Building a random forest machine learning model for Net Ecosystem Exchange estimation in agricultural fields using discontinuous atmospheric Eddy Covariance measurements.

Camille Pery[1,2,3], Pierre Polsenaere[1], Eric Lamaud[2], Natacha Volto[4], Nathalie Long[4], Jonathan Deborde[1], Shunko Bolsée[2], Olivier Philippine[5], Vona Méléder[6,7], Christine Dupuy[4]

[1] Ifremer, COAST, 17390 La Tremblade, France
[2] INRAe, ISPA, 33140 Villenave D'Ornon, France
[3] Univ. Bordeaux, CNRS, Bordeaux INP, UMR 5805, F-33600 Pessac, France
[4] LIENSs, UMR7266 CNRS- La Rochelle Université, 17000 La Rochelle, France
[5] Unima, 17180 Périgny, France
[6] Nantes Université, ISOMer, 44000 Nantes, France
[7] Institut Universitaire de France, IUF, France

Correspondence to: Camille Pery (camille.pery@bordeaux-inp.fr)

Abstract. Atmospheric CO_2 exchanges in agroecosystems are strongly controlled by plant phenology, management and microclimatic conditions; yet obtaining finely resolved flux estimates across heterogeneous agricultural landscapes remains challenging. This study evaluates the ability of a single Eddy Covariance (EC) system combined with wind-sector partitioning and Random Forest (RF) modelling to estimate the Net Ecosystem Exchange (NEE) of adjacent fields (wheat, mixed-grain, permanent grassland) in the Marais poitevin wetland (French Atlantic coast). Fluxes were measured over the period 2023-01-01-2024-01-31 and attributed to the different fields by wind sectors. Two modelling strategies were compared (i) a single global RF trained on all sectors and (ii) sector-specific RFs for each adjacent field. RF models globally showed good overall performance ($R^2 \approx 0.20$-0.94 depending on sector and period), but this apparent equivalence broke down once fluxes were reconstructed over the full annual cycle: the sector-specific approach then better reproduced phenological dynamics and management-driven flux reversals (harvest, grazing) than the global model, which compressed inter-sector variability by up to 67%. Annual NEE estimated from the sector-specific models indicate that the mixed-grain field acted as net source (+151 g C m^{-2} yr^{-1}), whereas the permanent grassland and the wheat field acted as net sinks (−145 and −308 g C m^{-2} yr^{-1}). A light-response-curve partitioning of these reconstructed fluxes into gross primary production (GPP) and ecosystem respiration (Reco) revealed contrasting phenological signatures behind different land-cover. Main limitations include spatial attribution uncertainty related to footprint mixing with ditch and riparian vegetation for sectors adjacent to these features, flux disturbances during stormy episodes, and the limited one-year observation period. This study highlights the novelty and practical value of coupling a single EC system with wind-sector partitioning and machine learning approaches to resolve carbon fluxes at the field scale within heterogeneous agricultural landscapes. This integrated approach provides a cost-effective alternative to traditional multi-tower setups, offering new opportunities to monitor spatial carbon dynamics and management effects in real agricultural mosaics. Beyond methodological innovation, this work provides one of the first field-scale, uncertainty-resolved

carbon flux datasets for a mosaic of contrasting agricultural land covers monitored by a single EC tower, offering a transferable framework to better constrain the terrestrial carbon balance of retro-littoral wetland landscapes.

Keywords: Eddy Covariance, Random Forest, Net Ecosystem Exchange, wind sectors, agriculture, gap-filling.

1 Introduction

Agricultural lands cover a vast fraction of the terrestrial biosphere with around 50% of the Earth's vegetated land surface (about 5 billion ha) under cultivation, grassland or rangeland; this extensive area makes agroecosystems a major component of the global carbon cycle. Indeed, since 1750 corresponding to the start of the industrial era, conversion of forests and meadows to cropland and grassland has released large amounts of CO_2 on the order of 214 ± 67 PgC from biomass and soils, contributing roughly to a quarter of anthropogenic greenhouse-gas emissions (Zomer et al., 2017). In the last decade, net CO_2 emissions from land-use change (mostly driven by agriculture) averaged about 4.2 ± 2.6 Gt CO_2 yr^{-1} (Friedlingstein et al., 2025). At the same time, agricultural ecosystems can remove atmospheric CO_2 through crop photosynthesis and store part of this carbon in plant biomass and soils. However, their overall role as a carbon sink or source depends largely on agricultural management practices too. Conventional practices, such as intensive tillage and crop residue removal, can reduce soil carbon stocks and increase greenhouse gas emissions, whereas more conservational practices can enhance soil carbon sequestration and mitigate CO_2 emissions. Globally, agricultural soils hold on the order of 8-10% of the world's soil organic carbon (Wu et al., 2024), so even small changes in land management can have measurable impacts on the atmospheric CO_2 concentration. Climate and land-use intensity control the Net Ecosystem Exchange (NEE) of agroecosystems in complex and interconnected ways: solar radiation and air temperature are the primary drivers of gross primary production (GPP) in many temperate croplands (Moreaux et al., 2020), while water availability, plant phenology and vapor pressure deficit (VPD) modulate the growing-season uptake. Crop type and anthropogenic management (e.g. fertilization, irrigation, harvest) further alter the balance between photosynthesis and respiration; for instance, deep-rooted or perennial crops typically achieve higher seasonal carbon uptake than shallow-rooted annuals, and no-harvest practices slow down soil CO_2 release. Thus, the net carbon budget of agricultural fields sensitively depends on crop species, phenological stage, and local environment variables, as well as on socio-economic factors determining land-use intensity (Moreaux et al., 2020; Zomer et al., 2017).

To quantify NEE over large spatial scales and fine time resolution, the atmospheric Eddy Covariance (EC) technique has become a widely used method in ecosystem science, although it remains costly and technically demanding. EC deploys fast response anemometers and CO_2/H_2O Infra-Red Gaz Analysers (IRGA) on towers to directly measure the turbulent flux of CO_2 between the biosphere and atmosphere (Baldocchi, 2020). Since the late 1990s, global networks of EC towers (e.g. FLUXNET, AmeriFlux, AsiaFlux, etc.) have accumulated long time series of carbon and water fluxes from hundreds of sites representing diverse biomes (Baldocchi, 2020). In Europe, the Integrated Carbon Observation System (ICOS) network similarly provides standardized CO_2 flux measurements through 180 stations deployed across 16 European countries. In total, these networks

have been invaluable for linking ecosystem carbon exchange to climate drivers; they include sites across continents (North America, Europe, Asia, Africa, etc.) and ecosystem types from boreal and temperate forests to grasslands, croplands, wetlands, and even urban landscapes. For example, Baldocchi (2020) notes that EC networks now span the main climate and ecological spaces, revealing how gross primary production (GPP) and ecosystem respiration (R_{eco}) respond to sunlight, temperature and moisture. These multi-decadal observations have elucidated several key environmental controls on ecosystem carbon cycling. Solar radiation emerges as the primary driver of photosynthetic carbon uptake, with light response curves showing saturation effects at high irradiance levels and demonstrating how different vegetation types exhibit varying light-use efficiencies (Reichstein et al., 2007; Wohlfahrt and Gu, 2015). Temperature controls operate through multiple pathways: while moderate warming generally enhances both photosynthesis and respiration, the temperature sensitivity of respiration, commonly expressed as the Q_{10} coefficient (i.e., the factor by which a biological process increases for a 10 °C rise in temperature), typically exceeds that of photosynthesis, leading to reduced net carbon uptake at higher temperatures (Kirschbaum, 1995; Lloyd and Taylor, 1994; Mahecha et al., 2010). Moisture availability represents another fundamental control, with vapor pressure deficit (VPD) strongly regulating stomatal conductance and thus photosynthetic rates, while soil water content directly affects both plant productivity and microbial respiration (Knauer et al., 2018; Novick et al., 2016).

EC measurements also have well-known limitations; for instance, raw flux time series typically contain gaps due to instrument maintenance, meteorological conditions (precipitation), low turbulence or filtering of spurious data. Then, filling these gaps is necessary, requiring statistical or machine-learning gap-filling routines, which can at the same time, introduce uncertainty if large or systematic gaps occur (Gao et al., 2023). Moreover, an EC tower integrates fluxes over an upwind footprint whose size and shape change with surface roughness, atmospheric stability and wind direction; in heterogeneous landscapes, a single tower may sample multiple land covers, complicating attribution of fluxes to specific fields or vegetation types. Such footprint heterogeneity can bias estimates when land cover is patchy, a particular concern for small agricultural fields in a mosaic of crops and fallows (Kumari et al., 2020a; Bou-Zeid et al., 2020).

Recently, Machine Learning (ML) techniques have been increasingly applied to model and gap-fill NEE, taking advantage of large EC datasets and remote-sensing/meteorological inputs. Unlike simple empirical regressions or process-based models that require detailed parameterization, ML methods can capture complex nonlinear interactions without explicit mechanistic assumptions. Typical algorithms include Random Forests (RF), artificial neural networks, support vector machines, and ensemble approaches. For example, Zeng et al. (2020) used a RF model to upscale EC-based NEE, GPP and R_{eco} from FLUXNET datasets across the globe, producing high-resolution carbon flux maps. In agricultural contexts, ML has proven effective for both gap-filling and predictive modelling. Gao et al. (2023) demonstrated that a RF-based gap-filling framework (combined with moving-average decomposition) outperformed conventional methods (including linear interpolation, mean diurnal variation, and look-up table approaches commonly used in the FLUXNET standardized processing) over extended CO_2 flux gaps. A recent study applied a knowledge-guided ML framework to agroecosystems in the U.S. Corn Belt and showed it

surpassed both pure process models and "black-box" ML models in quantifying carbon dynamics (Liu et al., 2024). These studies highlight key advantages of ML, i.e., flexibility in assimilating diverse inputs (weather, soil moisture, phenology, management) and the ability to identify subtle patterns in flux variability. In practice, RF are especially popular for flux prediction because they are robust to overfitting and provide variable importance metrics.

In this context, the La Rochelle Zero Carbon Territory (LRTZC) project (2019-2027) provides a regional framework for carbon monitoring and mitigation in the perimeter of La Rochelle Agglomeration. The project has pledged carbon neutrality by 2040 through combined mitigation and sequestration efforts. Under LRTZC, the Blue Carbon axis and associated research actions target carbon fluxes in coastal ("blue carbon") and retro-littoral ecosystems that correspond to freshwater, brackish and salt shallow systems separated from tidal immersion by dykes and locks, and are now numerous particularly along the French Atlantic coast. As part of these efforts, an EC tower has been established from February 2023 to February 2024 in a rural wetland (freshwater marsh, Marans) to monitor retro-littoral NEE. This wetland ecosystem is constituted by agricultural parcels (terrestrial compartment) separated by ditches (aquatic compartment). Because the EC system cannot fully isolate fluxes from the water ditches, this single tower overlooks three adjacent fields with different crops and management histories. By partitioning the 360° EC data into specific chosen wind-direction sectors, the "wind-sector approach" adopted in the present study allows analysing separate NEE time series for each parcel of interest depending on wind directions.

The objective of this study was to estimate the NEE of the three individual fields using the sector-specific EC data. RF models were built for each wind sector to predict CO_2 flux from meteorological and soil variables measured at the tower location. Annual NEE was then computed for each parcel. NEE was further partitioned into gross primary production (GPP) and ecosystem respiration (Reco) for each sector, using a light-response-curve approach applied to daytime NEE, in order to quantify the relative contribution of photosynthetic uptake and respiration to each parcel's annual carbon budget. To test the benefit of this disaggregated approach, a "pooled" model was trained using all data from all sectors (ignoring field boundaries), and its flux predictions were compared to those from the separate specific models. The hypothesis was that sector-specific models would more accurately capture each field's distinct NEE (reflecting crop and management differences) and thus yield more precise values than a single model that pooled all fields. By evaluating disaggregated versus pooled modelling, the study aimed to demonstrate how wind-sector partitioning could enhance carbon monitoring in heterogeneous agricultural landscapes. In this way, complete annual NEE values of neighbouring fields could be compared under the same climatic conditions without multiple tower needs.

2 Materials and methods

2.1 Study site

The study was conducted in the city of Marans, located in the Charente-Maritime department in southwestern France (46.31° N, 0.99° W). Charente-Maritime is characterized by a temperate oceanic climate and a predominantly agricultural landscape representing 62% of the total Charente-Maritime region surface, corresponding to 425,267 ha; the regional land use is dominated by cereal crops (63% of the utilised agricultural area, UAA), followed by permanent grasslands (18%), alongside viticulture and livestock farming (Agreste, 2020).

The experimental site comprises three adjacent land parcels representative of the regional land use structure. To the northeast, a parcel was cultivated with mixed grain, sown in autumn and harvested on 24 May 2023, after which the field remained as bare soil for the remainder of the measurement period. To the south, a second parcel was planted with wheat in February and harvested on 7 July 2023. In the northwest, a permanent grassland parcel, continuously grazed by cattle from June to November (i.e. cattle numbers of 10 in February 2023, 41 in April 2023 and 43 in July 2023), was maintained throughout the study. Each parcel is delineated by hydrological features, including a network of drainage ditches typical of the Marais poitevin wetland system (Butzer, 2002). The Brune Ditch runs between the northern parcels and the wheat parcel and is intersected by a smaller secondary ditch bordered by sparse woody vegetation separating the mixed grain and grassland fields.



Figure 1. Agricultural fields located in the Poitevin Marsh (Marans city) seen by the atmospheric Eddy Covariance station (red dot) deployed from February 2023 to April 2024. Wheat field in the south, mixed-grain field in the northeast and grassland in the northwest sectors (Drone orthomosaic image taken in February 2023). The Brune Ditch (in brown) orientated from west to east, separates wheat fields in the south from mixed-grain and grassland fields in the north.

The Eddy Covariance (EC) tower was installed on the 1st of February 2023 at a narrow-grassed strip approx. 5 m wide adjacent to the Brune Ditch, providing a central location from which flux measurements could be made across the multiple land covers. The site configuration allowed for the assessment of atmospheric CO_2 fluxes (NEE) across contrasting land uses within a complex agricultural landscape.

2.2 Atmospheric CO_2 exchanges and associated meteorological data measured at the ecosystem scale by Eddy Covariance method

2.2.1 Eddy Covariance data acquisition

The atmospheric EC technique is a micrometeorological method that quantifies net CO_2 fluxes at the ecosystem-atmosphere interface. In the atmosphere, CO_2 is transported by numerous eddies of varying size and frequency. These eddies occur in the

surface boundary layer. The method relies on measuring the covariance between fluctuations in vertical wind velocity and scalar concentrations within these eddies. The CO_2 flux can be expressed as:

$$F = \overline{\rho_d ws} \approx \overline{\rho_d w's'}$$

where the overbar indicates the time averaging and the prime represents fluctuations around this average (Burba, 2013; Reynolds, 1886). Here, (ρ_d) denotes the dry air density (kg m^{-3}), (w) the vertical wind velocity (m s^{-1}), and (s) the dry mole fraction of CO_2 (mol mol^{-1}, commonly expressed in µmol mol^{-1} or ppm). The covariance between these variables yields a mass flux with units of kg m^{-2} s^{-1}. For CO_2 exchange studies, this mass flux is converted into a molar flux by dividing by the molar mass of CO_2 (MCO_2 = 44.01 g mol^{-1}) and is therefore commonly reported in µmol m^{-2} s^{-1}. Negative CO_2 flux values indicate the ecosystem acts as a CO_2 sink, while positive flux values indicate ecosystem source behaviour.

