CC1: 'Comment on egusphere-2026-524, xiangmin sun, 24 Apr 20261

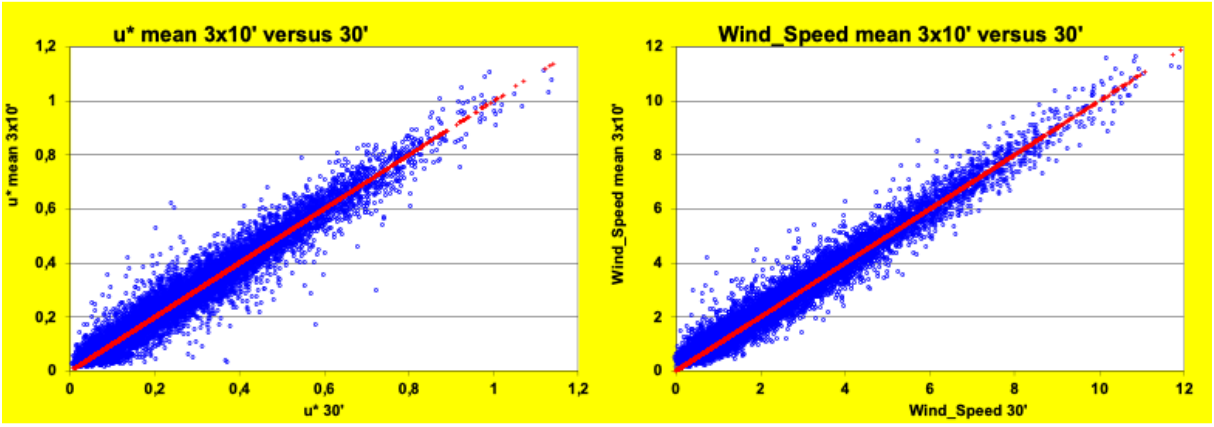
This study measured CO2 fluxes from one eddy covariance tower over a heterogeneous agricultural landscape, and separated fluxes from each sector based on wind direction and the associated radians. These CO2 fluxes, labeled by each land use, contained many gaps and were filled using a random forest algorithm for annual cumulative evaluation. A few of my questions related to eddy covariance data are:

Using a 10-minute averaging period for eddy covariance introduces random error, as it may fail to capture low-frequency turbulence. You may need to run an ogive analysis and a stationarity test for verification. In line 239, you mentioned that wind direction changes quickly. Can you validate based on your wind data?

We thank Professor Sun for his first comment also mentioned by Referees 2 and 3. We did not run ogive analysis but processed data were screened using tests for steady state and turbulent conditions through EddyPro following Foken and Wichura (1996), Foken et al. (2004) and Göckede et al. (2004) (see for instance Polsenaere et al., 2012, bg-9-249-2012 and Mayen et al., 2024, bg-21-993-2024); we will add this information on the revised version of the manuscript. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.2.1, lines 179-187]***

We understand this comment with regard to the 10 min. average flux window we now use since several years, as adapted for more coastal and dynamic systems such as wetlands (present study), salt marshes (i.e., Mayen et al., 2024, bg-21-993-2024) and tidal flats (i.e. Polsenaere et al., 2012, bg-9-249-2012), instead of the commonly used 30 min. averaged flux window for EC data processing over terrestrial ecosystems. The comment is fair since, our study site though being a wetland system, is surrounded by different crop and grassland fields and in any cases, we need, as for all atmospheric EC studies to be sure to measure and catch all eddy that contribute to the fluxes and be careful with the corresponding EC set-up in terms of measurement height and time averaging window for the chosen specific studied site.

To address this point and more strongly justify the 10 min. averaging time period chosen in the present study as rightly noticed, we did further computations comparing the friction velocity, wind speed and flux data between the mean of the three 10 min. values and the corresponding mean values obtained by reprocessing all data (February-September 2023 period) through EddyPro with an averaging time period of 30 min. (see attached graphs below). Although we can see some dispersion between both 10 min. and 30 min. mean values (u*, wind speeds, mean fluxes), we also clearly notice that we did not have any low frequency losses on the 10 min. flux data. ***[See Pery_et_al__2026-524_MS_V2, Appendix A, Fig. A1, referenced at lines 179-187]***



Furthermore, we decided to choose a 10 min. averaging time period instead of 30 min. to lose 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. The new computations and comparisons done between 10 min. and 30 min. data showed that, according to the land parcel, we would lose between 21% (wheat, sector 213°-283° with dominant wind directions) and 49% (wheat, sector 133°-173°) of data with a choice of 30 min. mean flux calculation instead of the 10 min. mean flux procedure used here.

All these considerations and new computations will be added to the revised version of the manuscript to specify and strengthen our choice.

Could you refer any study that used the u*/u to delineate or separate the sectors? I would suggest to include a brief interpretation of "u*/u" in the introduction section.

In the atmospheric surface layer, the presence of relief elements, whether they are plant-based as in our case or man-made as in Bolsée et al. (2026), causes an increase in surface roughness z_0, which leads to an increase in the roughness coefficient u_*/u, as can be inferred from the following equation: ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.1, lines 251-257 (u*/U rationale) and reference to Bolsée et al., 2026]***

$$u(z) = (u_* / k) \, Ln((z - d)/z_0)$$

The spatial distribution of u_*/u, obtained from momentum flux measurements by Eddy Covariance, therefore allows us to differentiate the areas responsible for the CO_2 fluxes measured by the same setup.

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

The ditch is very close to the tower and could be negligible. Have you investigated the along-wind distance providing the highest (peak) contribution x_peak and x_10% relative to turbulent fluxes? Footprint analysis along wind directions at different percentages can help you estimate whether the ditch's contribution is negligible.

We apologize for not mentioning any information on our footprint calculation in the submitted manuscript. The footprint 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 (u_mean, m s^{-1}), standard deviations of the lateral velocity fluctuations after rotation (sigma_v, m s^{-1}), the Obukhov length (L), friction velocities (u_star, m s^{-1}) and wind directions (°), measured by the station from February, 1^{rst} 2023 to April, 3rd 2024. We also checked footprint model distance outputs parametrized under EddyPro using Kljun et al. (2004) and Kormann and Meixner (2001) footprint models to help estimate whether the ditch's contribution is negligible or not, as rightly suggested by Professor Sun. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.1, lines 290-296 (Kljun et al., 2015, 2004 and Kormann and Meixner, 2001 footprint models)]***

EC measurements have limitations, including energy budget closure, equipment costs, and technical challenges in data processing. But I am inclined to argue that short gaps during maintenance and rainfall are the limitations, as these gaps are trivial. The assumption of EC measurements is homogeneous field across the fetch; thus EC is not applicable to sites with high footprint heterogeneity. This is not a limitation of EC measurement itself; the challenge is separating the footprints for this complicated mosaic.

We thank Professor Sun for his feedback and agree with him about EC measurement limitations/challenges due to i) sensor maintenance/issue period without any measured fluxes that need to be gapfilled and ii) heterogenous site with different habitats seen differently in time and space according to wind directions/speeds. It is particularly true after having deployed several EC towers these last years since years of 2010's either over homogenous (less heterogenous) tidal flats, salt marsh systems (Polsenaere et al., 2012, bg-9-249-2012 and Mayen et al., 2024, bg-21-993-2024) or hereafter over clearly heterogenous sites with different habitats around the EC mast as carried out in the present study for instance. This is the reason why, owing to RF modeling tool improvements, we propose our combined measured and modelled EC flux approach that, we believe, can overcome these limitations and proved EC measurements can be done over such heterogenous crop field to have at the end consistent estimated flux (NEE) values for each parcel seen by EC sensors. In this way and as risen by all Referees, we significantly modified the manuscript in turn with further computations especially for better constrain and quantify NEE values coming from each parcel.

