteorology, 103, 205–226, https://doi.org/10.1023/A:1014556300021, 2002.

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: A Simple Two-Dimensional Parameterisation for Flux Footprint Prediction (FFP), Geoscientific Model Development, 8, 3695–3713, https://doi.org/10.5194/gmd-8-3695-2015, 2015.

Kutsch, W. L., Liu, C., Hörmann, G., and Herbst, M.: Spatial Heterogeneity of Ecosystem Carbon Fluxes in a Broadleaved Forest in Northern Germany, Global Change Biology, 11, 70–88, https://doi.org/10.1111/j.1365-2486.2004.00884.x, 2005.

Launiainen, S., Vesala, T., Mölder, M., Mammarella, I., Smolander, S., Rannik, Ü., Kolari, P., Hari, P., Lindroth, A., and Katul, G. G.: Vertical Variability and Effect of Stability on Turbulence Characteristics down to the Floor of a Pine Forest, Tellus B: Chemical and Physical Meteorology, 59, 919, https://doi.org/10.1111/j.1600-0889.2007.00313.x, 2007.

Levy, P., Drewer, J., Jammet, M., Leeson, S., Friborg, T., Skiba, U., and Oijen, M. V.: Inference of Spatial Heterogeneity in Surface Fluxes from Eddy Covariance Data: A Case Study from a Subarctic Mire Ecosystem, Agricultural and Forest Meteorology, 280, 107 783, https://doi.org/10.1016/j.agrformet.2019.107783, 2020.

Lokupitiya, E., Denning, S., Paustian, K., Baker, I., Schaefer, K., Verma, S., Meyers, T., Bernacchi, C. J., Suyker, A., and Fischer, M.: Incorporation of Crop Phenology in Simple Biosphere Model (SiBcrop) to Improve Land-Atmosphere Carbon Exchanges from Croplands, Biogeosciences (Online), 6, 969–986, https://doi.org/10.5194/bg-6-969-2009, 2009.

Lokupitiya, E., Denning, A. S., Schaefer, K., Ricciuto, D., Anderson, R., Arain, M. A., Baker, I., Barr, A. G., Chen, G., Chen, J. M., Ciais, P., Cook, D. R., Dietze, M., El Maayar, M., Fischer, M., Grant, R., Hollinger, D., Izaurralde, C., Jain, A., Kucharik, C., Li, Z., Liu, S., Li, L., Matamala, R., Peylin, P., Price, D., Running, S. W., Sahoo, A., Sprintsin, M., Suyker, A. E., Tian, H., Tonitto, C., Torn, M., Verbeeck, H., Verma, S. B., and Xue, Y.: Carbon and Energy Fluxes in Cropland Ecosystems: A Model-Data Comparison, Biogeochemistry, 129, 53–76, https://doi.org/10.1007/s10533-016-0219-3, 2016.

Lucas-Moffat, A. M., Schrader, F., Herbst, M., and Brümmer, C.: Multiple Gap-Filling for Eddy Covariance Datasets, Agricultural and Forest Meteorology, 325, 109 114, https://doi.org/10.1016/j.agrformet.2022.109114, 2022.

Mahrt, L. and Bou-Zeid, E.: Non-Stationary Boundary Layers, Boundary-Layer Meteorology, 177, 189–204, https://doi.org/10.1007/s10546-020-00533-w, 2020.

Mammarella, I. ., Kolari, P., Rinne, J. ., Keronen, P., Pumpanen, J. ., and Vesala, T.: Determining the Contribution of Vertical Advection to the Net Ecosystem Exchange at Hyytiälä Forest, Finland, Tellus B: Chemical and Physical Meteorology, 59, 900, https://doi.org/10.1111/j.1600-0889.2007.00306.x, 2007.

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, Ü., and Vesala, T.: Relative Humidity Effect on the High-Frequency Attenuation of Water Vapor Flux Measured by a Closed-Path Eddy Covariance System, Journal of Atmospheric and Oceanic Technology, 26, 1856–1866, https://doi.org/10.1175/2009JTECHA1179.1, 2009.