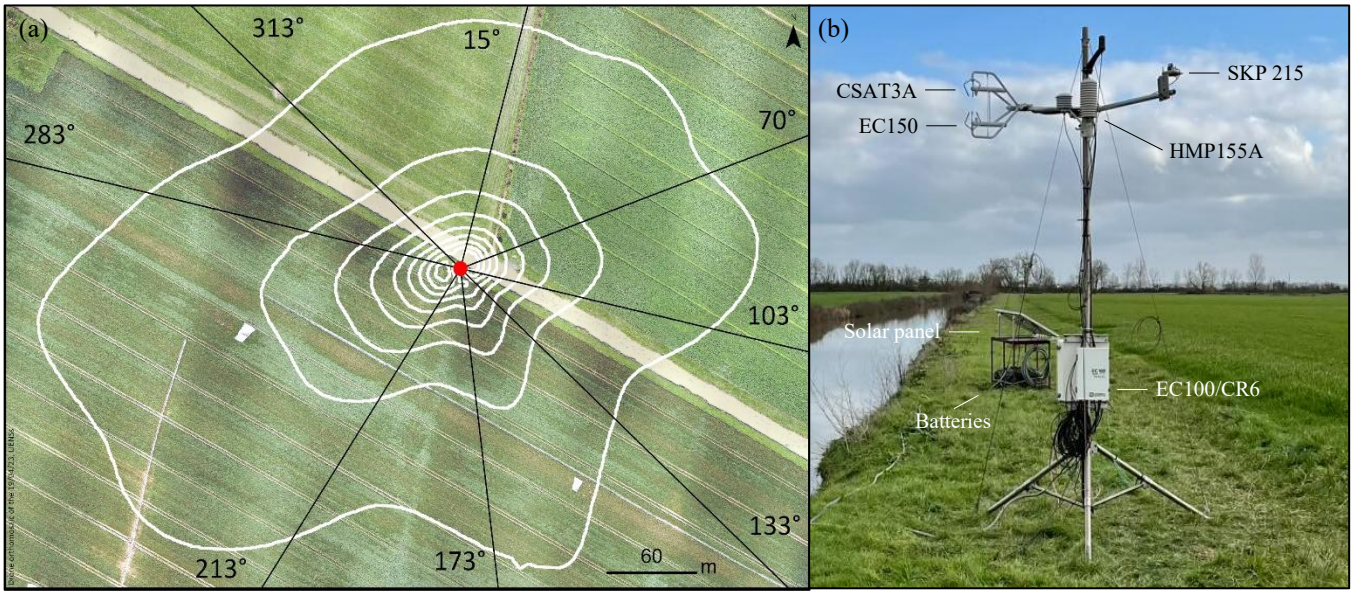


Figure 2. (a) Study site spatial partitioning into wind direction sectors for footprint-based attribution of CO_2 flux measurements to land cover types. High-resolution drone orthomosaic of the study area showing angular sectors centred on the EC tower (red dot). 15° to 70°: Sector 1 (hybrid transition zone), 70° to 103°: Sector 2 (mixed grain), 103° to 133°: Sector 3 (ditch), 133° to 283°: Sectors 4, 5 and 6 (wheat), 283° to 313°: Sector 7 (ditch), and 313° to 15°: Sector 8 (grassland).10% to 90% contour lines in white of the annual footprint. (b) The atmospheric EC station set-up with the two main anemometer and CO_2 EC sensors (CSAT3A and EC150, Campbell Sci.) linked to EC100 electronics module connected to the data acquisition central (CR6), with the associated environmental sensors (HMP155A, SKP 215 and TE525MM). The system is powered by batteries charged by solar panels (see main text for more details).

Mean fluxes were calculated over a 10 min. averaging time period we now use since several years, as particularly adapted for more coastal and dynamic systems such as wetlands (present study), salt marshes (i.e., Mayen et al., 2024) and tidal flats (i.e., Polsenaere et al., 2012). For the present study, computations done comparing the friction velocity, wind speed and flux data between the mean of the three 10 min. values and the corresponding mean values obtained with an averaging time period of

30 min. (see graphs in Appendix A, Fig. A1) showed no low frequency losses on the 10 min. flux data, though some dispersion between both 10 min. and 30 min. mean values (u^*, wind speeds, mean fluxes) were noticed. Choosing a 10 min. averaging time period (instead of 30 min.) permitted losing less data due to non-stationary conditions and/or electronical noises leading to potential and more data suppression (with the 30 min. choice) according to our data quality check procedure adopted in the present study (see Appendix A, Fig. A1 and detailed procedure below in Section 2.2.2).

The EC station (Campbellsci), installed and monitored from February, 1st 2023 to April, 3rd 2024 between three agricultural parcels (Fig. 1), includes the following sensors positioned at the top of a 3m tower mast (Fig. 2b): an ultrasonic anemometer (CSAT3A) that measures three-dimensional wind speed components, and an open-path IRGA (EC150, connected to the EC100 central) that measures CO_2 and H_2O concentrations in the air both at a frequency of 20 Hz. The station also included additional sensors measuring various environmental parameters every 10 minutes: a photosynthetically active radiation (PAR) sensor (SKP215), an air temperature and relative humidity sensor (HMP155A), and a rainfall sensor (TE525MM, Texas Electronics) positioned few meters away in front of the tower. Soil temperature and water content were measured by another probe (CS655). The data acquisition system was powered by two batteries (12 volts, 260 Ah) charged by monocrystalline solar panels (24 V, 200Wp module with MPPT 100 V/30 A controller; Victron Energy). High-frequency data were recorded on an SD micro-card and replaced every two weeks along sensor cleaning. Meteorological data were stored in the central acquisition system (CR6 datalogger, Campbellsci) (Fig. 2b).

2.2.2 EC data processing and filtering

Raw EC data were processed using the EddyPro software (EddyPro® v7.0.8, LI-COR Inc.). Following Vickers and Mahrt (1997) the processing included several key steps: (1) unit conversion to check that the units for instantaneous data are appropriate and consistent to avoid any errors in the calculation and correction of CO_2 fluxes; (2) despiking to remove outliers in the instantaneous data from the anemometer and gas analyser due to electronic and physical noise and replaced the detected spikes with a linear interpolation of the neighbouring values; (3) amplitude resolution to identify situations in which the signal variance is too low with respect to the instrumental resolution; (4) double coordinate rotation to align the x axis of the anemometer to the current mean streamlines, nullifying the vertical and cross-wind components; (5) time delay removal by detecting discontinuities and time shifts in the signal acquisition from the anemometer and gas analyser (i.e. covariance maximization option for the time lag detection method that calculates the most likely time lag based on the circular correlation procedure to avoid possible flux underestimations in the case of existing potential time lags); (6) detrending with removal of short-term linear trends to suppress the impact of low-frequency air movements; and (7) performing the Webb-Pearman-Leuning (WPL) correction to take into account the effects of temperature and water vapor fluctuations on the measured fluctuations in the CO_2 and H_2O densities (Burba, 2013). Linear detrending was used to calculate fluctuations as the deviations from a linear trend on the chosen 10-minute averaging time period (Gash and Culf, 1996). Data with more than 10% missing high-frequency data over these 10-minutes were excluded. Then, EddyPro software outputs contains 4% missing values.

Additionally, some extreme values in data were identified and removed using the median absolute deviation (MAD) method (Papale et al., 2006). The MAD was calculated over a 2-week window separating daytime and night-time periods. Data above 5.2×MAD were removed. Without extreme values, the number of missing values reaches 12% for the studied period from 2023-02-01 to 2024-01-31. In addition, data were filtered removing periods where friction velocity (u_*) values were less than 0.1 m s^{-1}. This parameter quantifies atmospheric turbulence then it allows identifying poorly mixed periods. The EC method often underestimates fluxes during these stable periods due to insufficient turbulent mixing. During these periods, the near-ground air can be decoupled from the air above, leading to an accumulation of CO_2 that is not exactly captured by measuring instruments (Barr et al., 2013; Gu et al., 2005). This u_* filter optimizes the accuracy of EC measurements particularly during night. After thisfiltration the data set contains a total of 32% missing values with 23% of filtered data corresponding to the night, compared with 9% during the day (for more detail, see Mayen et al., 2024; 2025). To go further and remove other extreme values in the dataset particularly during periods of strong meteorological disturbances observed in Autumn 2023 and taking into account crop managements as well, we introduced hard limits based on the observed range of realistic ecosystem fluxes. Before harvest, NEE values were constrained between +20 and −50 µmol m^{-2} s^{-1}, whereas after harvest, when ecosystem activity was substantially reduced, a range of +10 to −10 µmol m^{-2} s^{-1} was applied. These limits removed a small number of spurious observations, including the most extreme values associated with the November storm period in particular (Fig 3). After application of this additional hard-limits filter, the overall proportion of missing CO_2 flux data reached 34% for the studied period.It should be noted that the data-filtering procedure with associated steps described above applies exclusively to CO_2 flux measurements. This filtering does not affect the meteorological variables used as RF predictors, for which the fraction of missing data was minimal (≈0.03% for air and soil temperature and soil humidity, ≈0.005% for relative humidity, and ≈0.68% for wind direction, wind speed and atmospheric pressure), with gaps never exceeding seven consecutive missing values (70 minutes). Given their limited extent and short duration, these gaps were filled using a simple mean-based approach, ensuring temporal continuity of the meteorological inputs without introducing unnecessary complexity.

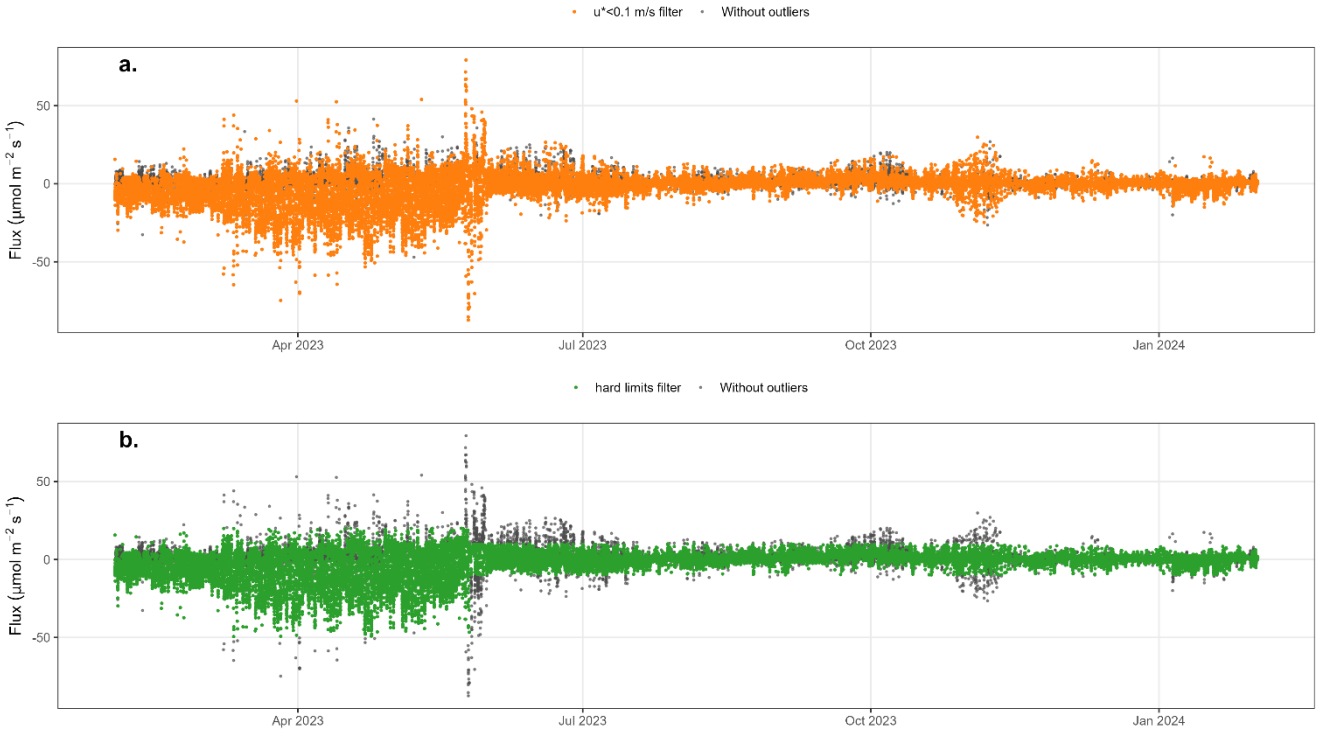


Figure 3. Comparison of NEE time series after application of (a) the $u*$ (< 0.1 m s^{-1}) filter and (b) the hard-limits filter. Grey points represent the dataset after outlier removal only, while coloured (orange and green) points indicate the retained fluxes after each filtering step ($u*$ and hard-limits filter, respectively). Missing data reached 32% after the u_x* filter alone and 34.25% after the additional hard-limits filter.

2.3 Intermittent EC data modelling

2.3.1 Landscape-based wind sector classification, vegetation dynamic analysis and footprint estimation

To spatially constrain the surface sources contributing to the measured NEE, we delineated a set of wind sectors that reflect both the surrounding land-cover distribution (Fig. 2a) and the roughness-related flow heterogeneity inferred from turbulence measurements (Fig. 4). The objective of the sectorisation was primarily to characterize the spatial heterogeneity of the flux footprint and the effects of contrasting management practices within the study area. In addition to the spatial analysis of landscape structure, we used the ratio between friction velocity ($u*$) and mean wind speed (U) as a function of wind direction. Because $u*$ responds sensitively to upwind vegetation height, discontinuities and surface roughness, directional variations in the $u*/U$ ratio provide an empirical signature of surface-roughness transitions, making it possible to identify, the position of ditch banks, riparian vegetation or flow disturbances induced by the tower infrastructure (Bolsée et al., 2026). Indeed, in the atmospheric surface layer, the presence of relief elements, whether they are plant-based as in our case or man-made, causes an increase in surface roughness (z_0), which leads to an increase in the roughness coefficient $u*/U$, as inferred from the following equation $u(z) = (u*/k)\ \ln((z - d)/z_0)$. Therefore, the spatial distribution of $u*/U$, obtained from momentum EC flux

measurements allows differentiating the areas responsible for the CO_2 fluxes measured by the same setup (Bolsée et al., 2026). In the present study, this approach was particularly effective after mowing, when all cropped fields were essentially bare soil while riparian corridors remained vegetated, thereby amplifying roughness contrasts between parcels and allowing a precise alignment between directional u_*/U features and landscape elements. The spatial origin of the measured NEE was determined under all atmospheric stability conditions (instead of selecting only neutral conditions) as it could give a biased view of the sectors where CO_2 fluxes come from, R_{eco} being linked to stable night-time conditions and GPP being the main component of daytime NEE under unstable conditions.

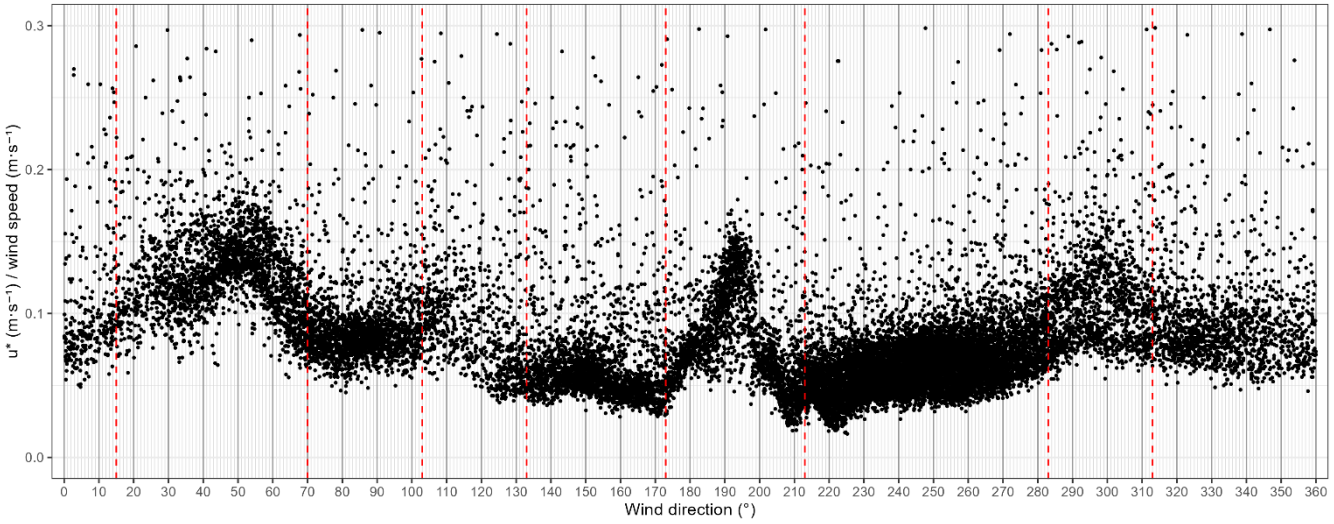


Figure 4. Directional variability of the u_*/U ratio after wheat harvest (10/07/2023), highlighting roughness transitions and disturbances used to define wind-sector boundaries (defined by dashed red lines) around the EC tower.