We also note that the interpretation of Eddy Covariance measurements in heterogeneous environments is fundamentally linked to the spatial structure of the flux footprint. In our study, this aspect is explicitly addressed through the sector-based framework introduced in Fig. 2a (submitted MS), which provides a directional decomposition of the contributing source areas. This representation is essential for our analysis, as it allows us

to directly relate the reconstructed fluxes to the surrounding land-use mosaic. This is particularly important for sector 2 (70°–103°), where multiple land-cover types are simultaneously present within the footprint, including grassland, mixed grain fields, and riparian vegetation along the Brune canal. As already outlined in the introduction section of the submitted MS, this configuration does not correspond to a single dominant source area, which complicates the attribution of the observed CO_2 signal and justifies a sector-specific treatment. By contrast, the wheat-dominated sector (133°–283°) presents a much more homogeneous footprint, with a single crop type largely controlling the measured fluxes. In this case, the sectorisation confirms that the Eddy Covariance assumptions are much more closely met, and the interpretation of the reconstructed fluxes is correspondingly more direct. Nevertheless, even in this more homogeneous configuration, additional sources of uncertainty remain, as discussed in detail in response to Reviewer 2. ***[See Sect. 2.3.1, lines 267-289]***

Since data availability is 39% (line 235) for your site. Thus, each land cover data will mostly be less than 20%. While the meteorological data is measured from the tower, not from each individual sector, the meteorological data carries the same gaps as the CO2 flux. How do you deal with gaps in meteorological variables from each sector/land use type?

We thank Professor Sun for this important question. In our modelling framework, meteorological variables are not assigned independently to each land cover sector. Instead, for sector-specific predictions, we use only the data corresponding to valid wind directions associated with that sector. For time steps where the wind direction does not fall within a given sector, these data are not directly used for that sector's modelling. ***[See Sect. 2.3.3, lines 345-352]***

To enable continuous predictions across time, we apply a simple sampling approach: for each sector, when a prediction is required outside its valid wind direction range, wind direction values are randomly sampled within the boundaries of that sector. This allows the meteorological input space to remain consistent while preserving sector-specific constraints. We will clarify this procedure in the Methods section of the coming revised MS.

Please see as well our Referee 3 response to similar comment (1-).

Some minor questions are listed below for the first half of this manuscript. A detailed proofreading is recommended before you resubmit.

We appreciate Professor Sun minor questions listed below that help, together with Referee 1, 2 and 3 comments help improving the submitted manuscript.

Line 43-44: I am not sure I understand the sentence. From my understanding, conventional agricultural practices, such as intensive tillage and heavy reliance on synthetic inputs, can lead to soil degradation and increased greenhouse gas emissions. On the other hand, sustainable farming practices, such as no-till farming and leaving crop residue (stalks, leaves) on the soil surface, can reduce soil erosion and boost organic matter/soil carbon. Residue retention can potentially reduce greenhouse gas emissions compared to traditional residue removal or burning.

We thank Professor Sun for this information we will add in our purpose rephrasing the corresponding sentence in the revised MS

Line 51-54: seasonal uptake of "carbon"? Forgot to add "carbon" before the "uptake"? What are "no-harvest practices"? Cover crop? Can you add some specifics? Maybe add a "comma ," before "and local environment"? Does "social-economic factors determining land-use intensity" refer to "farming common practices"?

The sentence will be rewritten in the revised MS for more clarity as suggested by Professor Sun.

Line 53: Maybe add "response" after "fast", or rephrase it to "high temporal frequency anemometer ". It should be "gas", not "gaz".

These modifications will be done in the revised version of the MS.

Line 63: Could you be more specific about “dozens of”? There should be more representative articles besides “Baldocchi, 2020” to refer to for the developments in flux research.

The sentence will be modified in the revised MS as: “In Europe, the Integrated Carbon Observation System (ICOS) network similarly provides standardized CO_2 flux measurements through 180 stations deployed across 16 European countries.

Line 73: Any reason for the abbreviation of Q10? Scenario when the temperature increases by 10 degrees.

The meaning of Q_{10} will be specified in the revised MS.

Line 105: not sure of the meaning of “retro-littoral”.

Retro-littoral ecosystems correspond to freshwater, brackish and salt shallow systems separated from tidal immersion by dykes and locks, that are now numerous particularly along the French Atlantic coast (Bergeon et al. 2023; Mayen et al. 2023; Xaus et al. 2025). Indeed, over the past centuries, anthropogenic activities have strongly modified coastal marsh typology and hydrological connectivity this way. Originally, the aim was to gain artificial land surfaces on the sea to develop salt- and oyster-farming activities through strict tidal control; now, many coastal marsh ecosystems are managed to promote biodiversity and patrimonial uses (Tortajada et al. 2011).

Line 114-115: If the meteorological and soil variables are measured at the tower location, you need to make sure they are representative of all three fields. Are these variables homogeneous among the three fields? Since soil cannot be measured at the tower, you may add “spot” or “location” after “the tower”.

We will specify it in the revised version of the manuscript.

Line 115: Not sure how many percent of gaps were filled for each parcel? I assume each site might have more than 50% gaps.

We will specify it in the revised version of the manuscript.

Line 128: “with 62% of its surface spent to …”, this expression seems not a common phrase.

We will rephrase this sentence as followed: “…agricultural landscape representing 62% of the total Charente-Maritime region surface, corresponding to 425,267 ha (Agreste, 2020).”

Line 145: Please check the grammar of this sentence.

The sentence will be rephrased as: “The Eddy Covariance (EC) tower was installed on the 1^{rst} of February 2023 within the 5 m wide narrow-grassed strip close to the Brune Ditch, corresponding to a safe and central location to run EC measurements coming from all multiple land covers around the mast. »

Figure 1: The left map can be smaller, showing the general location. Please mark the study site in a more visible and clear way in the left map. It might be helpful to add the wind rose map in this figure.

It will be done as suggested in the revised version of the manuscript.

Line 153: the eddies might not be “homogeneous”, with different frequencies and rotation sizes.

It is true, we will remove “homogeneous” in the sentence in the revised manuscript.

Line 154-161: the basic principle equation of eddy covariance measurement is common in textbook, and can be skipped for journal paper. In addition, the equation is not numbered and not correct technically. Air density should be dry air density. There should be separate parts, bar over rho, and bar over (w’ and s’).

It is right, as noticed by Referee 1 too. ρ given in the equation (l.155, p.6 in the submitted manuscript) was actually the dry air density air (ρ_d). We modified accordingly the equation associated text in this section as followed

"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:

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

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}. In this study, measured fluxes were averaged every 10 minutes to maximize the number of observations. Negative CO_2 fluxes indicate the ecosystem acts as a CO_2 sink, while positive fluxes indicate ecosystem source behaviour."

Line 177: I am afraid that I can't see the rainfall sensor in Figure 2.

It is true, the rainfall sensor is not present in Figure 2. For information, here is the complete photo of our EC set-up taken on February 2023, 1rst showing the sensor position:



Line 205: for more details of data processing protocol?

Indeed, please refer to the cited studies for more details on used data processing procedure, we add this precision in our revised version of the manuscript.

Figure 3: Note the meaning of the dashed red lines in the figure.

It will be noted in the revised version of the manuscript.

Line 224: usually the riparian corridor and the ditch is very close to the tower and should not be the major contributor. Did you check the flux footprint and the flux peak distance?

Please see our response above for this comment.