Mammarella, I., Peltola, O., Nordbo, A., Järvi, L., and Rannik, Ü.: Quantifying the Uncertainty of Eddy Covariance Fluxes Due to the Use of Different Software Packages and Combinations of Processing Steps in Two Contrasting Ecosystems, Atmospheric Measurement Techniques, 9, 4915–4933, https://doi.org/10.5194/amt-9-4915-2016, 2016.

Markwitz, C.: Micrometeorological Measurements and Numerical Simulations of Turbulence and Evapotranspiration over Agroforestry, Doctoral Thesis, Georg-August-University Göttingen, https://doi.org/10.53846/goediss-8477, 2021.

Markwitz, C. and Siebicke, L.: Low-Cost Eddy Covariance: A Case Study of Evapotranspiration over Agroforestry in Germany, Atmospheric Measurement Techniques, 12, 4677–4696, https://doi.org/10.5194/amt-12-4677-2019, 2019.

Markwitz, C., Knohl, A., and Siebicke, L.: Evapotranspiration over Agroforestry Sites in Germany, Biogeosciences (Online), 17, 5183–5208, https://doi.org/10.5194/bg-17-5183-2020, 2020.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A Strategy for Quality and Uncertainty Assessment of Long-Term Eddy-Covariance Measurements, Agricultural and Forest Meteorology, 169, 122–135, https://doi.org/10.1016/j.agrformet.2012.09.006, 2013.

Moncrieff, J., Malhi, Y., and Leuning, R.: The Propagation of Errors in Long-Term Measurements of Land-Atmosphere Fluxes of Carbon and Water, Global Change Biology, 2, 231–240, https://doi.org/10.1111/j.1365-2486.1996.tb00075.x, 1996.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Najibnia, S., Koocheki, A., Nassiri, M., and Porsa, H. M.: Water Capture Efficiency, Use Efficiency and Productivity in Sole Cropping and Intercropping of Rapeseed, Bean and Corn, European Journal of Sustainable Development, 3, 347–358, https://doi.org/10.14207/ejsd.2014.v3n4p347, 2014.

Oren, R., Hsieh, C.-I., Stoy, P., Albertson, J., Mccarthy, H. R., Harrell, P., and Katul, G. G.: Estimating the Uncertainty in Annual Net Ecosystem Carbon Exchange: Spatial Variation in Turbulent Fluxes and Sampling Errors in Eddy-Covariance Measurements, Global Change Biology, 12, 883–896, https://doi.org/10.1111/j.1365-2486.2006.01131.x, 2006.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a Standardized Processing of Net Ecosystem Exchange Measured with Eddy Covariance Technique: Algorithms and Uncertainty Estimation, Biogeosciences (Online), 3, 571–583, https://doi.org/10.5194/bg-3-571-2006, 2006.

Peltola, O., Hensen, A., Helfter, C., Belelli Marchesini, L., Bosveld, F. C., Van Den Bulk, W. C. M., Elbers, J. A., Haapanala, S., Holst, J., Laurila, T., Lindroth, A., Nemitz, E., Röckmann, T., Vermeulen, A. T., and Mammarella, I.: Evaluating the Performance of Commonly Used Gas Analysers for Methane Eddy Covariance Flux Measurements: The InGOS Inter-Comparison Field Experiment, Biogeosciences (Online), 11, 3163–3186, https://doi.org/10.5194/bg-11-3163-2014, 2014.

Peltola, O., Hensen, A., Belelli Marchesini, L., Helfter, C., Bosveld, F., Van Den Bulk, W., Haapanala, S., Van Huissteden, J., Laurila, T., Lindroth, A., Nemitz, E., Röckmann, T., Vermeulen, A., and Mammarella, I.: Studying the Spatial Variability of Methane Flux with Five Eddy Covariance Towers of Varying Height, Agricultural and Forest Meteorology, 214–215, 456–472, https://doi.org/10.1016/j.agrformet.2015.09.007, 2015.

Pohanková, E., Hlavinka, P., Orság, M., Takáč, J., Kersebaum, K. C., Gobin, A., and Trnka, M.: Estimating the Water Use Efficiency of Spring Barley Using Crop Models, The Journal of Agricultural Science, 156, 628–644, https://doi.org/10.1017/S0021859618000060, 2018.