Based on the combined interpretation of land-cover maps and directional u_*/U patterns, the area surrounding the EC tower was partitioned into eight directional sectors reflecting the main land-cover units within the flux footprint (Figs. 2a and 4). Sector 1 (15–70°) corresponds to the transition zone between grassland and mixed-grain parcels, together with the associated riparian vegetation. Sector 2 (70–103°) covers the mixed-grain field, up to the first ditch axis. Sector 3 (103–133°) isolates this first ditch and its bordering riparian corridor. The wheat field, extending from 133° to 283°, was further subdivided into three sub-sectors: Sector 4 (133–173°), Sector 5 (173–213°) and Sector 6 (213–283°), the combination of which (Sectors 4+5+6) reproduces the full wheat sector (133–283°). Sector 7 (283–313°) corresponds to the second part of ditch, centred around ~298°, and its riparian margin. Finally, Sector 8 (313–15°) corresponds to the permanent grassland. Sector 5 is centred on the wake of the sonic anemometer, whose peak is visible at 193° (Fig. 4). Although this wake does not directly bias CO_2 fluxes, it locally distorts the wind field in a way that renders the standard u_* filter (< 0.1 m s^{-1}) unreliable within this directional range. Retaining all eight sectors, was necessary to reduce the imbalance in sample size across sectors and sub-periods: as shown by the associated number of 10-minute observations (Table 1), the eight-sector scheme, together with the pre-/post-harvest temporal split, yields sub-groups ranging from a few hundred to several thousand observations, a spread that a coarser

classification would only widen further. In particular, subdividing the wide wheat sector (133–283°) into three narrower sub-sectors (Sectors 4, 5 and 6) allows each of them to be sampled with a comparable, wind-direction-driven density of observations, avoiding a single, very large wheat class from dominating the dataset relative to the smaller mixed-grain, hybrid and ditch sectors (Table 1); this subdivision also makes it possible to assess whether the wheat parcel behaves consistently across its angular extent, or whether spatial heterogeneity remains within a single, apparently homogeneous field. Sectors 3 and 7 (the two ditch/riparian corridors) were retained for a similar diagnostic purpose: rather than providing a stand-alone carbon-flux estimate for the ditches themselves, they serve as a point of comparison for the neighbouring mixed-grain (Sector 2) and grassland (Sector 8) sectors, helping to assess the extent to which riparian and open-water surfaces may contribute to, or "contaminate," the flux signal nominally attributed to these two crop and grassland sectors (see Sect. 3.3, 4.2).

In the present study, the footprint (Fig. 2a) was estimated with the Kljun et al. (2015) model and calculated using the constant measurement height (Zm=3m), the constant displacement height (d = 0.1 m), mean wind velocities, standard deviations of the lateral velocity fluctuations after rotation, the Obukhov length (L), friction velocities (u_*) and wind directions, measured by the station from February, 1st 2023 to 31 January 2024. No further specific parametrization like aerodynamic canopy height for instance (Chu et al., 2018) was done since averaged footprint we present was sufficient to our modelling approach. Footprint model distance outputs using Kljun et al. (2004) and Kormann and Meixner (2001) footprint models given by EddyPro, were also checked to help estimate whether the ditch's contribution was negligible or not.

NDVI time series were derived from 26 MODIS MOD13Q1.061 Vegetation Indices 16-Day composite products (Terra platform, 250 m spatial resolution, tile h17v04), covering the period from 1 January 2023 to 31 January 2024 and obtained from the NASA Land Processes Distributed Active Archive Center (LP DAAC) via the NASA Earthdata Search portal. To analyse and better understand vegetation dynamics according to phenology and agricultural management, three well-defined "pure" agricultural parcels were delineated, and for each of them, a single NDVI value per date was computed as the spatial mean of all MODIS pixels falling within the corresponding polygon, thereby representing the vegetation status of the parcel as a whole. Riparian vegetation located along the two canals as part of the EC footprint in certain sectors was explicitly excluded from the NDVI-based analysis.

2.3.2 Modelling Scenarios and Sector-Specific Approach

To complete and fully characterize the spatio-temporal dynamics of NEE measured by EC over the February 2023 - February 2024 period across the study site, RF models were implemented according to two distinct strategies. The first strategy (Scenario 1) involved training a single global RF model using pooled data from all sectors (0–360°), representing the standard approach when sector information is not exploited; this global model served as a baseline against which the added value of the sector-specific strategy could be quantified, both in terms of point-prediction performance and reconstructed annual cumulative NEE (Sect. 3.2).

In the second strategy (Scenario 2), separate RF models were trained for each landscape-based wind sector classification area and key temporal subdivision, accounting for phenological shifts such as crop harvest for instance that influence flux dynamics over the wheat and mixed-grained areas. Specifically, for each crop parcel, pre- and post-harvest periods were modelled independently to better capture ecosystem transitions, resulting in seven sector-specific models overall (Table 1). Both strategies used the same explanatory variables described in Sect. 2.3.3 to ensure comparability and comparison.

Table 1. Sector definition and temporal subdivision associated with the number of observations recorded each 10 minute from 01/02/2023 to 31/01/2024.

Sector	Land cover	Wind direction range (°)	Temporal split	Split criterion	Model ID	Number of observations	Percentage of gap-filled data
1	Hybrid (grassland/mixed grain transition)	15-70	Yes	24 May	Sector 1 - before Sector 1 - after	2981 3273	88%
2	Mixed grain	70-103	Yes	24 May	Sector 2 - before Sector 2 - after	1130 2307	94%
3	Ditch	103-133	No	-		1207	98%
4	Wheat	133-173	Yes	7 July	Sector 4 - before Sector 4 - after	625 1914	95%
5	Wheat	173-213	Yes	7 July	Sector 5 - before Sector 5 - after	1031 1563	95%
6	Wheat	213-283	Yes	7 July	Sector 6 - before Sector 6 - after	4931 7940	76%
7	Ditch	283-313	No	-		2207	96%
8	Grassland	313-15	No	-		3729	93%
4+5+6	Wheat (combined)	133-283	Yes	7 July	Sector 4+5+6 - before Sector 4+5+6 - after	6587 11417	66%

2.3.3 Selection variables for Random Forest modelling

Although variable selection procedures such as Variable Selection Using Random Forest (VSURF) (Genuer et al., 2010) are commonly used to reduce dimensionality prior to RF modelling, we did not apply automated selection in the present study. Instead, we opted for a targeted set of explanatory variables a priori chosen based on their established mechanistic links with CO_2 fluxes and their relevance in surface-atmosphere exchange studies. All models were therefore trained using the same set of predictors: photosynthetically active radiation (PAR), day of year (DOY), a qualitative day-night indicator, relative air humidity, soil moisture, soil temperature, wind direction, and air temperature. PAR directly regulates photosynthetic activity and light-driven variations in ecosystem CO_2 uptake; DOY captures seasonal changes in vegetation state and radiation regime. The qualitative day-night indicator distinguishes periods of daytime and nighttime activity by assigning "day" to time steps with PAR > 10 μmol m^{-2} s^{-1} and "night" to those with PAR ≤ 10 μmol m^{-2} s^{-1}, and exhibit in turn contrasting drivers of respiration and assimilation. Relative humidity and air temperature influence stomatal conductance and microbial respiration, while soil moisture and soil temperature jointly modulate below-ground respiration and substrate availability. Wind direction was included to account for directional heterogeneity in source areas. The importance of these variables chosen and checked for each model are provided in the Appendix B (see Sect. 2.3.4 below and Figs. B2, B3). Explicit seasonal or monthly cyclical time variables though useful in some modelling contexts (Vekuri et al., 2023), were not included here since the main temporal dynamics were already well captured by PAR, DOY and the day/night indicator, and would therefore provide largely redundant information. Moreover, the objective was to keep the model as simple and interpretable as possible, especially since RF models are already able to handle non-linear temporal patterns without requiring explicit seasonal transformations. Similarly, NDVI observations (Sect. 2.3.1) could not be directly used as predictors in the RF model, which was trained on footprint-integrated flux measurements rather than pure-field vegetation indices. At the end, this expert-driven approach ensured that all major biophysical drivers of NEE were represented while maintaining comparability across sectors and modelling exercises.

It should also be clarified how wind direction, one of the meteorological predictors, was assigned when reconstructing a continuous NEE time series for each sector. All meteorological variables come from the same single EC tower, whatever the wind direction. Each sector-specific RF model was trained and evaluated only on the time steps when the wind actually blew from within that sector's own angular range. To reconstruct NEE continuously for a sector, however, predictions were also needed for time steps when the wind blew from elsewhere, and the RF model still required a wind-direction value as input to produce a prediction. For these time steps, a wind-direction value was drawn at random from within the sector's angular range, while all other meteorological variables (PAR, air and soil temperature, humidity, etc.) were taken directly from the actual tower measurements at that time step.

2.3.4 Model training, performance evaluation and uncertainty quantification

Two complementary modelling scenarios were implemented to reconstruct NEE values across the different wind sectors. In the first scenario (global model), a single Random Forest (RF) model was trained on the entire dataset, pooling observations across all wind directions (0–360°) without any sector- or period-specific stratification. In the second scenario (sector-specific models), independent RF models were trained separately for each land-cover sector, and, where relevant, for each sub-period defined by a phenological or management transition (e.g., harvest date; Tables 1 and B1), using only the data associated with wind directions falling within the corresponding sector's angular range.

Both scenarios shared an identical RF parameterization. The number of trees was fixed at ntree = 300, consistent with previous evidence that RF performance stabilizes well below this threshold (Breiman, 2001), and the number of predictors sampled at each split (mtry) followed the standard heuristic for regression trees, mtry = $\lfloor\sqrt{p}\rfloor$, where p is the number of candidate predictors (PAR, air temperature, relative humidity, wind direction, day of year, soil temperature, soil moisture content, and a day/night indicator; p = 8, yielding mtry = 2).

Uncertainty was quantified identically for both scenarios too, using a bootstrap resampling procedure as providing a more robust basis for evaluating differences among vegetation types and modelling approaches, particularly during disturbance and post-harvest periods when fluxes are highly variable and non-stationary. For each retained model, B = 100 RF models were trained on independent bootstrap resamples of the corresponding training dataset (sampling with replacement, resample size equal to the original dataset size), using the same ntree and mtry values throughout. The resulting ensemble of 100 bootstrap models was used to generate, for each time step, a distribution of predicted CO_2 fluxes: the median of this distribution was retained as the point estimate, and its 2.5^{th} and 97.5^{th} percentiles defined an empirical 95% prediction interval. Model performance (RMSE, MAE, R^2) was assessed using out-of-bag (OOB) predictions: for each observation, predictions were derived only from the subset of trees, within the ntree = 300 trees of each RF model, for which that observation had not been included in the corresponding internal bootstrap training sample, following the standard OOB procedure inherent to the Random Forest algorithm (Breiman, 2001). This OOB evaluation was applied consistently across the B = 100 bootstrap-resampled models of each scenario, providing an internal, unbiased estimate of predictive performance for both the global and sector-specific strategies, and forming the basis of the OOB validation framework used to compare the two strategies in Sect. 3.2 (Table 2).

For flux reconstruction, continuous predictions were generated separately for each wind sector, using only the wind-direction range corresponding to that sector; whenever the observed wind direction fell outside this range, wind direction was randomly resampled within the sector's angular boundaries to preserve a consistent input space, while all other meteorological predictors were retained from the tower measurements. In the sector-specific scenario, each sector was predicted using its own dedicated bootstrap-ensemble model (and, where applicable, the model corresponding to the relevant sub-period). In the global-model scenario, the same bootstrap-ensemble model, trained on the pooled 0–360° dataset, was applied uniformly to generate predictions for every sector. This design allowed a direct comparison between sector-specific and global reconstructions of CO_2 fluxes, together with their associated uncertainty, across the different wind sectors.

2.4 NEE flux partitioning

Net Ecosystem Exchange (NEE) reconstructed by the sector-specific RF models (Sect. 2.3.4) was further partitioned into Gross Primary Production (GPP) and Ecosystem Respiration (Reco) for each sector and for the combined wheat sector (Sectors 4+5+6). Following Kowalski et al. (2003, 2004), daytime NEE (NEEd) can be expressed as a curvilinear (rectangular hyperbolic) function of the incident photosynthetic photon flux density (PAR):

$$NEEd = \frac{a1 - PAR}{(a2 + PAR) - Reco}$$

where a1 is the maximum photosynthetic uptake rate at light saturation (µmol m^{-2} s^{-1}), a2 is the PAR value at which photosynthetic uptake reaches half of a1 (µmol m^{-2} s^{-1}), and Reco is the mean ecosystem respiration rate (µmol m^{-2} s^{-1}) over the time window used to fit the relationship. In the present study, the three parameters (a1, a2, Reco) were estimated by non-linear least-squares regression of daytime NEE (PAR > 10 µmol m^{-2} s^{-1}, following the day/night indicator defined in Sect. 2.3.3) against PAR, applied independently at a daily scale for each sector-specific RF reconstruction. This daily-resolution fitting allows Reco, a1 and a2 to vary from day to day, thereby capturing short-term changes in photosynthetic capacity and respiration associated with phenological development, senescence, harvest and other management events, rather than assuming stationary parameters over longer periods. For each day and sector, Reco obtained from the fit was taken as the estimate of ecosystem respiration for that day (assumed constant over the 24 h cycle), and daily GPP was then derived as the residual between the reconstructed daily NEE and the fitted Re: GPP = NEE − Reco (with GPP negative, by convention, indicating net carbon uptake by vegetation, consistent with the sign convention adopted in Sect. 2.2.1). Daily Reco and GPP estimates were then summed to obtain cumulative annual Reco and GPP for each sector (Sect. 3.4, Table 3, Fig. 9).

3 Results

3.1 Site condition, measured NEE dynamic and phenology of the studied land parcels

Over the full monitoring period (1 February 2023 – 31 January 2024), the EC system recorded continuous atmospheric CO_2 exchanges together with the associated meteorological and soil variables (Fig. 5). NEE followed a clear seasonal pattern with studied site behaving predominantly as a carbon sink from February to the end of May, reflecting the progressive development of crop and grassland canopies under increasing PAR and air temperature values (Figs. 5a, b, d). A pronounced minimum in NEE (i.e. maximum net uptake) occurred around mid-May, coinciding with peak vegetative development across the three underlying land uses (wheat, grassland, mixed grain, Fig. 6). From mid-May onward, NEE values progressively shifted toward a source behaviour, associated with crop senescence, harvest managed operations and increasing soil respiration under warmer, drier conditions (Fig. 5c–g), before stabilising into a moderate source regime during autumn and winter. A distinct disturbance

was recorded in November 2023, when an intense storm produced a sharp rainfall peak and an abrupt increase in soil moisture (Fig. 5a, 5e, f; red box), coinciding with large-amplitude, high-frequency fluctuations in the measured NEE (see Sect. 2.2.2 and Appendix A, Fig. A2). Wind direction (Fig. 5h) confirmed that the tower sampled all eight defined sectors (Table 1) over the year, although unevenly, consistent according to the prevailing wind regime at the site.