RC1: 'Comment on egusphere-2026-524, Anonymous Referee #1, 17 Mar 2026

This manuscript presents a novel gap-filling method for CO_2 fluxes measured by a single eddy covariance tower located in the middle of four fields with different vegetation types. The proposed approach relies on machine learning techniques, specifically a random forest algorithm, to reconstruct fluxes when wind direction does not originate from the target area. The random forest model is trained using a set of environmental variables, including temperature, radiation, soil moisture, as well as information related to crop management practices and the day of year. This approach shows strong potential for reconstructing CO_2 fluxes, as well as energy fluxes, in heterogeneous landscapes using measurements from a single flux tower, and is therefore of great interest for extending the spatial representativeness of Eddy Covariance.

We are grateful to Reviewer 1 for her/his positive and encouraging feedback on our work and for all her/his constructive following comments we addressed below.

However, although the method is promising and deserves particular attention, the manuscript would benefit from more details regarding the technicality of the method. It would in particular benefit from the computation of the uncertainties in reconstructed/gap-filled CO2 flux in each wind sector. This would allow to assess the statistical significance of the differences in CO2 fluxes observed in each wind sector (specific to a vegetation/soil type).

We thank the reviewer for this constructive suggestion. We agree that the manuscript would benefit from a clearer description of the reconstruction procedure and from an explicit quantification of uncertainty associated with the gap-filled CO_2 fluxes. In the revised version, we will add a detailed description of the sector-specific modelling framework and we will estimate uncertainty for each reconstructed wind sector using a bootstrap approach. Specifically, we will refit the random forest models on multiple bootstrap resamples within each sector, generate a distribution of reconstructed fluxes, and derive 95% prediction intervals from the empirical distribution. This will also allow us to test whether the differences in reconstructed CO_2 fluxes between wind sectors are statistically significant. We will report these uncertainties and clarify their interpretation in the Methods and Results sections of the revised MS. ***[Pery_et_al__2026-524_MS_V2, Now implemented: Methods Sect. 2.3.4, lines 366-379 (bootstrap procedure); Results Sect. 3.3 and Table 2, Fig. 8]***

The authors could also try to use spatially explicit data to evaluate if the observed CO2 flux dynamics is sound. The use of NDVI or EVI from Sentinel data may be of interest. Due to the lack in uncertainty and significance analysis, in its current form, the discussion remains somewhat speculative regarding the reconstructed CO_2 fluxes for the plots surrounding the tower.

We agree that additional independent information on vegetation dynamics would help support the interpretation of the reconstructed CO_2 fluxes. To address this point, the revised manuscript will include a new analysis based on Sentinel-2 NDVI time series for the three study plots over the entire study period as shown in Figure 1. below: ***[Pery_et_al__2026-524_MS_V2, Now implemented: Methods, NDVI paragraph, lines 298-305; Results Sect. 3.1, Fig. 6]***

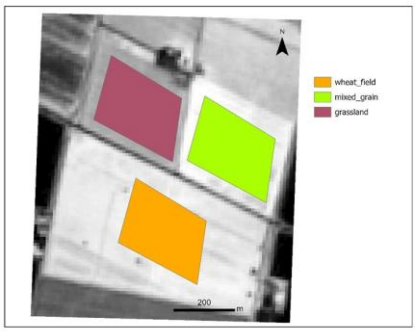


Figure 1. Polygon delineation uses for NDVI estimation

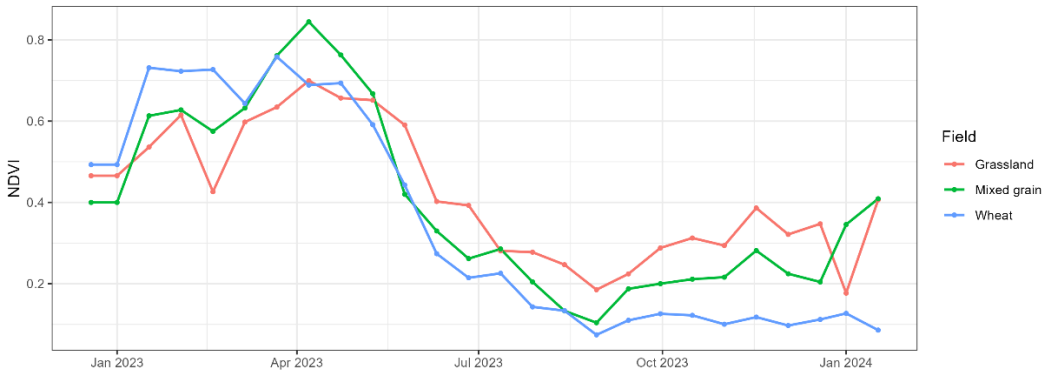


Figure 2. NDVI values of the tree different fields

Also, as shown in Figure 2, we will provide a detailed description of the phenological trajectories associated with the different land uses and management practices. In particular, it clearly highlights the contrasting dynamics of the wheat and mixed grain fields. The mixed grain field exhibits a sharp decrease in NDVI immediately after harvest on May 24th, while the vegetation is still green and physiologically active. Following the autumn sowing, NDVI increases again, reflecting the rapid re-establishment of vegetation cover. In contrast, the wheat field maintains high NDVI values until late June before gradually senescing. Harvest occurred on July 7th when the crop was already dry, resulting in a much smaller NDVI discontinuity. Furthermore, unlike the mixed grain field, no autumn sowing was performed, leaving the soil largely bare until the following spring, as reflected by persistently low NDVI values.

Importantly, these NDVI time series are computed on well-defined "pure" agricultural parcels (see Figure 1, showing the three polygon delineations), and therefore do not include the influence of the riparian vegetation located along the two canals. This distinction is essential, as this riparian vegetation is part of the EC footprint in certain sectors (as discussed in Figure 2a in the submitted MS), but is explicitly excluded from the NDVI-based analysis. For this reason, these NDVI observations cannot be directly used as predictors in the RF model, which is trained on footprint-integrated flux measurements rather than pure-field vegetation indices. ***[See revised manuscript lines 298-305]***

These contrasting vegetation trajectories provide an independent framework for interpreting the reconstructed CO_2 fluxes. For example, the abrupt reduction in vegetation cover caused by the mowing of the mixed grain field should coincide with a marked decrease in CO_2 uptake, whereas the delayed harvest of the wheat crop is expected to have a more limited impact because photosynthetic activity had already largely ceased before harvest. Similarly, the autumn regrowth observed in the mixed grain field and the prolonged bare-soil period observed in the wheat field should be reflected in the reconstructed carbon exchange dynamics.

I would therefore recommend major revision by integrating uncertainties and significance analysis, before interpreting and discussing potential differences between crop types.

Again, we thank Reviewer 1 for all her/his constructive following comments we addressed below to be able to submit after the revised manuscript according to the Co-Editor decision.

Specific comments:

L155-164: in equation ρ should be "dry" air density (ρ_d). and there is a conversion from g to mol missing. Also give units of w and s.

Referee 1 is right and ρ given in the equation (l.155, p.6 in the submitted manuscript) was actually the dry air density air (ρ_d). We modified accordingly the equation and gave w and s units as requested:

"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}. In this study, measured fluxes were averaged every 10 minutes to maximize the number of observations. Negative CO_2 fluxes indicate the ecosystem acts as a CO_2 sink, while positive fluxes indicate ecosystem source behaviour."

L190: why using a time delay with colocated CO2/H2O and sonic instruments?

Referee 1 comment is fair especially in the case of the Campbell Irgason sensor (spatially close CSAT3A and EC150 sensors) as used in the present study. Actually, in EddyPro, we used the covariance maximization option for the time lag detection method that calculates the most likely time lag based on the circular correlation procedure that is recommended in most situations. Thus, using this option instead of no time lags compensation method at all, we avoided possible flux underestimations in the case of existing potential time lags.

L200: the u* threshold is usually adapted to local conditions. Jusitify the use of a single value.