Prabha, T. V., Leclerc, M. Y., and Baldocchi, D.: Comparison of In-Canopy Flux Footprints between Large-Eddy Simulation and the Lagrangian Simulation, Journal of Applied Meteorology and Climatology, 47, 2115–2128, https://doi.org/10.1175/2008jamc1814.1, 2008.

Quandt, A., Neufeldt, H., and Gorman, K.: Climate Change Adaptation through Agroforestry: Opportunities and Gaps, Current Opinion in Environmental Sustainability, 60, 101 244, https://doi.org/10.1016/j.cosust.2022.101244, 2023.

Ran, Y., Li, X., Sun, R., Kljun, N., Zhang, L., Wang, X., and Zhu, G.: Spatial Representativeness and Uncertainty of Eddy Covariance Carbon Flux Measurements for Upscaling Net Ecosystem Productivity to the Grid Scale, Agricultural and Forest Meteorology, 230–231, 114–127, https://doi.org/10.1016/j.agrformet.2016.05.008, 2016.

Rannik, Ü. and Vesala, T.: Autoregressive Filtering versus Linear Detrending in Estimation of Fluxes by the Eddy Covariance Method, Boundary-Layer Meteorology, 91, 259–280, https://doi.org/10.1023/A:1001840416858, 1999.

Rannik, Ü., Kolari, P., Vesala, T., and Hari, P.: Uncertainties in Measurement and Modelling of Net Ecosystem Exchange of a Forest, Agricultural and Forest Meteorology, 138, 244–257, https://doi.org/10.1016/j.agrformet.2006.05.007, 2006.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir,

D., and Valentini, R.: On the Separation of Net Ecosystem Exchange into Assimilation and Ecosystem Respiration: Review and Improved Algorithm, Global Change Biology, 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Richardson, A. D. and Hollinger, D. Y.: A Method to Estimate the Additional Uncertainty in Gap-Filled NEE Resulting from Long Gaps in the CO2 Flux Record, Agricultural and Forest Meteorology, 147, 199–208, https://doi.org/10.1016/j.agrformet.2007.06.004, 2007.

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., William Munger, J., Ricciuto, D. M., Stoy, P. C., Suyker, A. E., Verma, S. B., and Wofsy, S. C.: A Multi-Site Analysis of Random Error in Tower-Based Measurements of Carbon and Energy Fluxes, Agricultural and Forest Meteorology, 136, 1–18, https://doi.org/10.1016/j.agrformet.2006.01.007, 2006.

Sabbatini, S., Mammarella, I., Arriga, N., Fratini, G., Graf, A., Hörtnagl, L., Ibrom, A., Longdoz, B., Mauder, M., Merbold, L., Metzger, S., Montagnani, L., Pitacco, A., Rebmann, C., Sedlák, P., Šigut, L., Vitale, D., and Papale, D.: Eddy Covariance Raw Data Processing for CO2 and Energy Fluxes Calculation at ICOS Ecosystem Stations, International Agrophysics, 32, 495–515, https://doi.org/10.1515/intag-2017-0043, 2018.

Soegaard, H.: Carbon Dioxide Exchange over Agricultural Landscape Using Eddy Correlation and Footprint Modelling, Agricultural and Forest Meteorology, 114, 153–173, https://doi.org/10.1016/S0168-1923(02)00177-6, 2003.

Stoy, P. C., Chu, H., Dahl, E., Cala, D. S., Shveytser, V., Wiesner, S., Desai, A. R., and Novick, K. A.: The Global Distribution of Paired Eddy Covariance Towers, https://doi.org/10.1101/2023.03.03.530958, 2023.

Sun, F., Roderick, M. L., Farquhar, G. D., Lim, W. H., Zhang, Y., Bennett, N., and Roxburgh, S. H.: Partitioning the Variance between Space and Time, Geophysical Research Letters, 37, 2010GL043 323, https://doi.org/10.1029/2010GL043323, 2010.