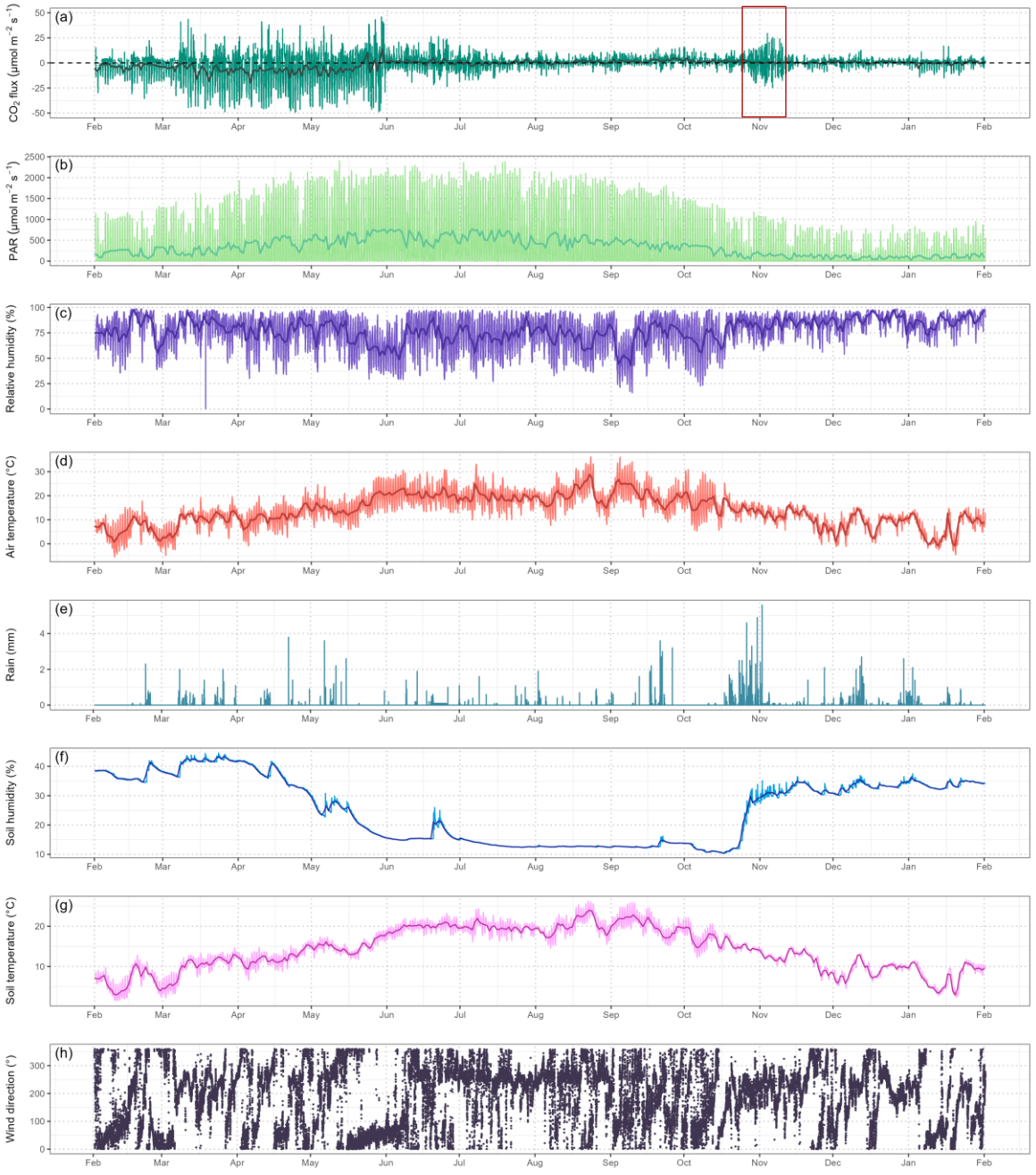


Figure 5. Atmospheric Eddy Covariance measured time series from 01 February 2023 to 31 January 2024 of (a.) atmospheric CO_2 exchanges NEE, (b.) Photosynthetically Active Radiation (PAR), (c.) relative humidity, (d.) air temperature, (e.) rain, (f.) soil humidity, (g.) soil temperature and (h.) wind direction. For CO_2 fluxes, PAR, relative humidity, air temperature, soil humidity and soil temperature, the continuous line represents the daily average obtained from the 10 min values. CO_2 flux values shown in Fig. 5 correspond to EC data without hard limits method filtering as presented for illustrative purposes only. The red box highlights the November storm period.

The Sentinel-2 NDVI time series (Fig. 6) illustrated contrasting phenological trajectories among the three underlying land uses. Both the wheat and mixed-grain parcels showed a steady NDVI increase from late winter, reaching maximum canopy greenness in spring. The mixed-grain parcel exhibited an abrupt NDVI decline following its harvest on 24th May 2023, after which the field remained largely bare until autumn re-sowing that produced a renewed NDVI increase. The wheat parcel maintained high NDVI values later into the season, consistent with its later phenological cycle, and showed a comparatively smaller NDVI discontinuity at harvest because the crop was already senescent and dry at that stage. The permanent grassland parcel displayed the smallest seasonal amplitude, with NDVI values remaining moderate and relatively stable year-round, consistent with continuous grazing and the absence of a single, abrupt senescence event. One can notice that these independent phenological trajectories, derived from well-defined "pure" parcel polygons that exclude riparian vegetation (Sect. 2.3.2), provide an external, remote-sensing-based validation of the crop calendars used to define the pre-/post-harvest temporal subdivisions of the sector-specific models (Table 1).

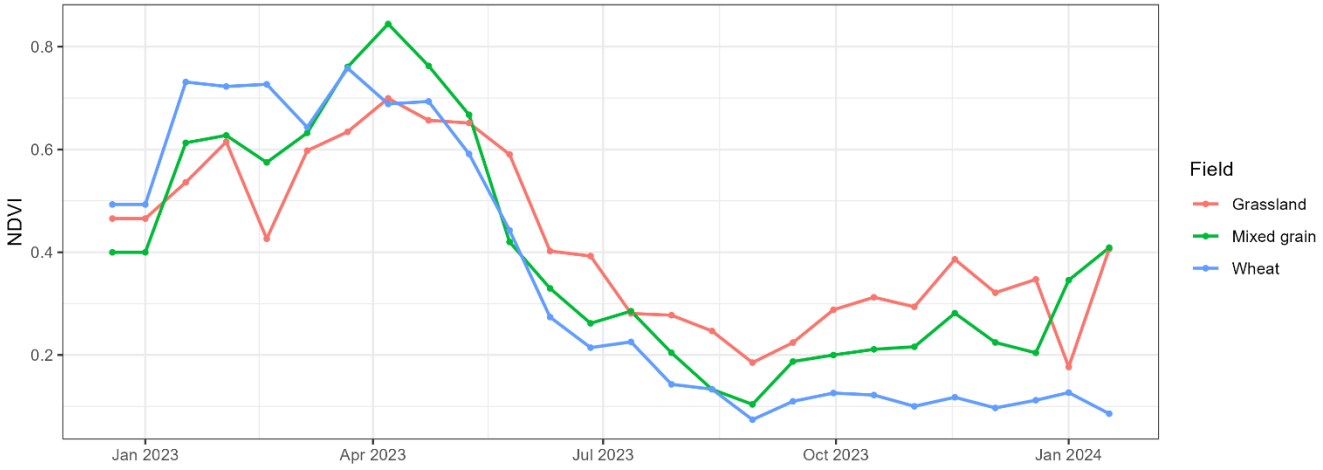


Figure 6. Temporal evolution of MODIS-derived NDVI for the three study fields from January 2023 to January 2024.

3.2 Global versus sector-specific RF modelling: point performance and annual flux reconstruction

The two modelling strategies (Scenario 1, global; Scenario 2, sector-specific) were first compared in an out-of-bag (OOB) validation framework, in which both models were evaluated on exactly the same observations for each of the 13 sector × period sub-groups (n = 34,838 in total; Table 2). Point-prediction performance was overall very similar between the two strategies: R^2 differed by less than 0.02–0.03 and RMSE by less than 5% for most sub-groups (Table 2). The global model nonetheless showed a systematically narrower 95% confidence interval (IC95) than the corresponding sector-specific model in every sub-group, with a reduction of 5 to 16% depending on the sub-group, reflecting the much larger effective bootstrap training sample available to the global model (n ≈ 34,800 vs. 600–8,000 for individual sector-specific models). The sign of the point-performance difference between strategies depended on sub-group sample size: for the most heavily sampled sub-groups (n > ~3000; e.g. Sector 6-after, Sector 8, Sector 1-after), the sector-specific model matched or marginally outperformed the global model, whereas for smaller sub-groups, the global model showed a modest advantage (e.g. R^2 + 0.06 for Sector 2-after and + 0.07 for Sector 4-after relative to the sector-specific model). R^2 varied widely among sub-groups (0.20–0.94; Table 2), but this range was driven primarily by differences in the amplitude of observed flux variability rather than by absolute model error, which remained comparatively stable across sub-groups (RMSE ≈ 1.5–2.6 µmol m^{-2} s^{-1} for the more stable post-harvest periods and ≈ 3–4.6 µmol m^{-2} s^{-1} for the higher-variability pre-harvest periods).

Table 2. Out-of-bag (OOB) performance of the sector-specific and global RF models for each of the 13 sector × period sub-groups, evaluated on the same held-out observations (n, sample size; Mean and SD, mean and standard deviation of observed flux, µmol m^{-2} s^{-1}; RMSE, root-mean-square error, µmol m^{-2} s^{-1}; IC95, width of the bootstrap 95% confidence interval, µmol m^{-2} s^{-1}).

Model ID	Observations			RMSE	RMSE	R^2	R^2	IC width	IC width
	Number	Mean	Sd	sector	global	sector	global	sector	global
Sector 1 - before	2981	-5.12	10.75	2.98	2.97	0.92	0.92	2.71	2.53
Sector 1 - after	3273	0.97	4.7	2.05	2.09	0.81	0.81	1.66	1.55
Sector 2 - before	1130	-5.99	9.46	2.35	2.24	0.94	0.94	2.41	2.23
Sector 2 - after	2307	0.73	3.11	1.79	1.63	0.67	0.73	1.5	1.4
Sector 3	1207	-1.24	5.64	2.61	2.48	0.79	0.81	2.39	2.09
Sector 4 - before	625	-5.10	11.37	3.73	3.46	0.89	0.91	3.81	3.21
Sector 4 - after	1914	1.20	2.02	1.57	1.53	0.39	0.43	1.30	1.24
Sector 5 - before	1031	-5.14	10.79	4.63	4.56	0.82	0.82	4.12	3.64
Sector 5 - after	1563	0.95	2.18	1.95	1.96	0.20	0.20	1.63	1.42
Sector 6 - before	4931	-6.04	13.54	4.10	4.11	0.91	0.91	3.26	3.10
Sector 6 - after	7940	0.84	2.35	1.62	1.63	0.53	0.52	1.28	1.21
Sector 7	2207	-0.93	5.59	2.38	2.26	0.82	0.84	1.90	1.69
Sector 8	3729	-0.59	7.03	2.24	2.30	0.90	0.89	1.80	1.74

This apparent equivalence between the two strategies did not hold, however, once fluxes were reconstructed continuously over the full annual cycle for each sector (i.e. annual cumulative NEE, Fig. 7). Both strategies reproduced the same qualitative

seasonal pattern (a growing sink from February to late May, a marked minimum around June, and a progressive return to source behaviour in late summer/autumn), but diverged markedly in amplitude and relative sector ranking. Under the sector-specific strategy, annual cumulative NEE values ranged from +259 gC m^{-2} (Sector 1, Hybrid) to −370 gC m^{-2} (Sector 6, Wheat), with an inter-sector range of 629 gC m^{-2}. Under the global strategy, the same sectors ranged from +164 gC m^{-2} (Sector 1) to −227 gC m^{-2} (Sector 7), a range of 391 gC m^{-2}, corresponding to an overall compression of about 38% in the amplitude of inter-sector variability (Fig. 7). This compression was uneven across sectors: comparatively small for Sector 8 (grassland, −15%) and Sector 6 (wheat, the best-sampled sector, −45%), but much stronger for Sector 4 (−70%), Sector 5 (−61%) and Sector 2 (mixed grain, −59%) (the less-sampled sub-sectors). Sector 7 (Ditch East) was a clear exception: rather than being compressed, its cumulative NEE value was amplified under the global strategy (−187 to −227 gC m^{-2}), to the point of inverting its ranking relative to Sector 6 between the two strategies. Taken together, these results show that the global and sector-specific models, despite near-equivalent OOB point performance, yield substantially different spatial pictures of the site's NEE values cumulated (extrapolated) over the full annual cycle for each sector. Because the present site combines strong sub-field heterogeneity in land cover, phenology and management, and because the objective of this study was precisely to resolve field-scale NEE contrasts, the global model was not retained for the remainder of the analysis: all results and interpretations presented hereafter (Sects. 3.3–3.4) are then based on the sector-specific (Scenario 2) reconstructions (see Sect. 4.1 for further discussion).

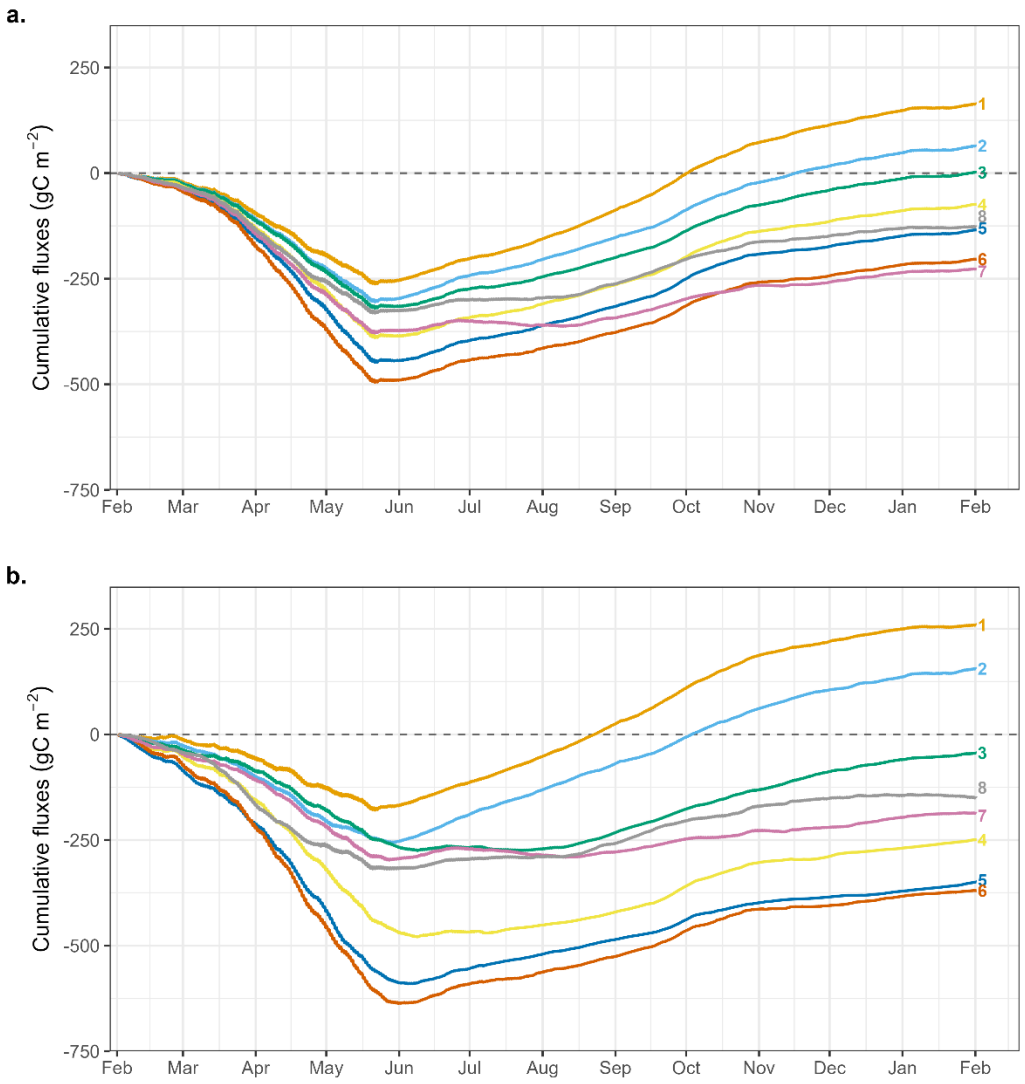


Figure 7. Annual cumulative CO_2 fluxes (cumulative NEE) reconstructed for each of the eight sectors (Fig. 2a) with the global strategy (a, Scenario 1) and the sector-specific strategy (b., Scenario 2).