As responded to Referee 3 too, in the revised manuscript, we applied additional quality-control procedures to the CO_2 flux time series to remove physically unrealistic values that were not adequately detected by the standard filtering approach. While the conventional filters based on the median absolute deviation (MAD) and the commonly used threshold of u* < 0.1 m s^{-1} successfully removed a large fraction of low-quality measurements, several extreme flux values remained in the dataset, particularly during periods of strong meteorological disturbances.

To address this issue and as it will be presented in the revised MS, we now 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 (see Fig. 3).

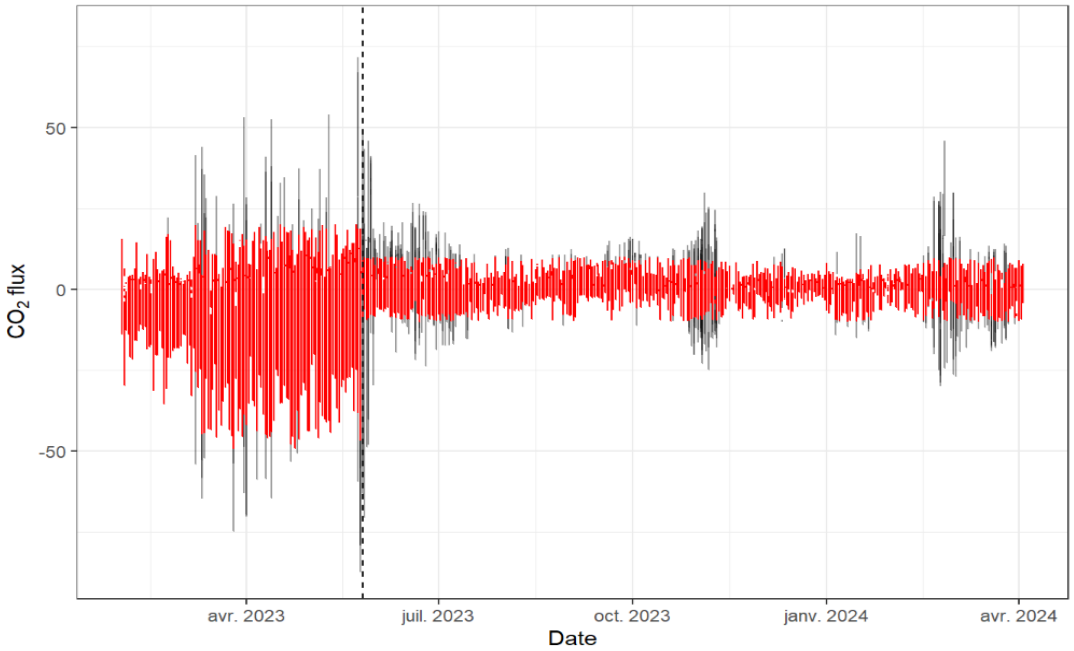


Figure 3. CO_2 fluxes presented in the submitted manuscript (black) and CO_2 with hard limits filtration (red) introduced in the revised manuscript.

The effectiveness of this additional filtering is particularly evident for the wheat sector after harvest, which was the most strongly affected by these extreme values. In the original analysis, the RF model achieved an R^2 of 0.44 for this period. After applying the hard limits, the model performance increased to an R^2 of 0.97, indicating that a limited number of unrealistic observations were disproportionately degrading the model calibration.

Finally, to better understand the RF model's performances on wheat, we tested it separately on the 3 subsectors we had defined in Fig. 2a (Sector 4: 133°-173°; sector 5: 173°-213°; sector 6: 213°-283°), including sector 5, that we previously excluded due to u* contamination from the instrumentation wake; this inclusion required recalculating the u* values, needed to filter the NEE data. At the end, the comparison of these 3 subsectors with those of the entire wheat sector (133°-283°) is now possible. All these changes will be specified in the revised MS

Figure 3 (L219): you could select only neutral conditions here to have a closer relationship with the roughness z0.

Estimating the roughness length z_0 from the roughness coefficient u_*/u would indeed require using the following equation under neutral conditions:

$u(z) = (u_* / k) \, Ln((z - d)/z_0)$

However, that's not the focus of this study, where we need to determine the spatial origin of the measured CO_2 fluxes (NEE) under all atmospheric stability conditions, with Reco being linked to stable night-time conditions and GPP being the main component of daytime NEE under unstable conditions. Using only neutral conditions to determine u*/u could therefore give a biased view of the sectors where CO_2 fluxes come from.

L244-250: this is somewhat redundant with previous paragraph.

We will modify it in the revised MS to avoid possible redundancy with regards to this information.

L265-270: the temporal subdivision, and all parameters in the RF appraoch, should be explicited in a table (in appendix eventually).

We will explicit it adding a new table in the revised MS as suggested.

L280: could you explicit what is a qualitative day-night indicator?

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}.

L282: PAR also captures seasonal changes in radiation regimes? Is this not redundant?

The choice and the importance of all these variables used in each model was checked as specified in the submitted MS and the associated Appendix A.

Section 2.3.3 : in the predictive model, a measure of the leaf area index would be very valuable and would integrate spatial variations. I would recomend trying using an NDVI or EVI high resolution satellite product (from Sentinelle data).

Please see our response and information given above according to this comment.

Secion 2.3.4: this is the critical point in this manuscript. I would suggest the authors to compute an uncertainty for the RF reconstructed CO2 fluxes. This would be necessary to interpret the results and determine if significant differences can be retreived from each wind sector. I am not an expert in RF modelling but I guess some uncertainty could be computed. At least from the comparison between model and observed data uncertainties could be deduced.

Please see our response and information given above according to this comment too.

L314-318: I would recommend avoiding any use of GPP here as NEEday is not GPP. This paragraph could be rephrased to clarify that NEE was simply split between day and night

Since we finally carried out the partitioning of simulated NEE into GPP and Reco, following Referee recommendations, this paragraph will be completely rewritten in the revised MS.

3.1 section: The Figures 7 and 8 would benefit from showing the difference between the modelled and observed CO2 flux rather than each flux on top of each other. Actually the two may be shown : the modelled fluxes and the difference with observations. This graph should include uncertainties in the modelled fluxes.

We thank Referee 1 for her/his comment and modified Figures 7 and 8 accordingly adding modelled-measured flux differences.

The rest of the results and the discussion provides valuable arguments but would really need to have uncertainties to evaluate if the differences discussed are significant or not, especially when comparing annual budgets.

We agree with Referee 1 with regards to the need to add uncertainties for annual budget comparisons; accordingly, we have done further computations in this way to address this comment and be quantitatively more accurate showing atmospheric CO_2 flux dynamics over each parcel and corresponding annual budgets (see information within our response given above).

RC2: 'Comment on egusphere-2026-524, Anonymous Referee #2, 17 Apr 2026

The manuscript “Building a random forest machine learning model for carbon budget estimation in agricultural fields using discontinuous atmospheric Eddy Covariance measurements” by Pery et al. applies Random Forest models to eddy covariance CO_2 flux data from a heterogeneous agricultural landscape over a full year, comparing a single global model with sector and crop specific models for different crop types. Both approaches use the same meteorological and soil predictors and are evaluated over one year using standard performance metrics. Results show that both strategies capture broad seasonal NEE patterns with generally similar predictive performance, while the crop specific model outperforms the global model when representing phenological stages and harvest events. A major strength of the study is the very thorough and careful analysis of wind sectors and turbulence characteristics, which provides a solid and transparent foundation for footprint attribution and sector based modelling. At the same time, a major limitation of the study is that several key conclusions are driven by differences in model structure, data availability, and temporal segmentation rather than independent evidence of process representation, while incomplete carbon accounting, sparse data coverage for many vegetation types, and the absence of uncertainty quantification limit the strength and generality of the interpretations.