Trouwloon, D., Streck, C., Chagas, T., and Martinus, G.: Understanding the Use of Carbon Credits by Companies: A Review of the Defining Elements of Corporate Climate Claims, Global Challenges, 7, 2200 158, https://doi.org/10.1002/gch2.202200158, 2023.

van Ramshorst, J. G. V., Siebicke, L., Baumeister, M., Moyano, F. E., Knohl, A., and Markwitz, C.: Reducing Wind Erosion through Agroforestry: A Case Study Using Large Eddy Simulations, Sustainability, 14, 13 372, https://doi.org/10.3390/su142013372, 2022.

van Ramshorst, J. G. V., Knohl, A., Callejas-Rodelas, J. Á., Clement, R., Hill, T. C., Siebicke, L., and Markwitz, C.: Lower-Cost Eddy Covariance for CO $_2$ and H $_2$ O Fluxes over Grassland and Agroforestry, https://doi.org/10.5194/amt-2024-30, 2024.

Vekuri, H., Tuovinen, J.-P., Kulmala, L., Papale, D., Kolari, P., Aurela, M., Laurila, T., Liski, J., and Lohila, A.: A Widely-Used Eddy Covariance Gap-Filling Method Creates Systematic Bias in Carbon Balance Estimates, Scientific Reports, 13, 1720, https://doi.org/10.1038/s41598-023-28827-2, 2023.

Veldkamp, E., Schmidt, M., Markwitz, C., Beule, L., Beuschel, R., Biertümpfel, A., Bischel, X., Duan, X., Gerjets, R., Göbel, L., Graß, R., Guerra, V., Heinlein, F., Komainda, M., Langhof, M., Luo, J., Potthoff, M., Van Ramshorst, J. G. V., Rudolf, C., Seserman, D.-M., Shao, G., Siebicke, L., Svoboda, N., Swieter, A., Carminati, A., Freese, D., Graf, T., Greef, J. M., Isselstein, J., Jansen, M., Karlovsky, P., Knohl, A., Lamersdorf, N., Priesack, E., Wachendorf, C., Wachendorf, M., and Corre, M. D.: Multifunctionality of Temperate Alley-Cropping Agroforestry Outperforms Open Cropland and Grassland, Communications Earth & Environment, 4, 20, https://doi.org/10.1038/s43247-023-00680-1, 2023.

Vesala, T., Kljun, N., Rannik, Ü., Rinne, J., Sogachev, A., Markkanen, T., Sabelfeld, K., Foken, Th., and Leclerc, M.: Flux and Concentration Footprint Modelling: State of the Art, Environmental Pollution, 152, 653–666, https://doi.org/10.1016/j.envpol.2007.06.070, 2008.

Vuichard, N. and Papale, D.: Filling the Gaps in Meteorological Continuous Data Measured at FLUXNET Sites with ERA-Interim Reanalysis, Earth System Science Data, 7, 157–171, https://doi.org/10.5194/essd-7-157-2015, 2015.

Wang, H., Jia, G., Zhang, A., and Miao, C.: Assessment of Spatial Representativeness of Eddy Covariance Flux Data from Flux Tower to Regional Grid, Remote Sensing, 8, 742, https://doi.org/10.3390/rs8090742, 2016.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic Anemometer Tilt Correction Algorithms, Boundary-Layer Meteorology, 99, 127–150, https://doi.org/10.1023/A:1018966204465, 2001.

Winck, B. R., Bloor, J. M. G., and Klumpp, K.: Eighteen Years of Upland Grassland Carbon Flux Data: Reference Datasets, Processing, and Gap-Filling Procedure, Scientific Data, 10, 311, https://doi.org/10.1038/s41597-023-02221-z, 2023.

Wohlfahrt, G., Hörtnagl, L., Hammerle, A., Graus, M., and Hansel, A.: Measuring Eddy Covariance Fluxes of Ozone with a Slow-Response Analyser, Atmospheric Environment, 43, 4570–4576, https://doi.org/10.1016/j.atmosenv.2009.06.031, 2009.

Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., Menzer, O., and Reichstein, M.: Basic and Extensible Post-Processing of Eddy Covariance Flux Data with REddyProc, Biogeosciences (Online), 15, 5015–5030, https://doi.org/10.5194/bg-15-5015-2018, 2018.