3.3 Uncertainties associated with sector-specific models and flux footprint

Uncertainty in the sector-specific reconstructions, quantified through the bootstrap 95% confidence interval (IC95) built from B = 100 resampled RF models (Sect. 2.3.4), varied significantly among sectors and between the pre- and post-harvest periods (Tables 1, 2, Fig. 8). Confidence intervals were systematically wider before harvest/mowing (IC95 = 2.41–4.12 µmol m^{-2} s^{-1}) than after (IC95 = 1.28–1.63 µmol m^{-2} s^{-1}), reflecting both the higher intrinsic flux variability during the active growth period and, for several sectors, smaller pre-harvest sample sizes (i.e., n = 625 for Sector 4-before vs. n = 1914 for Sector 4-after).

Among the post-harvest/stable periods, the narrowest intervals were obtained for the wheat sub-sectors 4 and 6 (IC95 = 1.28–1.30 µmol m^{-2} s^{-1}), and the widest for Sector 3 (IC95 = 2.39 µmol m^{-2} s^{-1}) and Sector 7 (IC95 = 1.90 µmol m^{-2} s^{-1}).

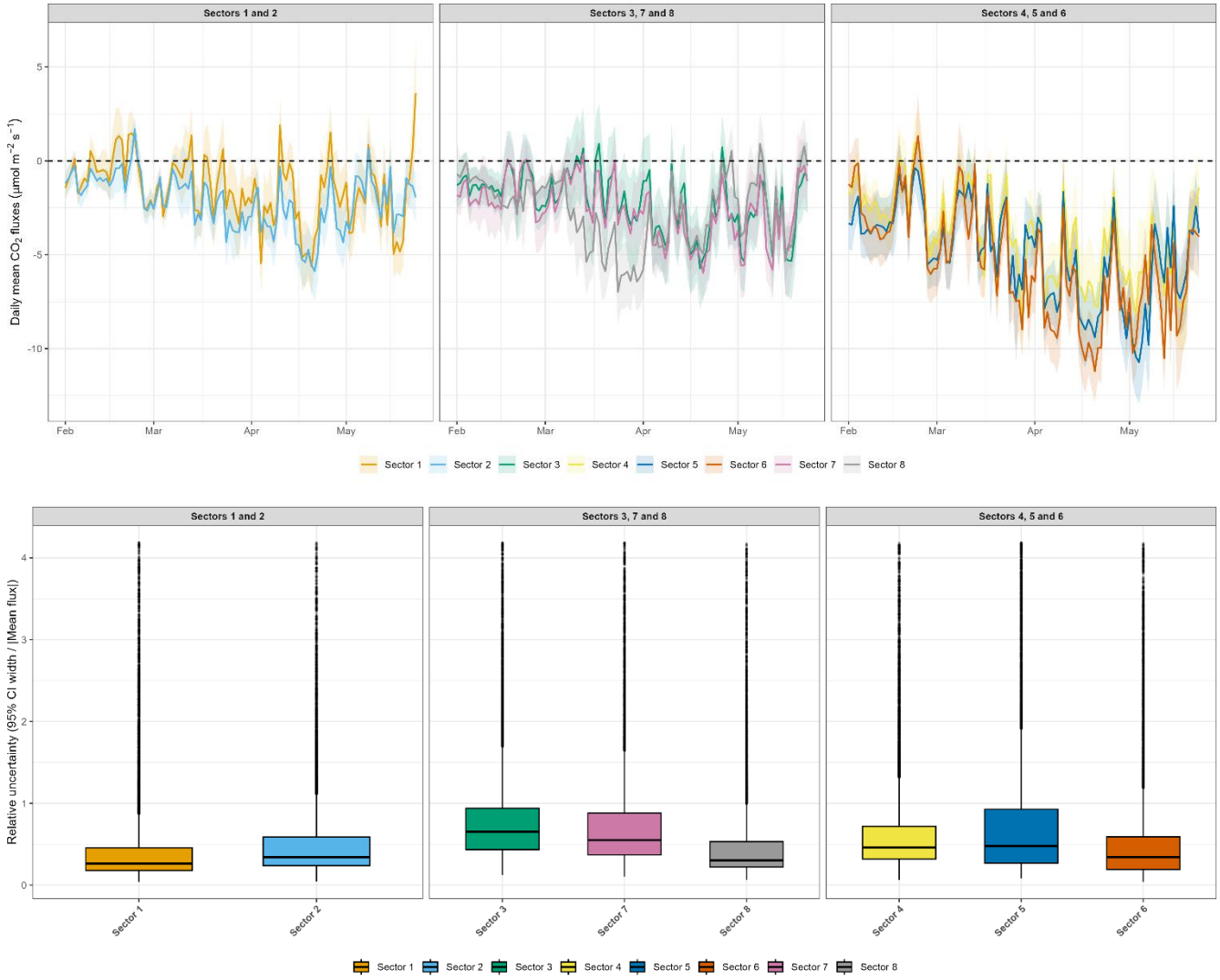


Figure 8. Reconstructed NEE dynamics and bootstrap 95% confidence intervals for each sector, before harvest/mowing (1 Feb.–24 May) (above) and distribution of boostrap 95% confidence interval width before harvest/mowing (below).

This pattern reflects both statistical sampling effects and the spatial structure of the flux footprint relative to the position of the EC tower; the implications of this footprint geometry for the interpretation of each sector's cumulative NEE are examined in detail in the Discussion (Sect. 4.2).

3.4 NEE Partitioning into GPP and ecosystem respiration (R_{eco})

Partitioning of the sector-specific RF-reconstructed NEE into GPP and Reco, following Kowalski et al. (2003, 2004) (Sect. 2.4), was performed for each of the eight sectors and for the combined wheat sector (Sectors 4+5+6) (Table 3, Fig. 9). Across all sectors, annual GPP ranged from −1037 gC m⁻² (Sector 5) to −1601 gC m⁻² (Sector 8, grassland), and annual Reco from 684 gC m⁻² (Sector 5) to 1700 gC m⁻² (Sector 1) (Table 3).

The three wheat sub-sectors (4, 5, 6) shared a broadly similar phenological signature, i.e. moderate GPP (−1037 to −1419 gC m⁻²) combined with the lowest Reco values of the whole dataset (684 to 1050 gC m⁻²), consistent with a single, well-defined crop cycle sampled without interference from adjacent land covers. When combined, Sectors 4+5+6 (i.e. the full wheat field, 133–283°) yielded an annual NEE of −308 gC m⁻², GPP of −1298 gC m⁻² and Reco of 990 gC m⁻² (Table 3); these combined values fall, as expected, within the range spanned by the three individual sub-sectors, and are closest to Sector 6 (the widest and best-sampled sub-sector, contributing the largest share of observations to the combined estimate; Table 1), while Sector 5 (the sub-sector centred on the sonic-anemometer wake) showed the most negative NEE and the lowest Reco of the three, and Sector 4 the least negative NEE. This internal consistency between the combined estimate and its three constituent sub-sectors supports the reliability of the wheat-field carbon budget as a whole (see also Sect. 3.3, 4.2).

By contrast, Sectors 1 (hybrid transition zone), 2 (mixed grain) and 8 (grassland), despite all lying within angular ranges vegetated for most of the year, displayed markedly different GPP/Reco balances. Sector 8 (grassland) combined the strongest gross uptake of all sectors (GPP = −1601 gC m⁻²) with a comparatively moderate Reco (1456 gC m⁻²), resulting in the second most negative cumulative annual NEE of the whole site (−145 gC m⁻²) after the wheat sectors. Sector 2 (mixed grain) showed a markedly lower GPP (−1141 gC m⁻²) combined with a high Reco (1292 gC m⁻²), reflecting the loss of photosynthetic activity after the early harvest (24 May) and the subsequent bare-soil respiration; overall, Sector 2 remained a modest net source (+151 gC m⁻²). Sector 1 (hybrid) showed both a high gross uptake (GPP = −1505 gC m⁻², closer to that of the grassland than of the mixed-grain sector) and, simultaneously, the highest Reco of all eight sectors (1700 gC m⁻²), so that despite a photosynthetic capacity comparable to Sector 8, its respiratory term outweighed it and Sector 1 remained the largest net source of the whole site (+195 gC m⁻²). This contrast between Sectors 1, 2 and 8, three sectors adjoining the grassland/mixed-grain area yet showing distinctly different NEE, GPP and Reco combinations , illustrates that superficially similar land-cover labels do not guarantee similar carbon dynamics once the underlying GPP/Reco decomposition is examined (see Sect. 4.3).

Sectors 3 and 7, both defined to isolate the two ditch/riparian corridors (Sect. 2.3.1), occupied an intermediate position relative to the other sectors: their GPP (−1147 and −1295 gC m⁻², respectively) and Reco (1115 and 1114 gC m⁻²) were both closer to the mixed-grain and hybrid values than to the wheat or grassland ones, and both sectors were net, moderate sinks (−32 and −181 gC m⁻², respectively). Given their position immediately adjacent to Sectors 2/1 and 8 (Sect. 3.3), this intermediate GPP/Reco signature is consistent with a footprint integrating both riparian vegetation and edge contributions from the neighbouring crop and grassland fields (see Sect. 4.3).

Table 3. Annual cumulative estimates of NEE, GPP and Reco (gC m^{-2} yr^{-1}) obtained from the Kowalski et al. (2003, 2004) partitioning, for each sector and for the combined wheat sector (Sectors 4+5+6).

	Sector 1	Sector 2	Sector 3	Sector 7	Sector 8	Sector 4	Sector 5	Sector 6	Sector 4+5+6
NEE	195	151	-32	-181	-145	-256	-353	-366	-308
GPP	-1505	-1141	-1147	-1295	-1601	-1056	-1037	-1419	-1298
Reco	1700	1292	1115	1114	1456	800	684	1050	990

535

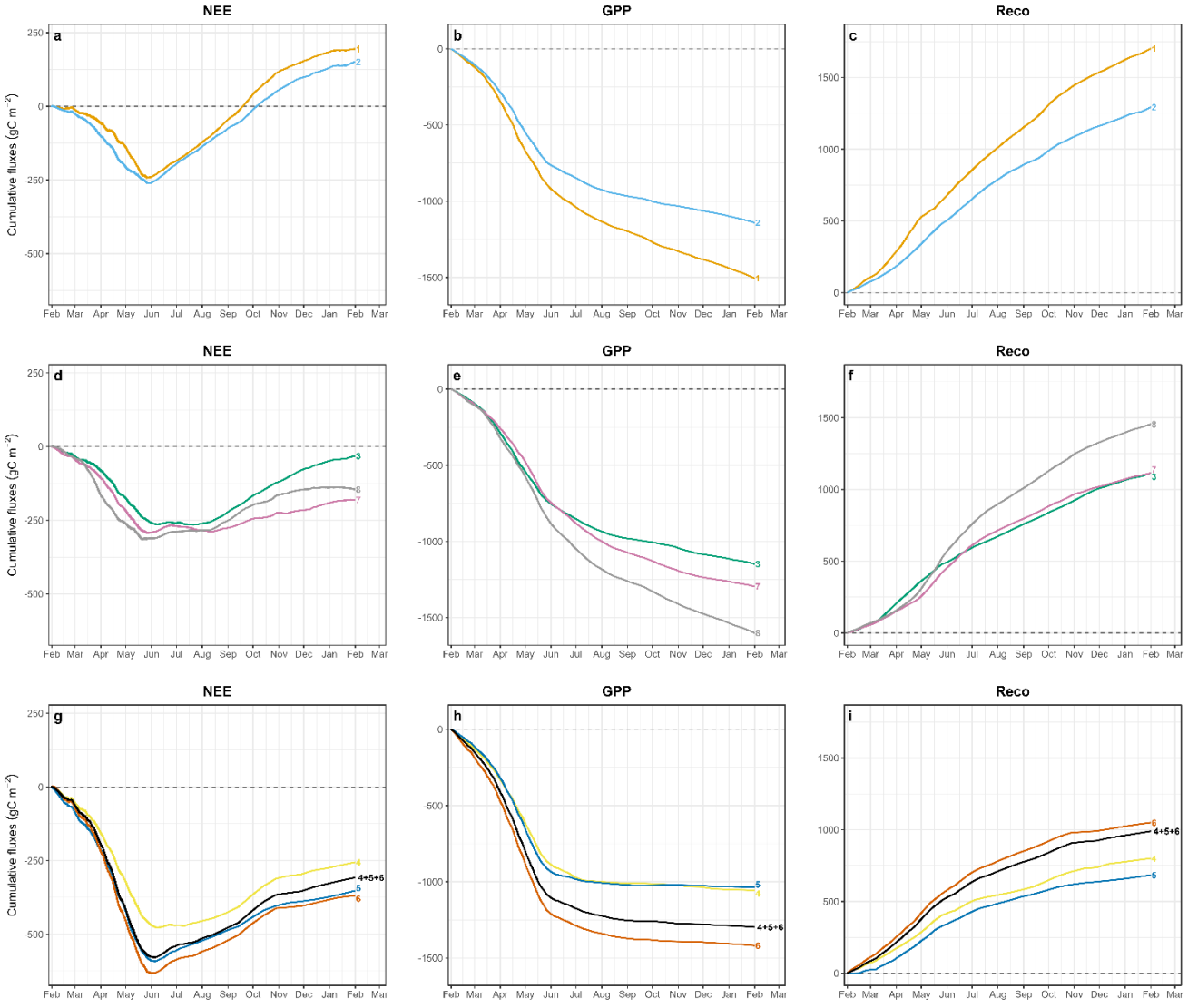


Figure 9. Cumulative estimates of NEE, GPP and Reco for each of the eight sectors, obtained from the Kowalski et al. (2003, 2004) partitioning of the sector-specific Random Forest flux reconstructions.

The consistency between this GPP/Reco-based partitioning and the direct RF-based cumulative NEE reconstructed in Sect. 3.2 is examined in the Discussion (Sect. 4.3), together with the implications of the sector-specific phenological and management-driven contrasts identified above.

4 Discussion

4.1 Why sector-specific modelling is required for a heterogeneous single-EC tower site set-up?

The near-equivalence and occasional slight advantage of the global RF model in the OOB point-prediction exercise (Sect. 3.2, Table 2) is best understood as a manifestation of the classical bias-variance trade-off inherent to ensemble tree methods (Breiman, 2001). In OOB validation, both strategies are evaluated only on wind directions that were actually sampled within each sector, i.e. an interpolation exercise within a domain of environmental conditions already encountered during training. In this regime, the global model benefits from a much larger effective training sample ($n \approx 34{,}800$ vs. 600–8,000 for individual sector-specific models), which mechanically narrows its bootstrap confidence intervals (Table 2) and allows the smallest, least-sampled sectors to "borrow" information from the wider dataset, a form of implicit partial pooling. Because wind direction remains an explanatory variable in the global model, the ensemble is, in principle, able to recover part of the sectoral structure from this proxy variable alone; but with mtry = $\lfloor\sqrt{8}\rfloor = 2$, wind direction competes with seven other predictors for selection at each split and is therefore only partially and diffusely "learned" by the global ensemble.