We are grateful to Referee 2 for her/his positive comments on our submitted manuscript in particular on the attention we done with regard to the wind sectors and turbulence characteristics analysis. At the same time and as noticed by Referee 2, we recognize the main limitation of our study approach based in majority on modelled flux computations that definitely need further uncertainties computations/quantifications (as generally noticed by Referees) to avoid too speculative discussion and conclusion and to the contrary to give the real strength of our data interpretations deserved by the approach we propose. Owing to the significant work done by Referee 2 and the others, we deeply address all Referee comments through further computations and deep modifications in the associated coming revised manuscript according to the Co-Editor decision.

The document has (very) few very small grammatical or spelling errors (i.e., line 58 Gaz, instead of Gas, Campbellsci, instead of Campbell Scientific, Logan Utah, etc.). Furthermore, several sections reintroduce acronyms such as NEE or GPP, which only need to be defined at their first mention.

We took into account Referee comment above and modified in the submitted manuscript in this way.

Major concerns

The comparison between the global and sector specific Random Forest models is difficult to interpret. The sector specific models are trained on more homogeneous data and are independently optimized, which inherently increases their flexibility. Their improved performance is therefore largely expected and does not provide evidence for distinct mechanistic controls on NEE. Rather, the comparison demonstrates spatial and management driven heterogeneity within the flux footprint and should be interpreted as descriptive rather than explanatory. This issue is compounded by the strong imbalance in data coverage across vegetation types. Vegetation classes with sparse observations, representing only 2 to 4 percent of the dataset and showing limited variability, exhibit the highest R^2 values, whereas wheat, which represents the largest and most dynamic subset, shows substantially lower performance, particularly post harvest. High R^2 values in sparsely sampled systems are therefore more likely to reflect limited signal complexity than superior predictive skill, and cross vegetation comparisons should be interpreted with caution.

The comparison between the global and sector-specific RF models will be substantially revised in the revised version of the manuscript to avoid any interpretation in terms of distinct mechanistic controls on NEE. The objective of the sectorisation is primarily to characterize the spatial heterogeneity of the flux footprint and the effects of contrasting management practices within the study area. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.2 and Sect. 4.1]***

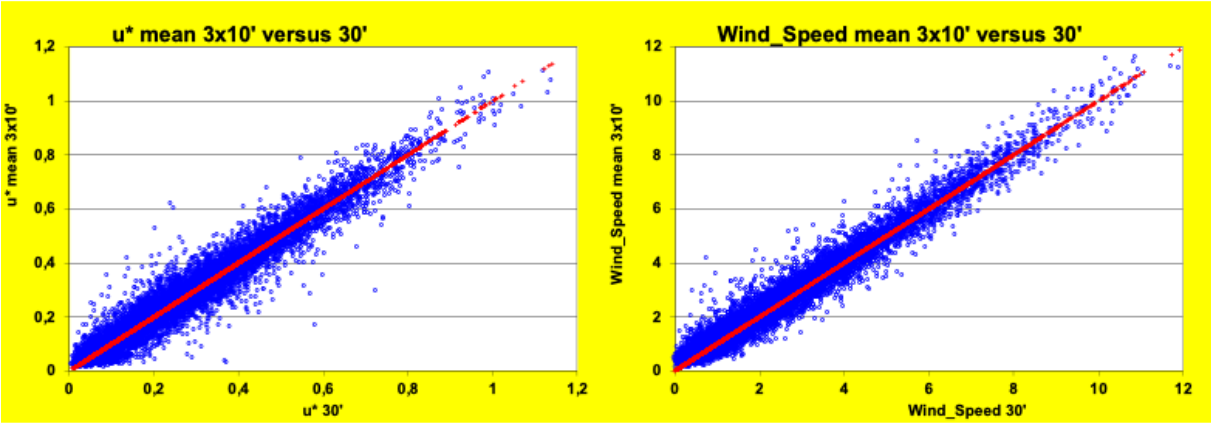
In addition, the revised analysis we carried out includes several methodological improvements as described above, namely the application of hard limits to remove spurious flux values and the inclusion of the three wheat subsectors (sectors 4, 5 and 6). These modifications significantly improved the robustness of the RF models and provide a more balanced basis for evaluating the impact of footprint heterogeneity on model performance. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.1and Sect. 2.3.2, table 1]***

These results suggest the lower performances reported in the original manuscript were largely driven by a limited number of extreme flux values and by methodological choices related to footprint partitioning rather than by intrinsic differences in ecosystem predictability. Consequently, the revised manuscript will emphasize that sector-specific models should be viewed as a descriptive framework for exploring footprint heterogeneity and management effects, while comparisons of model skill across vegetation types must be interpreted with caution because of differences in sample size, variability, and phenological dynamics. ***[See Pery_et_al__2026-524_MS_V2, Sect. 3.2]***

The use of 10-minute averaging intervals instead of the standard 30 minute eddy covariance averaging period requires stronger justification. Shorter averaging windows may exclude low frequency flux contributions and bias NEE estimates, particularly under stable or disturbed conditions. The manuscript does not provide supporting diagnostics such as ogives or spectral analyses to demonstrate that flux carrying frequencies are adequately retained within the 10-minute interval. In the absence of such evidence, the validity of the calculated fluxes and their subsequent interpretation remains uncertain.

We thank Referee 2 for her/his first comment also mentioned by Referee 3 and Professor Sun. We understand this comment with regard to the 10 min. average flux window we now use since several years, as adapted for more coastal and dynamic systems such as wetlands (present study), salt marshes (i.e. Mayen et al., 2024, bg-21-993-2024) and tidal flats (i.e. Polsenaere et al., 2012, bg-9-249-2012), instead of the commonly used 30 min. averaged flux window for EC data processing over terrestrial ecosystems. The comment is fair since, our study site though being a wetland system, is surrounded by different crop and grassland fields and in any cases, we need, as for all atmospheric EC studies to be sure to measure and catch all eddy that contribute to the fluxes and be careful with the corresponding EC set-up in terms of measurement height and time averaging window for the chosen specific studied site. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.2.1 and Appendix A, Fig. A1 (10-min vs 30-min comparison)]***

To address this point and more strongly justify the 10 min. averaging time period chosen in the present study as rightly noticed, we did further computations comparing the friction velocity, wind speed and flux data between the mean of the three 10 min. values and the corresponding mean values obtained by reprocessing all data (February-September 2023 period) through EddyPro with an averaging time period of 30 min. (see attached graphs below). Although we can see some dispersion between both 10 min. and 30 min. mean values (u*, wind speeds, mean fluxes), we also clearly notice that we did not have any low frequency losses on the 10 min. flux data. ***[See Pery_et_al__2026-524_MS_V2, Appendix A, Fig. A1, referenced at lines 179-187]***



Furthermore, we decided to choose a 10 min. averaging time period instead of 30 min. to lose 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. The new computations and comparisons done between 10 min. and 30 min. data showed that, according to the land parcel, we would lose between 21% (wheat, sector 213°-283° with dominant wind directions) and 49% (wheat, sector 133°-173°) of data with a choice of 30 min. mean flux calculation instead of the 10 min. mean flux procedure used here. ***[See Pery_et_al__2026-524_MS_V2, at lines 179-187]***

All these considerations and new computations will be added to the revised version of the manuscript to specify and strengthen our choice.