However, this apparent equivalence broke down once fluxes were extrapolated over the full annual cycle for each sector (Sect. 3.2, Fig. 6) — an exercise that, by construction, requires the model to predict flux behaviour for combinations of PAR, air temperature, day-of-year and soil conditions for which each specific sector is often poorly represented in the training data (i.e. a genuine extrapolation regime rather than interpolation). In this regime, the advantage of pooling reverses: when a particular combination of environmental drivers is rare for a given target sector but frequent for the dominant sector in the pooled dataset (Sector 6/wheat, representing the largest single share of the dataset), the corresponding leaves of the global RF trees become populated mainly by the dominant sector, so that the global model's prediction for the target sector regresses toward the pooled, dominant-sector response — a shrinkage effect conceptually related to partial pooling in hierarchical models and, in its net numerical consequence, to Simpson's-paradox-type aggregation bias. This mechanism is consistent with the observation that the least-sampled sectors (2, 4, 5) showed the strongest compression of their annual NEE values under the global strategy (−59 to −70%; Sect. 3.2), while the best-sampled sector (Sector 6) showed the smallest compression (−45%). Sector 3, whose sector-specific NEE was already close to zero (−44 gC m^{-2}), showed an even more extreme manifestation of this shrinkage, with its cumulative NEE collapsing to near zero (+2 gC m^{-2}) under the global strategy rather than being merely compressed in magnitude. The exceptional amplification observed for Sector 7 indicates that this shrinkage is not a simple function of sample size alone, but also depends on the proximity, in covariate space, between the target sector and the dominant sectors that "contaminate" its predictions — a proximity that can act in either direction depending on the local structure of the fitted trees.

Ecologically, this divergence is consistent with the fact that day-of-year is only an imperfect proxy for phenological stage: at a given day-of-year, the wheat, mixed-grain and grassland covers can be at very different phenological stages, and a global model that only weakly resolves sector identity through wind direction will tend to average out these physiologically distinct crop responses. This supports the conclusion, drawn in Sect. 3.2, that a single pooled RF model is not adequate for reconstructing field-scale annual cumulative NEE values at a sub-field-heterogeneous site such as ours, even though it may appear competitive - or even preferable in terms of confidence-interval width - in a standard point-validation exercise as done in the present study. This distinction between the two validation regimes (interpolation vs. extrapolation) has a broader methodological implication for flux-partitioning studies relying on machine-learning gap-filling at heterogeneous sites: OOB or cross-validation metrics computed only on already-sampled conditions may substantially under-estimate the practical cost, in terms of annual NEE bias, of pooling data across ecologically distinct source areas. This finding also reinforces the argument, raised in the Introduction, that resolving spatial heterogeneity within EC footprints - whether through footprint modelling and land-cover attribution (Chen et al., 2011), multi-tower designs (Oren et al., 2006), or alternative statistical approaches (Wang et al., 2016) - remains a prerequisite for deriving physically meaningful, field-scale annual NEE from single-tower EC records over heterogeneous agricultural mosaics. In this respect, the wind-sector RF approach used here, combined with the increasing availability of low-cost EC hardware (Cunliffe et al., 2022; van Ramshorst et al., 2024; Callejas-Rodelas et al., 2025; Rannik et al., 2026), offers a practical, scalable pathway to improve the spatial representativeness of single-tower flux monitoring in mosaic landscapes, without the cost and logistical burden of a full multi-tower network. For this reason, and consistent with Sect. 3.2, the remainder of this Discussion (Sects. 4.2–4.3) focuses exclusively on the sector-specific (Scenario 2) reconstructions.

4.2 Model and experimental uncertainties: the role of the flux footprint and tower position

Beyond the purely statistical uncertainty captured by the bootstrap confidence intervals (Sect. 3.3, Table 2, Fig. 8), the reliability of the sector-specific NEE values derived in this study is fundamentally constrained by the spatial structure of the EC flux footprint and by the fixed position of the tower relative to the surrounding land-cover mosaic (Fig. 2a). Because a single EC system was used, the "sector" attribution of each 10-min flux value is only as reliable as the assumption that the footprint for that wind direction is dominated by the target land cover; this assumption holds to very different degrees across the eight sectors defined in this study.

For the wheat field, sampled through the wide uninterrupted angular range of Sectors 4, 5 and 6 (133–283°), this assumption is comparatively well supported: no ditch, canal or riparian corridor lies within this angular range, and the footprint can reasonably be considered to sample the wheat canopy (and, secondarily, bare or senescent soil after harvest) almost exclusively. Confidence intervals for these three sub-sectors, particularly after harvest (IC95 = 1.28–1.30 µmol m^{-2} s^{-1}; Table 2), are also among the narrowest of the whole dataset. Together, these two independent lines of evidence - a spatially "clean" footprint and comparatively narrow statistical uncertainty - support together the interpretation that the reconstructed NEE, GPP and R_{eco}

values for the wheat field (Sects. 3.2–3.4; Table 3) represent a genuinely well-constrained cumulative annual NEE value for this parcel, one of the more robust outcomes of the present study.

The situation is markedly less favourable for Sectors 1, 2, 3 and 7. Sectors 3 and 7 were explicitly defined to isolate the two ditch corridors and their riparian margins (Sect. 2.3.3); even so, given the narrow width of these features relative to the extent of the EC footprint (Kljun et al., 2015; Kumari et al., 2020), some contribution from the flanking crop or grassland edges cannot be excluded, and, conversely, some contribution from the ditch/riparian vegetation is very likely to persist within the crop-labelled sectors immediately adjacent to it (Sectors 1 and 2). This is consistent with the comparatively wide confidence intervals obtained for Sectors 3 and 7 (IC95 = 2.39 and 1.90 μmol m^{-2} s^{-1}, respectively; Table 2) and with the largest discrepancy found between the RF-based and Kowalski-partitioned cumulative NEE value for Sector 1 ($\Delta \approx 64$ gC m^{-2}, $\approx 25\%$; Sect. 3.4). We therefore consider that, for the mixed-grain (Sector 2) and hybrid transition zone (Sector 1) in particular, the reconstructed yearly cumulative NEE values cannot be interpreted as unambiguous parcel-specific values in the same way as for the wheat sectors: they are best regarded as composite NEE values integrating a variable, and currently unquantified, contribution from riparian vegetation and open-ditch surfaces.

The permanent grassland (Sector 8) occupies an intermediate position. Visually and statistically, its footprint appears comparatively little affected by the adjacent ditch network, and its confidence interval (IC95 = 1.80 μmol m^{-2} s^{-1}) is only slightly wider than that of the wheat sectors, while the discrepancy between its RF-based and partitioned NEE estimates was the smallest of all eight sectors ($\Delta \approx 1$ gC m^{-2}; Sect. 3.4). While this sector's cumulative NEE value cannot be considered as unambiguously "pure" as that of the wheat field, given its position adjacent to the drainage-ditch network, the convergence of these independent lines of evidence suggests that it is comparatively reliable, and markedly less exposed to footprint contamination than the mixed-grain, hybrid or ditch sectors.

Overall, these results indicate that the practical outcome of a single-tower, wind-sector RF approach in a heterogeneous agricultural landscape does not give a uniformly precise annual cumulative NEE value for every land cover, but rather a gradient of confidence directly related to the local footprint geometry: high confidence for spatially "isolated" covers such as the wheat field (and, to a slightly lesser extent, the grassland), and comparatively lower confidence for covers immediately adjacent to linear landscape features such as drainage ditches. Explicit footprint modelling with land-cover attribution (e.g. Kljun et al., 2015; Chen et al., 2011), applied at finer temporal resolution than the sector-averaged approach used here, or complementary plot-scale chamber measurements deployed within specific sectors, would be valuable next steps to further quantify, and potentially correct for, this residual footprint-mixing uncertainty in the mixed-grain, hybrid and ditch sectors.

4.3 Consistency between direct NEE reconstruction and GPP/R_{eco} partitioning

Comparing the yearly cumulative NEE deduced from the GPP/Reco partitioning (Table 3, Sect. 3.4) with the direct RF-based cumulative NEE obtained under the sector-specific strategy (Sect. 3.2, Fig. 7) shows overall good agreement, with sector-dependent discrepancies. For most sectors, the two estimates differ by less than 15 gC m^{-2} (< 10%): Sector 3 (−30 vs. −32 gC m^{-2}), Sector 4 (−244 vs. −256 gC m^{-2}), Sector 5 (−348 vs. −353 gC m^{-2}), Sector 6 (−370 vs. −366 gC m^{-2}), Sector 7 (−187 vs.

−181 gC m⁻²) and Sector 8 (−146 vs. −145 gC m⁻²), and the combined wheat sector shows essentially identical values (−308 gC m⁻² under both approaches). Larger, though still moderate, discrepancies are found for Sector 1 (+259 vs. +195 gC m⁻², a difference of 64 gC m⁻², ≈ 25%) and, to a lesser extent, Sector 2 (+136 vs. +151 gC m⁻², a difference of 15 gC m⁻², ≈ 11%). These two sectors are precisely those identified in Sect. 3.3 and 4.2 as most exposed to footprint mixing with riparian vegetation and ditch surfaces, suggesting that the Kowalski partitioning, which separates fluxes on the basis of a day/night, PAR-driven criterion rather than by land-cover contribution, may be differentially sensitive to the additional respiratory or photosynthetic signal contributed by non-target surfaces within the footprint. This overall convergence between two independent methods (direct RF reconstruction vs. light-response partitioning) provides confidence that the underlying GPP/Reco decomposition (Table 3, Fig. 9) is a robust basis for comparing the carbon-flux dynamics of the different sectors.

Beyond this consistency check, the GPP/Reco decomposition offers a more mechanistic view of why sectors bearing superficially similar land-cover labels can nonetheless display markedly different net carbon budgets (Sect. 3.4). The three wheat sub-sectors (4, 5, 6) shared a common phenological trajectory dominated by a single, well-defined annual crop cycle, so that their individual GPP and Reco values, and those of the combined Sectors 4+5+6, remained close to one another and reflected essentially the same underlying process: a strong, synchronous photosynthetic pulse during spring growth followed by a sharp decline in both GPP and Reco after harvest (7 July). This internal consistency across spatial subdivisions of a single, homogeneous field supports treating Sectors 4+5+6 as a genuine field-scale carbon budget for winter wheat.

By contrast, Sectors 1, 2 and 8 illustrate how similar apparent vegetation cover can conceal contrasting carbon dynamics. Sector 8 (grassland) combined the highest GPP of the whole site with a moderate Reco, consistent with a continuously vegetated, perennial cover subject to grazing rather than harvest, which sustains photosynthetic activity later into the season without the abrupt post-harvest respiration pulse seen in annual crops. Sector 2 (mixed grain), despite sharing an annual-crop phenology broadly similar to wheat before its early harvest (24 May), displayed a much lower GPP and a comparatively high Reco, indicating a shorter and less productive growing period combined with sustained soil respiration on bare ground for most of the year. Sector 1 (the hybrid transition zone between grassland and mixed grain, including riparian vegetation) combined a GPP close to that of the grassland (Sector 8) with the highest Reco of the whole site, i.e. it behaves neither as a "pure" grassland nor as a "pure" cropland, but as a composite signal integrating perennial photosynthetic capacity with elevated, likely riparian- and ditch-driven, respiration; this explains why Sector 1, despite substantial gross uptake, is the largest net carbon source identified in this study (Sect. 3.4).

Sectors 3 and 7 (the two ditch/riparian corridors) further illustrate this gradient: their GPP and Reco values lie between those of the crop/grassland sectors and those of Sector 1, and both remain modest net sinks. Given their position directly adjoining Sectors 2 and 8 respectively (Sect. 2.3.1, 3.3), this intermediate signature is consistent with these sectors sampling a mixture of open-ditch/riparian surfaces and edge contributions from the neighbouring vegetated parcels, rather than a fully homogeneous, non-vegetated ditch surface. Taken together, these results show that the wind-sector RF/partitioning approach, beyond providing annual NEE, GPP and Reco values, allows a differentiated reading of within-site phenological and management-driven contrasts that a single pooled model, or a coarser land-cover classification, would not resolve (Sect. 4.1).

4.4 Annual NEE estimates of the studied parcels and implications for carbon accounting and limitations

Annual cumulative NEE, GPP and Reco values derived from the Kowalski et al. (2003, 2004) partitioning (Table 3, Sect. 3.4) provide the reference carbon-flux estimates discussed here, given their close agreement with the direct RF-based reconstructions for every sector (Sect. 4.3). Considered together, these values reveal contrasting carbon budgets among the four land covers of the site: the permanent grassland (Sector 8) and the winter wheat field (Sectors 4+5+6) acted as net carbon sinks (−145 and −308 gC m^{-2} yr^{-1}, respectively), whereas the mixed-grain field (Sector 2) and the hybrid transition zone (Sector 1) behaved as net carbon sources (+151 and +195 gC m^{-2} yr^{-1}). However, the confidence that can be placed in each of these values differs markedly according to the footprint considerations discussed in Sect. 3.3 and 4.2.

The winter wheat carbon budget can be considered the most robust of the site (Sect. 3.3, 4.2). The combined estimate for Sectors 4+5+6 (NEE = −308 gC m^{-2} yr^{-1}) is well bracketed by its three constituent sub-sectors (NEE from −244 to −370 gC m^{-2}; Sect. 3.4), whose differences mainly reflect the angular width, and therefore the sampling density, each sub-sector received (Table 1) rather than genuine differences in the underlying crop: Sector 6, the widest and best-sampled sub-sector, dominates the combined estimate, while Sector 5 (centred on the sonic-anemometer wake, and thus more strongly affected by the u* filtering constraints described in Sect. 2.3.1) shows the most negative NEE of the three. This internal agreement between the combined wheat estimate and its sub-sectors, together with a footprint essentially free of ditch or riparian contamination (Sect. 3.3), supports a genuinely field-scale wheat carbon budget. This sink strength is broadly consistent with other reported values for winter wheat (Table 4); for instance, Schmidt et al. (2012) found ≈ −270 gC m^{-2} yr^{-1} over two years of EC measurements, and Wang et al. (2013) reported −251 gC m^{-2} for a growing season using EC, in close agreement (within 7%) with automatic-chamber estimates.

The permanent grassland (Sector 8) carbon budget can be considered more or less robust: its footprint (Sect. 3.3) is comparatively little exposed to the drainage-ditch network relative to Sectors 1, 2, 3 and 7, since only the margin of the secondary ditch lies within its angular range, so that the reconstructed budget predominantly reflects genuine grassland photosynthesis and respiration rather than a substantial mixed riparian or aquatic contribution. Its confidence interval (Table 2) is only slightly wider than that of the wheat sectors, and the discrepancy between its RF-based and partitioned NEE estimates is the smallest of all eight sectors (Sect. 4.3). This grassland sink strength compares favourably with other managed, humid or grazed temperate grasslands (Table 4; Jaksic et al., 2006; Senapati et al., 2014), though it should still be interpreted with some residual caution given its position adjacent to the drainage network.

By contrast, the carbon budgets derived for the mixed-grain (Sector 2, NEE = +151 gC m^{-2} yr^{-1}) and hybrid transition zone (Sector 1, NEE = +195 gC m^{-2} yr^{-1}) sectors should be interpreted with considerably more caution. Given the site configuration — both sectors lie directly adjacent to the Brune ditch and/or its secondary branch and their riparian vegetation (Sect. 2.3.1, 3.3) — the corresponding NEE, GPP and Reco values cannot be considered fully representative of the mixed-grain crop, or of the grassland/mixed-grain transition, alone. The exceptionally high Reco recorded in Sector 1 (1700 gC m^{-2} yr^{-1}, the highest

of the whole site; Sect. 4.3) most likely integrates a substantial riparian and/or ditch-adjacent respiratory signal rather than reflecting the transition vegetation alone. We therefore consider that, as constrained by a single EC tower, the mixed-grain carbon budget is not fully acceptable as a stand-alone, field-specific estimate, and recommend that it be interpreted qualitatively (net source behaviour, driven by a short and comparatively unproductive growing period followed by prolonged bare-soil respiration) rather than quantitatively; the hybrid-zone estimate should be treated with even greater caution, as a composite signal rather than a genuine parcel-specific budget (Sect. 3.3, 4.2, 4.3).