The separation of NEE into daytime and nighttime subsets does not constitute a valid flux partitioning. Daytime NEE cannot be assumed to represent GPP, and nighttime NEE cannot be equated with ecosystem respiration without an explicit partitioning approach. Respiration occurs continuously throughout the day and night, and eddy covariance does not directly measure GPP. Consequently, cumulative NEEday and NEEnight sums cannot be interpreted as physiological components of the carbon balance. The justification based on changing wind direction does not resolve this conceptual limitation, and the resulting interpretation of metabolic fluxes and annual component budgets therefore requires substantial revision.

We totally agree with Referee 2, that is why we finally carried out the partitioning of simulated NEE into GPP and Reco for each of the 8 sectors in Fig. 2a as well as for the entire wheat sector. All this work will be presented in the coming revised MS. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.4 and Sect. 3.4]***

Given the uneven data coverage, the lack of flux partitioning, and the absence of uncertainty quantification, the study would benefit from a clearer discussion of limitations. Without uncertainty estimates, it is difficult to assess whether differences between crops or modelling strategies are meaningful, particularly during disturbance and post harvest periods when eddy covariance data are highly non stationary.

We thank the reviewer for this valuable comment. We agree that the limitations of the study should be more explicitly discussed. In the revised manuscript, we will expand the Discussion section to better acknowledge the uneven distribution of observations among vegetation types and the implications of using NEE without flux partitioning. ***[See Pery_et_al__2026-524_MS_V2, Discussion Sect. 4.4]***

In addition, as detailed in our response concerning uncertainty quantification, we will include a bootstrap-based uncertainty analysis and report confidence intervals for the reconstructed fluxes. This additional analysis will provide 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. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.4 and 3.3]***

Finally, some of the literature cited in support of the interpretations is more than 15 years old. While not inherently problematic, more recent studies, particularly those addressing grazing systems, may provide more relevant context given the management practices present at the site.

We thank Referee 2 for this suggestion. We agree that incorporating more recent literature would strengthen the discussion and place our findings in a more up-to-date scientific context. In the revised manuscript, we will review and update the references used to support our interpretations, with particular attention to recent studies addressing grazing systems, grasslands, and managed agricultural ecosystems. We believe that these additions will improve the discussion and better contextualize our results within the current literature.

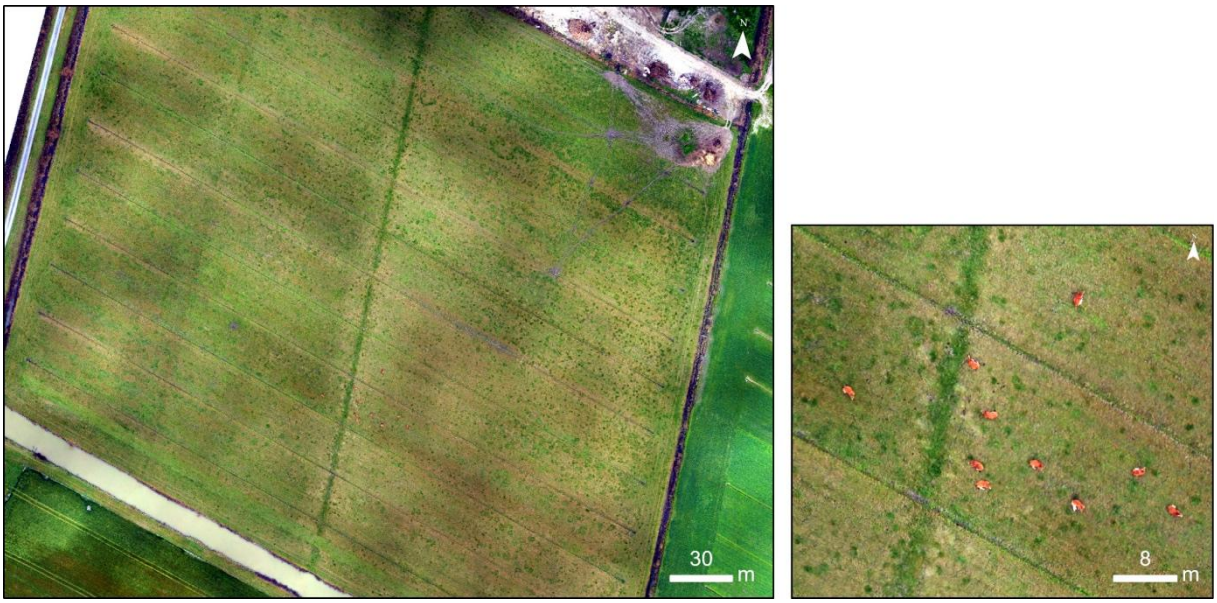
Specific notes

Line 43: Intensive agricultural is typically associated with heavy inputs, which generally don't correlate with increased carbon storages in soils. Perhaps a clarification of which systems are meant here would help

This clarification will be done in the revised version of the MS.

Line 135 please indicate how many cattle were present, as their density might influence the CO2 balance as well.

This information will be specified in the revised manuscript. 10 in February 2023, 41 in April 2023 and 43 in July 2023 cattles were present in the grassland (see the attached February 2023 file for example below).



L 161, measured fluxes were averaged every 10 minutes. This needs more expansion

Indeed, information given above will be added in the revised manuscript.

Section 2.3.2 The previous section describes in great detail how different fields were isolated based on turbulence wind sector analysis, which included some areas such as shading by the anemometer, etc, but in this section the first scenario RF is phrased as if all sectors were now included. But shouldn't at least the ditch and the tower shaded area be excluded from this model? Perhaps this is what was meant here. Further explanation is needed.

Further explanation will be given in the revised MS as rightly suggested by Referee 2 and as explained in our Referee comment responses.

Results

Section 3.1 The manuscript states that vegetation type is included in the global RF model via a qualitative variable (Scenario 1), but this variable is not clearly defined in the predictor list. Clarification is needed, which should be included in the methods section 2.3.3.

Indeed, clarification will be done in the revised manuscript.

Line 369: The term maximum measured productivity could be misleading, please use the standard NEE instead. Further, since this is mentioned for Scenario 1, it should also be mentioned for Scenario 2, as the goal is to compare their performance.

We thank Referee 2 for her/his suggestion that will be used in the revised MS

Line 373: sentence starting with "However, " do you mean here actually observed in the field, or in a dataset. If dataset, which one is referenced here?

We mean in each corresponding hybrid and mixed grained dataset. This will be more clearly specified in the revised MS.

Line 377: less productive is again a bit misleading, perhaps better "reduces photosynthetic activity"

We thank Referee 2 for her/his more appropriated suggestion here that will be used in the revised MS.

Section 3.3 Given the noise and non-stationarity of EC data, particularly during disturbance and post-harvest periods, the absence of uncertainty quantification limits interpretation and makes it difficult to assess the significance of differences between vegetation types and modelling strategies.

We agree with Referee 2 and as explained above and in our Referee's responses (Referee 1), all these aspects will be added in the revised MS.

Line 422 The description of cumulative NEEdaytime is confusing and potentially misleading. The reported "peak" value refers to a cumulative minimum rather than a flux, and the shift from -844 to -764 gC m^{-2} indicates a reversal toward daytime carbon release after harvest, not merely a slowdown in accumulation. Moreover, interpreting daytime NEE as productivity or photosynthetic uptake is conceptually incorrect without flux partitioning.

Following our partitioning of simulated NEE into GPP and Reco done for these reviews, section 3.3 will be completely rewritten in the revised MS.

Section 3.4 Again for this section including uncertainties is very important in order to distinguish between sinks and sources.

Indeed, and accordingly all these aspects done will be added in the revised MS.

Discussion

Line 509 The interpretation that reduced model performance during storm periods is "not problematic" because the RFs do not reproduce "physically unrealistic fluxes" is not justified. Failure of a Random Forest model to predict certain flux values indicates that these values are not explainable by the chosen predictors, but does not demonstrate that they are artefacts rather than genuine storm-driven ecosystem responses or unmodelled processes. Using model behaviour to infer data validity is circular and should be avoided. This passage would benefit from substantially more cautious language and explicit acknowledgement that these periods represent increased uncertainty rather than invalid fluxes.