Comparing the four sectors together, the wheat and grassland carbon budgets, the two most robust estimates of this study, both point to net carbon sinks broadly consistent with other reported values for comparable crop and grassland systems (Table 4), whereas the mixed-grain and hybrid sectors, whose footprints are more exposed to riparian and ditch surfaces, both appear as net sources of comparable, moderate magnitude. A key factor underlying the contrast between the wheat/grassland and mixed-grain systems, beyond the footprint considerations above, is land-use history. The grassland and wheat plots have been under their respective management for many years, whereas the mixed grain was grassland before the measurement year and reverted to grassland afterward (Guignard and Pouponnot, personal communications); the mixed-grain system was thus a one-off cereal rotation rather than an established crop. Crop-rotation studies indicate that continuous annual cropping tends to reduce productivity, and thus carbon uptake, relative to long-term rotations or perennial cover, owing to changes in soil structure, nutrient cycling and microbial communities (Liebman and Dyck, 1993), and biodiversity/rotation experiments have similarly shown that systems with longer-lived, diverse plant communities achieve greater and more stable productivity than single-year crops (Tilman et al., 2006).

The modest carbon uptake found in the mixed-grain sector in the present study is thus consistent with its transient status: the recent conversion from grassland is associated with lower soil organic matter quality and microbial carbon-use efficiency, which can reduce photosynthetic capacity while still sustaining substantial post-harvest soil respiration, as reported for short-term grassland-to-cropland conversions elsewhere (Mukumbuta et al., 2019).

The established winter wheat field, on the other hand, was almost neutral in full carbon accounting: once grain removal (≈ 250 gC m^{-2} via harvest) is taken into account, the sink behaviour reported here (Sects. 3.2, 3.4) would be largely offset. This agrees with other studies of long-term cereal monocultures (e.g. Schmidt et al., 2011) that find harvest removals can balance the high productivity of wheat. The hybrid sector (mixed-grain/grassland transition, riparian buffer) produced the largest net CO_2 source of the site; its amplified emissions arose from both the large respiration pulse of the mixed-grain crop after harvest and the continued respiration of the perennial grassland and riparian vegetation components integrated within its footprint (Sect. 4.3).

The present study solely focuses on NEE values and harvested biomass, without incorporating all carbon pathways, in particular stocks and associated changes in soil organic carbon and lateral hydrological fluxes, which are particularly important to draw a precise integrated carbon budget of the site. Simultaneous water pCO_2 measurements conducted in November 2023 within the LRTZC project indicated mean water CO_2 emissions of 12.42 ± 14.54 mmol m^{-2} h^{-1} (equivalent to approximately 3.45 ± 4.04 µmol m^{-2} s^{-1}, data not shown) from the ditch surface, which contrast with the cumulative NEE values obtained in the present study for the surrounding parcels; from previous bimonthly water carbon and discharge measurements carried out

in 2017 and 2018 in the different watercourses of the Marais poitevin watershed flowing to the Aiguillon bay, an annual terrestrial carbon export of 15 gC m^{-2} yr^{-1} was estimated (Coignot et al., 2020). These carbon fluxes (i.e. water-air CO_2 flux and carbon export values) highlight the importance of lateral aquatic fluxes that need to be taken into account to draw the whole regional carbon budget of the land-sea continuum formed by the Marais poitevin wetland and the Aiguillon tidal bay (Polsenaere, 2025).

Table 4. Annual net ecosystem exchange (NEE, in gC m^{-2} yr^{-1}) values found from this study and comparable literature. Site information and methodology are specified. Negative values indicate net CO_2 uptake (sink). See text for details of site conditions.

Reference (Location)	Vegetation/Crop	Method	Annual NEE (gC m^{-2} yr^{-1})	Notes
Jaksic et al. 2006 (Ireland)	Managed humid grassland	EC (2 years)	-193 (wet year); -258 (dry)	No moisture stress (soil > wilting point)
Senapati et al. 2014 (France)	Sown grassland (mowed)	EC (1 yr)	-476	"Mowing" management; net C storage ~+23 gC (after harvest)
Senapati et al. 2014 (France)	Sown grassland (grazed)	EC (1 yr)	-231	"Grazing" management; net C storage ~+141 gC
Wang et al., 2013 (China)	Winter wheat (crop season)	EC & auto-chamber	-251 (EC, seasonal)	Chamber EC comparison; EC flux matched AC flux within 7%
Schmidt et al. 2012 (Germany)	Winter wheat (annual)	EC (2 yrs)	-270 (Yr1); -270 (Yr2)	High uptake: field was a net source if harvest C removed
This study (France)	Permanent grassland	EC + RF model	-145	Sector 8; permanent grazed grassland; measured 2023-2024
This study (France)	Winter wheat	EC + RF model	-308	Sectors 4+5+6 combined; harvested 7 July 2023
This study (France)	Mixed grain	EC + RF model	+151	Sector 2; one-off rotation, ex-grassland; harvested 24 May 2023; footprint partly exposed to ditch
This study (France)	Hybrid grassland/mixed-grain transition	EC + RF model	+195	Sector 1; composite signal incl. riparian vegetation; highest Reco of all sectors

5 Conclusions

This study shows that a single, cost-effective Eddy Covariance tower, combined with wind-sector partitioning and sector-specific Random Forest modelling, can resolve contrasting annual NEE dynamics across a heterogeneous agricultural mosaic without multiple flux towers. While a pooled, sector-agnostic model matched the sector-specific models in point-prediction performance, it strongly compressed, and in one case inverted, the inter-sector contrasts once fluxes were reconstructed over the full year — a shrinkage effect toward the dominant sector's response that makes the sector-specific approach a necessary, not merely preferable, choice for such heterogeneous sites. Partitioning the fluxes into GPP and Reco confirmed these results and showed that sectors with similar land-cover labels can still have markedly different carbon dynamics.

Over the study year, the permanent grassland and the winter wheat field acted as net carbon sinks (−145 and −308 gC m^{-2} yr^{-1}), while the mixed-grain field acted as net source (+151 gC m^{-2} yr^{-1}). These estimates are not equally reliable: the wheat and grassland budgets are well constrained, whereas the mixed-grain sectors, adjacent to drainage ditches and riparian vegetation, yield more composite, cautiously-interpreted estimate. Key limitations include the one-year observation period and the exclusion of soil carbon and lateral aquatic fluxes from the budget. Beyond these limitations, this single-tower, wind-sector framework offers a transferable, low-cost alternative to multi-tower setups for resolving field-scale carbon dynamics in other heterogeneous agricultural landscapes.

Appendix A

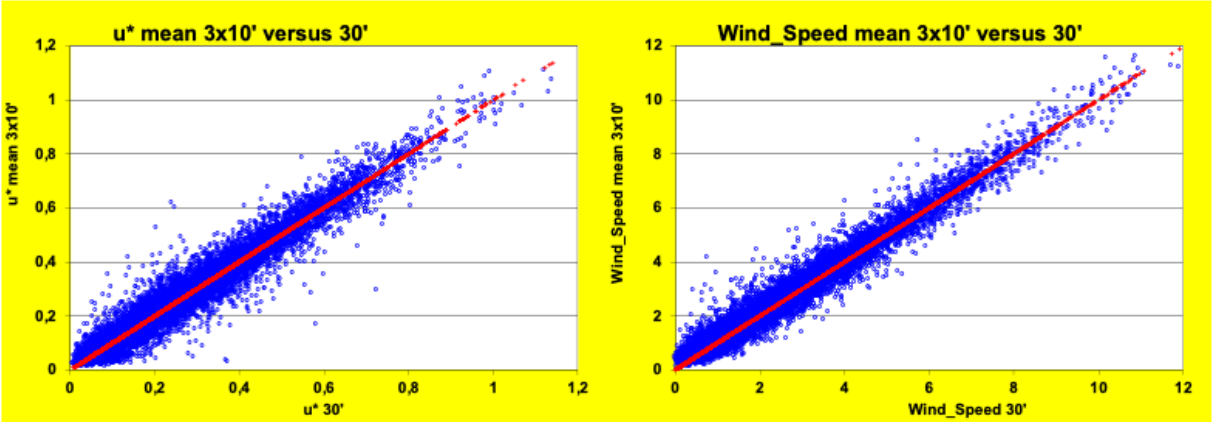


Figure A1. Friction velocity (u^*) and wind speed data as the mean of the three 10 min. values versus corresponding mean values obtained with an averaging time period of 30 min. (Comparisons done between 10 min. and 30 min. data showed that, according to the land parcel, a loose between 21% (wheat, sector 213°-283° with dominant wind directions) and 49% (wheat, sector 133°-173°) of data would be obtained with a choice of 30 min. mean flux calculation instead of the 10 min. mean flux procedure used here)

Appendix B

Table B1. Random Forest model parametrization.

Parameter	Value description
Algorithm	Random Forest regression (RandomForest package, R)
Number of trees (ntree)	300
Number of variables tried at each split (mtry)	$\lfloor\sqrt{(\text{number of predictors})}\rfloor = \lfloor\sqrt{8}\rfloor = 2$
Node/leaf size	Default RF settings (not modified)
Predictor variables (8)	PAR, relative humidity, air temperature, wind direction, day of year (DOY), soil moisture, soil temperature, day/night category (PAR > 10)
Target variable	Filtered NEE (CO_2 flux after hard-limit/QA filtering)
Uncertainty estimation	Bootstrap aggregation, B = 100 resamples (sampling with replacement) per sector/period model
Ensemble prediction	Median of the 100 bootstrap RF predictions
95% prediction interval	2.5^{th}–97.5^{th} percentile of the 100 bootstrap predictions
Performance metrics	RMSE, MAE, R^2 (computed between observed flux and bootstrap-median prediction)

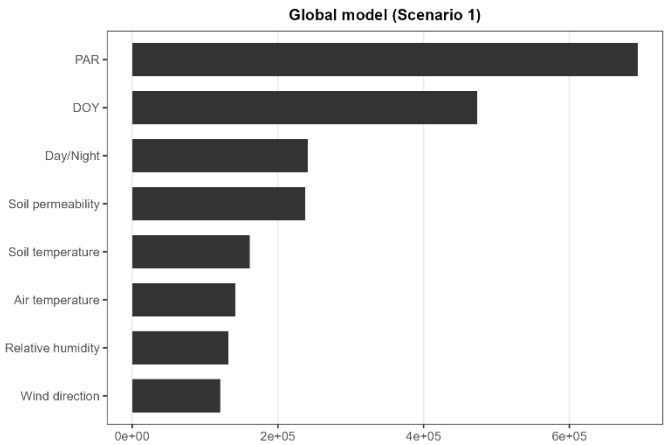


Figure B2. Variable importance of the global model (Scenario 1), expressed as percentage increase in mean squared error (%IncMSE)

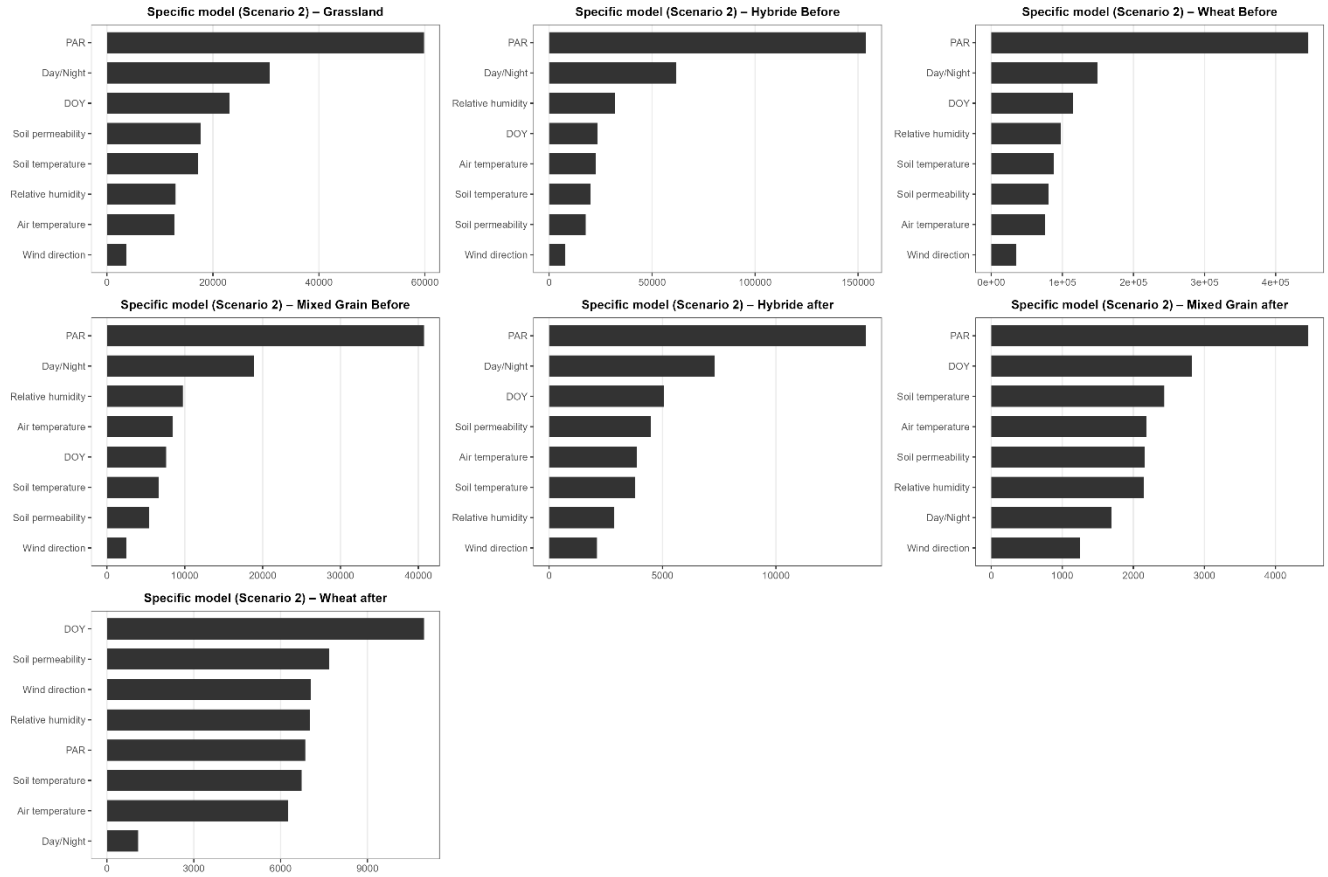


Figure B3. Variable importance across individual models in Scenario 2, expressed as percentage increase in mean squared error (%IncMSE)

Code/Data availability

All raw data can be provided by the corresponding authors upon request.

Author contribution

PP, EL, CP designed the study. PP, CD obtained the funding. PP, CP, JD performed the field work (EC tower deployment, maintenance, data retrieval). CP, PP, EL investigated and processed the data. PP, EL, NV, SB provided resources for data analysis. CP, PP, EL, NV, VM validated the data. CP made the graphics, wrote the draft. CP, PP led the review process assisted by CD, NV, OP.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work is a contribution of Camille Pery's two years work-study contract at Ifremer (Sept. 2022-Sept. 2024) in relationships with the LRTZC project (2019-2027: *La Rochelle Territoire Zero Carbone,* PIA TIGA #25875420; Caisse des Dépôts, Région Nouvelle-Aquitaine) and related to the ANR-22-PEXF-0006 - CarboNium project. We are grateful to our Ifremer colleague Louis Costes for his help in setting up the EC station, and to Wilfried Guignard and Rodolphe Pouponnot for the culture (management) information they provided us. We thank Maxime Paschal for his work on NDVI data during his internship and Nicolas Lachaussée for organizing drone flights.

References

Agreste: Mémento de la statistique agricole – Nouvelle-Aquitaine, Ministère de l'Agriculture et de la Souveraineté alimentaire, DRAAF Nouvelle-Aquitaine, 2020.

Bajgain, R., Xiao, X., Basara, J., Wagle, P., Zhou, Y., Mahan, H., Gowda, P., McCarthy, H. R., Northup, B., Neel, J., and Steiner, J.: Carbon dioxide and water vapor fluxes in winter wheat and tallgrass prairie in central Oklahoma, Sci. Total Environ., 644, 1511–1524, https://doi.org/10.1016/j.scitotenv.2018.07.010, 2018.

Baldocchi, D. D.: How eddy covariance flux measurements have contributed to our understanding of Global Change Biology, Glob. Change Biol., 26, 242–260, https://doi.org/10.1111/gcb.14807, 2020.

Barr, A. G., Richardson, A. D., Hollinger, D. Y., Papale, D., Arain, M. A., Black, T. A., Bohrer, G., Dragoni, D., Fischer, M. L., Gu, L., Law, B. E., Margolis, H. A., McCaughey, J. H., Munger, J. W., Oechel, W., and Schaeffer, K.: Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies, Agric. For. Meteorol., 171–172, 31–45, https://doi.org/10.1016/j.agrformet.2012.11.023, 2013.

Bolsée, S., Dupont, S., Lamaud, E., Irvine, M., Lafont, S., Chipeaux, C., Garrigou, C., Bonnefond, J.M., Domec, J.C., Loustau, D. : Wind flow dynamics over a photovoltaic power plant: comparison with a vegetated canopy, in preparation, 2026

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

Breiman, L.: Random Forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Burba, G.: Eddy Covariance Method for Scientific, Industrial, Agricultural and Regulatory Applications: A Field Book on Measuring Ecosystem Gas Exchange and Areal Emission Rates, https://doi.org/10.13140/RG.2.1.4247.8561, 2013.

Butzer, K. W.: French Wetland Agriculture in Atlantic Canada and Its European Roots: Different Avenues to Historical Diffusion, Ann. Assoc. Am. Geogr., 92, 451–470, https://doi.org/10.1111/1467-8306.00299, 2002.

Callejas-Rodelas, J. Á., Knohl, A., Mammarella, I., Vesala, T., Peltola, O., and Markwitz, C.: Does increased spatial replication above heterogeneous agroforestry improve the representativeness of eddy covariance measurements?, Biogeosciences, 22, 4507–4529, https://doi.org/10.5194/bg-22-4507-2025, 2025.

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, Agric. For. Meteorol., 151, 87–100, https://doi.org/10.1016/j.agrformet.2010.09.005, 2011.

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, Geophys. Res. Lett., 45, 9275–9287, https://doi.org/10.1029/2018GL079306, 2018.

Coignot, E., Polsenaere, P., Soletchnik, P., Le Moine, O., Souchu, P., Joyeux, E., Le Roy, Y., Guéret, J.-P., Froud, L., Gallais, R., Chourré, E., Chaigneau, L., Soletchnik, P., Le Moine, O., Souchu, P., Joyeux, E., Le Roy, Y., Guéret, J.-P., Froud, L., Gallais, R., Chourré, E., and Chaigneau, L.: Variabilité spatio-temporelle des nutriments et du carbone et flux associés le long d'un continuum terrestre-aquatique tempéré (Marais poitevin – Baie de l'Aiguillon – Pertuis Breton), 2020.

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, J. Geophys. Res.-Biogeosciences, 127, e2021JG006240, https://doi.org/10.1029/2021JG006240, 2022

Flanagan, L. B. and Adkinson, A. C.: Interacting controls on productivity in a northern Great Plains grassland and implications for response to ENSO events, Glob. Change Biol., 17, 3293–3311, https://doi.org/10.1111/j.1365-2486.2011.02461.x, 2011.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Arneth, A., Arora, V., Bates, N. R., Becker, M., Bellouin, N., Berghoff, C. F., Bittig, H. C., Bopp, L., Cadule, P., Campbell, K., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Colligan, T., Decayeux, J., Djeutchouang, L. M., Dou, X., Duran Rojas, C., Enyo, K., Evans, W., Fay, A. R., Feely, R. A., Ford, D. J., Foster, A., Gasser, T., Gehlen, M., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Lan, X., Lauvset, S. K., Lefèvre, N., Liu, Z., Liu, J., Ma, L., Maksyutov, S., Marland, G., Mayot, N., McGuire, P. C., Metzl, N., Monacci, N. M., Morgan, E. J., Nakaoka, S.-I., Neill, C., Niwa, Y., Nützel, T., Olivier, L., Ono, T., Palmer, P. I., Pierrot, D., Qin, Z., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Steinhoff, T., et al.: Global Carbon Budget 2024, Earth Syst. Sci. Data, 17, 965–1039, https://doi.org/10.5194/essd-17-965-2025, 2025.

Gao, D., Yao, J., Yu, S., Ma, Y., Li, L., and Gao, Z.: Eddy Covariance CO2 Flux Gap Filling for Long Data Gaps: A Novel Framework Based on Machine Learning and Time Series Decomposition, Remote Sens., 15, 2695, https://doi.org/10.3390/rs15102695, 2023.

Gash, J. H. C. and Culf, A. D.: Applying a linear detrend to eddy correlation data in realtime, Bound.-Layer Meteorol., 79, 301–306, https://doi.org/10.1007/BF00119443, 1996.

Genuer, R., Poggi, J.-M., and Tuleau-Malot, C.: Variable selection using random forests, Pattern Recognit. Lett., 31, 2225–2236, https://doi.org/10.1016/j.patrec.2010.03.014, 2010.

Gilmanov, T. G., Verma, S. B., Sims, P. L., Meyers, T. P., Bradford, J. A., Burba, G. G., and Suyker, A. E.: Gross primary production and light response parameters of four Southern Plains ecosystems estimated using long-term CO2-flux tower measurements, Glob. Biogeochem. Cycles, 17, undefined-undefined, https://doi.org/10.1029/2002gb002023, 2003.

Gu, L., Falge, E. M., Boden, T., Baldocchi, D. D., Black, T. A., Saleska, S. R., Suni, T., Verma, S. B., Vesala, T., Wofsy, S. C., and Xu, L.: Objective threshold determination for nighttime eddy flux filtering, Agric. For. Meteorol., 128, 179–197, https://doi.org/10.1016/j.agrformet.2004.11.006, 2005.

Heimsch, L., Lohila, A., Tuovinen, J.-P., Vekuri, H., Heinonsalo, J., Nevalainen, O., Korkiakoski, M., Liski, J., Laurila, T., and Kulmala, L.: Carbon dioxide fluxes and carbon balance of an agricultural grassland in southern Finland, Biogeosciences, 18, 3467–3483, https://doi.org/10.5194/bg-18-3467-2021, 2021.

Hernandez Rodriguez, L. C., Goodwell, A. E., and Kumar, P.: Inside the flux footprint: The role of organized land cover heterogeneity on the dynamics of observed land-atmosphere exchange fluxes, Front. Water, 5, https://doi.org/10.3389/frwa.2023.1033973, 2023.

Jaksic, V., Kiely, G., Albertson, J., Oren, R., Katul, G., Leahy, P., and Byrne, K. A.: Net ecosystem exchange of grassland in contrasting wet and dry years, Agric. For. Meteorol., 139, 323–334, https://doi.org/10.1016/j.agrformet.2006.07.009, 2006.

Kirschbaum, M. U. F.: The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage, Soil Biol. Biochem., 27, 753–760, https://doi.org/10.1016/0038-0717(94)00242-S, 1995.

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

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP), Geosci. Model Dev., 8, 3695–3713, https://doi.org/10.5194/gmd-8-3695-2015, 2015.

Knauer, J., Zaehle, S., Medlyn, B. E., Reichstein, M., Williams, C. A., Migliavacca, M., De Kauwe, M. G., Werner, C., Keitel, C., Kolari, P., Limousin, J.-M., and Linderson, M.-L.: Towards physiologically meaningful water-use efficiency estimates from eddy covariance data, Glob. Change Biol., 24, 694–710, https://doi.org/10.1111/gcb.13893, 2018.

Kowalski, S., Sartore, M., Burlett, R., Berbigier, P., and Loustau, D.: The annual carbon budget of a French pine forest (Pinus pinaster) following harvest, Glob. Change Biol., 9, 1051–1065, https://doi.org/10.1046/j.1365-2486.2003.00627.x, 2003.

Kowalski, A. S., Loustau, D., Berbigier, P., Manca, G., Tedeschi, V., Borghetti, M., Valentini, R., Kolari, P., Berninger, F., Rannik, Ü., Hari, P., Rayment, M., Mencuccini, M., Moncrieff, J., and Grace, J.: Paired comparison of carbon exchange between undisturbed and regenerating stands in four managed forests in Europe, Glob. Change Biol., 10, 1707–1723, https://doi.org/10.1111/j.1365-2486.2004.00846.x, 2004.

Kumari, S., Kambhammettu, B. V. N. P., and Niyogi, D.: Sensitivity of Analytical Flux Footprint Models in Diverse Source-Receptor Configurations: A Field Experimental Study, J. Geophys. Res. Biogeosciences, 125, e2020JG005694, https://doi.org/10.1029/2020JG005694, 2020.

Liebman, M. and Dyck, E.: Crop Rotation and Intercropping Strategies for Weed Management, Ecol. Appl., 3, 92–122, https://doi.org/10.2307/1941795, 1993.

Liu, L., Zhou, W., Guan, K., Peng, B., Xu, S., Tang, J., Zhu, Q., Till, J., Jia, X., Jiang, C., Wang, S., Qin, Z., Kong, H., Grant, R., Mezbahuddin, S., Kumar, V., and Jin, Z.: Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems, Nat. Commun., 15, 357, https://doi.org/10.1038/s41467-023-43860-5, 2024.

Lloyd, J. and Taylor, J. A.: On the Temperature Dependence of Soil Respiration, Funct. Ecol., 8, 315–323, https://doi.org/10.2307/2389824, 1994.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L., and Richardson, A. D.: Global Convergence in the Temperature Sensitivity of Respiration at Ecosystem Level, Science, 329, 838–840, https://doi.org/10.1126/science.1189587, 2010.

Mayen, J., Polsenaere, P., Lamaud, É., Arnaud, M., Kostyrka, P., Bonnefond, J.-M., Geairon, P., Gernigon, J., Chassagne, R., Lacoue-Labarthe, T., Regaudie de Gioux, A., and Souchu, P.: Atmospheric CO_2 exchanges measured by eddy covariance over a temperate salt marsh and influence of environmental controlling factors, Biogeosciences, 21, 993–1016, https://doi.org/10.5194/bg-21-993-2024, 2024.

Mayen, J., Polsenaere, P., Regaudie de Gioux, A., Deborde, J., Collin, K., Le Merrer, Y., Foucault, É., Ouisse, V., André, L., Arnaud, M., Kostyrka, P., Lamaud, É., Abril, G., and Souchu, P.: Aquatic metabolism influences temporal variations of water carbon and atmospheric carbon dioxide fluxes in a temperate salt marsh, Biogeosciences, 22, 5387–5411, https://doi.org/10.5194/bg-22-5387-2025, 2025.

Moreaux, V., Longdoz, B., Berveiller, D., Delpierre, N., Dufrêne, E., Bonnefond, J.-M., Chipeaux, C., Joffre, R., Limousin, J.-M., Ourcival, J.-M., Klumpp, K., Darsonville, O., Brut, A., Tallec, T., Ceschia, E., Panthou, G., and Loustau, D.: Environmental control of land-atmosphere CO2 fluxes from temperate ecosystems: a statistical approach based on homogenized time series from five land-use types, Tellus B Chem. Phys. Meteorol., 72, 2020.

Mukumbuta, I., Shimizu, M., and Hatano, R.: Short-term land-use change from grassland to cornfield increases soil organic carbon and reduces total soil respiration, Soil Tillage Res., 186, 1–10, 2019.

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., Papuga, S. A., Blanken, P. D., Noormets, A., Sulman, B. N., Scott, R. L., Wang, L., and Phillips, R. P.: The increasing importance of atmospheric demand for ecosystem water and carbon fluxes, Nat. Clim. Change, 6, 1023–1027, https://doi.org/10.1038/nclimate3114, 2016.

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, Glob. Change Biol., 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, 3, 571–583, https://doi.org/10.5194/bg-3-571-2006, 2006.

Polsenaere, P., Lamaud, E., Lafon, V., Bonnefond, J.-M., Bretel, P., Delille, B., Deborde, J., Loustau, D. and Abril, G.: Spatial and temporal CO2 exchanges measured by Eddy Covariance over a temperate intertidal flat and their relationships to net ecosystem production, Biogeosciences, 9, 249–268, https://doi.org/10.5194/bg-9-249-2012, 2012.

Polsenaere, P.: Dynamique du carbone en zone côtière tempérée: processus biogéochimiques et échanges aux interfaces terre – zone côtière – atmosphère. HDR. https://archimer.ifremer.fr/doc/00993/110419/, 2025.

Rannik, Ü., Mammarella, I., Vesala, T., Väkimies, P., Heiskari-Tuohiniemi, H., and Korkiakoski, M.: Good performance of low-cost carbon dioxide sensor based on intercomparisons with the standard eddy-covariance system, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2026-144, 2026.

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 CO2 and H2O fluxes over grassland and agroforestry, Atmos. Meas. Tech., 17, 6047–6071, https://doi.org/10.5194/amt-17-6047-2024, 2024.

Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogée, J., Allard, V., Aubinet, M., Bernhofer, Chr., Buchmann, N., Carrara, A., Grünwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. m., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.-F., Sanz, M.-J., Vesala, T., and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, Glob. Change Biol., 13, 634–651, https://doi.org/10.1111/j.1365-2486.2006.01224.x, 2007.

Reynolds: I. On the theory of lubrication and its application to Mr. Beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil, Proc. R. Soc. Lond., 40, 191–203, https://doi.org/10.1098/rspl.1886.0021, 1886.

Schmidt, M., Reichenau, T. G., Fiener, P., and Schneider, K.: The carbon budget of a winter wheat field: An eddy covariance analysis of seasonal and inter-annual variability, Agric. For. Meteorol., 165, 114–126, https://doi.org/10.1016/j.agrformet.2012.05.012, 2012.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S., and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, https://doi.org/10.1038/nature10386, 2011.

Senapati, N., Chabbi, A., Gastal, F., Smith, P., Mascher, N., Loubet, B., Cellier, P., and Naisse, C.: Net carbon storage measured in a mowed and grazed temperate sown grassland shows potential for carbon sequestration under grazed system, Carbon Manag., 5, 131, https://doi.org/10.1080/17583004.2014.912863, 2014.

Tilman, D., Reich, P. B., and Knops, J. M. H.: Biodiversity and ecosystem stability in a decade-long grassland experiment, Nature, 441, 629–632, https://doi.org/10.1038/nature04742, 2006.

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.

Vickers, D. and Mahrt, L.: Quality Control and Flux Sampling Problems for Tower and Aircraft Data, 1997.

Wagle, P., Gowda, P. H., Northup, B. K., Turner, K. E., Neel, J. P. S., Manjunatha, P., and Zhou, Y.: Variability in carbon dioxide fluxes among six winter wheat paddocks managed under different tillage and grazing practices, Atmos. Environ., 185, 100–108, https://doi.org/10.1016/j.atmosenv.2018.05.003, 2018.

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 Sens., 8, 742, https://doi.org/10.3390/rs8090742, 2016.

Wang, K., Liu, C., Zheng, X., Pihlatie, M., Li, B., Haapanala, S., Vesala, T., Liu, H., Wang, Y., Liu, G., and Hu, F.: Comparison between eddy covariance and automatic chamber techniques for measuring net ecosystem exchange of carbon dioxide in cotton and wheat fields, Biogeosciences, 10, 6865–6877, https://doi.org/10.5194/bg-10-6865-2013, 2013.

Wohlfahrt, G. and Gu, L.: The many meanings of gross photosynthesis and their implication for photosynthesis research from leaf to globe, Plant Cell Environ., 38, 2500–2507, https://doi.org/10.1111/pce.12569, 2015.

Wu, L., Miao, H., and Liu, T.: Development in Agricultural Ecosystems' Carbon Emissions Research: A Visual Analysis Using CiteSpace, Agronomy, 14, 1288, https://doi.org/10.3390/agronomy14061288, 2024.

Zeng, J., Matsunaga, T., Tan, Z.-H., Saigusa, N., Shirai, T., Tang, Y., Peng, S., and Fukuda, Y.: Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest, Sci. Data, 7, 313, https://doi.org/10.1038/s41597-020-00653-5, 2020.

Zomer, R. J., Bossio, D. A., Sommer, R., and Verchot, L. V.: Global Sequestration Potential of Increased Organic Carbon in Cropland Soils, Sci. Rep., 7, 15554, https://doi.org/10.1038/s41598-017-15794-8, 2017.