This passage will be re-written more cautiously in the revised MS as rightly suggested by Referee 2.

Line 516 This should be noted in the methods, that this period was excluded from cum sum evaluations in section 2.4. Further description is also needed regarding how the models were run with their exclusion.

It will be noted in the M&M section of the revised MS along with the further model descriptions as suggested by Referee 2.

Line 521 The discussion of static and automated chamber measurements is confusing, as no such data are presented or used in this study. While chamber studies from the literature are cited to illustrate potential agreement with EC fluxes, this does not constitute validation of the EC measurements or RF model applied here. This section should be removed or rephrased.

Referee 2 is right and we better specified this paragraph in the revised manuscript as our intention was to say deploying simultaneously static chambers at our study site at the different seasons according to associated crop management and phenology could have been useful to validate our modelling approach and this method should be done in the future.

"In the future, to validate the EC-based flux measurements and Random Forest (RF) modelling approaches at the studied site, static and automated chamber measurements could be deployed simultaneously over each crop at different locations inside the footprint and key time (seasonal, diurnal) periods linked to phenology and management. Indeed, some studies indicate that chamber measurements can reliably reflect EC fluxes at plot scale and can also capture temporal dynamics (e.g. drought responses or post-harvest changes). For instance, Wang et al. (2013) reported near-identical NEE values using EC and automated chambers in winter wheat, with seasonal EC estimates of -251 gC m^{-2} and chamber-derived values of -205 gC m^{-2}. Similarly, Drewer et al. (2017) found consistent patterns between chamber and EC fluxes in intensively managed grasslands. Although chamber data have limited spatial coverage, along EC measurements, both techniques run simultaneously can offer crucial

plot-scale insight and could help verify RF model accuracy, especially under flux footprint conditions that are highly localized."

Line 546 I'm not sure that Reichstein et al. 2005 is an appropriate citation here, since the focus is on riparian zones and higher soil moisture contents.

Referee 2 is right; we will remove the corresponding citation here in the revised MS.

Line 596 The term "behaved" is imprecise and anthropomorphic when referring to a winter wheat carbon balance.

Indeed, it will be modified in the revised manuscript as followed: "By contrast, the established winter wheat field was almost neutral in full carbon accounting."

Line 596 the harvest data should be introduced in the results section.

It will be introduced in the result section of the revised MS as proposed.

Line 604 this sentence is very hard to understand and should be broken up into different sections. Further, it is hard to see the connection to the present study, when the measurement timeframes differ by at least 5 years. Further, different flux pathways, spatial domains, and time scales are mixed without clear signposting, giving the impression of a quantitative comparison with cumulative NEE that is not actually performed. Clearer framing is needed to distinguish contextual background from results included in the present carbon balance.

All this section will be modified accordingly in the revised MS.

Conclusion

The concluding claim that the approach provides a "scalable template for integration into larger-scale carbon inventories" appears overstated, as no spatial or temporal scaling is demonstrated and key components of full carbon accounting are acknowledged to be incomplete. This statement would benefit from more cautious, prospective wording.

It will be rephrased accordingly in the revised MS.

RC3: 'Comment on egusphere-2026-524, Anonymous Referee #3, 30 Apr 2026

The manuscript by Pery et al. presents an interesting approach to gap-fill CO_2 flux densities based on wind-sector partitioning, to account for field heterogeneity at the experimental site. This method employs the same set of predictors, while separating the input data to the model in wind sectors and gap-filling them separately. While the method is sound and it has potential to improve NEE estimates over heterogeneous sites, without the costs and logistic difficulties associated to the deployment of several EC stations, there are some hindrances in this approach and how it was validated and presented. Therefore, I would recommend major revisions before this manuscript is accepted for publication.

We thank Referee 3 for her/his appreciation and particularly for all her/his comments we address below and take into account for the coming revised manuscript that greatly strengthen the validation of the proposed approach.

Major comments

1.- Were meteorological data gap-filled? How did you deal with these data for the specific wind sectors and periods? Please provide all these details as they are very relevant for the RF parametrization.

We thank the Referee for this important clarification request. We agree that the treatment of meteorological data is relevant for the RF parametrization and we provide additional details below. It is important to note the data filtering procedures described in the manuscript apply only to CO_2 flux measurements. As a result, the reported 32% of missing data concerns flux data exclusively and does not affect the meteorological drivers used in the random forest models.

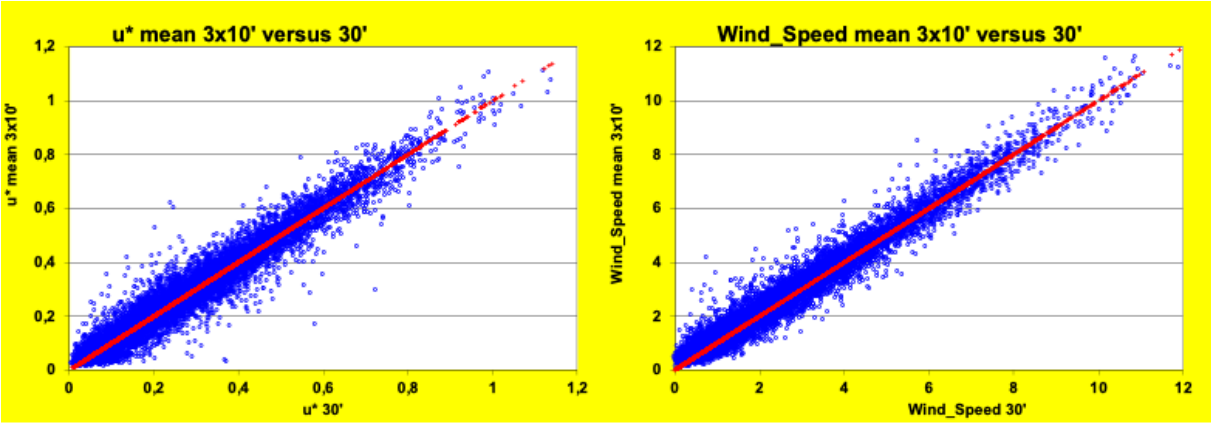
Regarding meteorological variables, a very small fraction of missing data remained after processed Eddy Pro output: approximately 0.03% for air temperature, soil temperature, and soil humidity, 0.005% for relative humidity, and 0.68% for wind direction, wind speed, and atmospheric pressure. These gaps were generally short, with a maximum duration of 7 consecutive missing values (i.e., 70 minutes).

Given the very limited extent and short duration of these gaps, we applied a simple mean-based gap-filling approach to ensure temporal continuity of the meteorological drivers used in the Random Forest models. This choice was made to avoid introducing unnecessary complexity while preserving the physical consistency of the predictors. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.3]***

2.- Using 10-min resolution for processing high-frequency data would require stronger justification; heterogeneity itself could be one, as rapidly shifting wind conditions could drive measuring over different sectors and land covers, but this can come with hindrances that should be accounted for, unless it is adequately demonstrated.

We understand the Referee 3 (and other Referees as well) comment with regard to our 10 min. average flux window we now use since several years, as adapted for more coastal and dynamic systems such as wetlands (present study), salt marshes (i.e., Mayen et al., 2024, bg-21-993-2024) and tidal flats (i.e., Polsenaere et al., 2012, bg-9-249-2012), instead of the commonly used 30 min. averaged flux window for EC data processing over terrestrial ecosystems. The comment is fair since, our study site though being a wetland system, is surrounded by different crop and grassland fields and in any cases, we need, as for all atmospheric EC studies to be sure to measure and catch all eddy that contribute to the fluxes and be careful with the corresponding EC set-up in terms of measurement height and time averaging window for the chosen specific studied site.

To address this point and more strongly justify the 10 min. averaging time period chosen in the present study as rightly noticed, we did further computations comparing the friction velocity, wind speed and flux data between the mean of the three 10 min. values and the corresponding mean values obtained by reprocessing all data (February-September 2023 period) through EddyPro with an averaging time period of 30 min. (see attached graphs below). Although we can see some dispersion between both 10 min. and 30 min. mean values (u*, wind speeds, mean fluxes), we also clearly notice that we did not have any low frequency losses on the 10 min. flux data.



Furthermore, we decided to choose a 10 min. averaging time period instead of 30 min. to lose 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. The new computations and comparisons done between 10 min. and 30 min. data showed that, according to the land parcel, we would lose between 21% (wheat, sector 213°-283° with dominant wind directions) and 49% (wheat, sector 133°-173°) of data with a choice of 30 min. mean flux calculation instead of the 10 min. mean flux procedure used here.

All these considerations and new computations will be added to the revised version of the manuscript to specify and strengthen our choice. ***[See Pery_et_al__2026-524_MS_V2, Appendix A, Fig. A1 (10-min vs 30-min comparison), referenced at lines 179-187]***

3.- In the results section, the figures should be presented following a logical order and in a more complete but concise way. Currently, this section is a bit unorganized and mentions to Figures 5 or 6 are included before any explanation of Figure 4. Please try to re-write this to help readers to follow the story.

Thank you for this helpful comment. We agree that the current structure of the Results section may not be optimal. In the revised manuscript, we will reorganize the section to ensure a more logical flow of figures and their corresponding descriptions. This restructuring will be implemented in line with the other modifications made in response to the reviewers' comments, resulting in a clearer and more coherent presentation of the results. ***[Results section reorganized]***

4.- Perhaps the RF could benefit from adding time cycle variables, for day/night and for seasons. This was implemented e.g. in Vekuri et al. (2023), using cyclical indicators day-night and for months of the year, as well as a linear indicator of the day of year. Maybe this removes in an easier way the uncertainty in gap filling related to changes in the ecosystem properties due to management events such as harvest.

We thank Referee 3 for this suggestion. We agree that cyclical time variables can be useful in some modelling contexts. However, in our case we chose not to include explicit seasonal or monthly cyclical encodings because the main temporal dynamics are already well captured by PAR, day-of-year (DOY), and the day/night indicator. Adding additional cyclic variables would therefore provide largely redundant information. Moreover, we aimed to keep the model as simple and interpretable as possible, especially since Random Forests are already able to handle non-linear temporal patterns without requiring explicit seasonal transformations. We will clarify this point in the revised manuscript. ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.3.3]***

5.- The inclusion of some events, like the November storm, could be interesting if these specific phenomena would be studied. However, as this is not the scope of the text and this can bias the results of the modeling, I would suggest to either remove it, or to address the specific effect it may have on the gap-filled data. Furthermore, it the storm was relevant for CO2, but no physiological activity was taking place at the moment, could it be an effect of a bad performance of the analyzer which then justifies to remove it from the time series? Or which explanation did you find for it?

In the revised manuscript, we applied additional quality-control procedures to the CO_2 flux time series to remove physically unrealistic values that were not adequately detected by the standard filtering approach. While the

conventional filters based on the median absolute deviation (MAD) and the commonly used threshold of u* < 0.1 m s^{-1} successfully removed a large fraction of low-quality measurements, several extreme flux values remained in the dataset, particularly during periods of strong meteorological disturbances.

To address this issue, 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 (see Fig. 3). ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.2.2, lines 225-233 (Fig. 3)]***

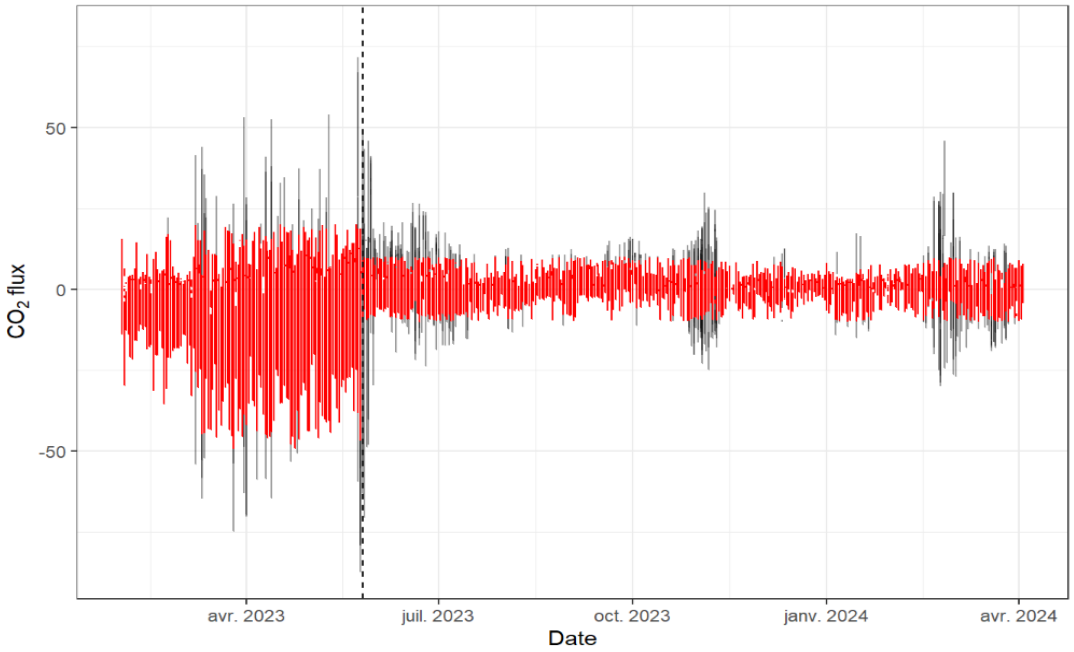


Figure 3. CO_2 fluxes presented in the submitted manuscript (black) and CO_2 with hard limits filtration (red) introduced in the revised manuscript.

The effectiveness of this additional filtering is particularly evident for the wheat sector after harvest, which was the most strongly affected by these extreme values. In the original analysis, the RF model achieved an R^2 of 0.44 for this period. After applying the hard limits, the model performance increased to an R^2 of 0.97, indicating that a limited number of unrealistic observations were disproportionately degrading the model calibration.

6.- The manuscript could benefit from adding the partitioning of NEE into GPP and RECO using either Reichstein et al. (2005) or Lasslop et al. (2010), or both. This would allow to skip the separation into nighttime and daytime NEE. These models could be applied to the filled data for the whole dataset and for the different wind sectors and periods. Please also be careful with naming daytime NEE as productivity or production, because this is the net CO_2 flux density and not photosynthesis.

We thank Referee 3 for encouraging us to carry out the partitioning of simulated NEE into GPP and Reco, which we regret not having done in the submitted MS. We've now done it for each of the 8 sectors shown in Fig. 2a as well as for the entire wheat sector, using the method from Kowalski et al. (2003, 2004). ***[See Pery_et_al__2026-524_MS_V2, Sect. 2.4]***

The current Figure 9 will therefore be replaced in the revised MS by a new figure showing the cumulative GPP, Reco, and NEE for all the sectors considered. As a result, sections 3.3, 3.4, and 4 will be largely rewritten accordingly.

Kowalski, A. S., Loustau, D., Berbigier, P., Manca, G., Tedeschi, V., Borghetti, M., Valentini, R., Kolari, P., Berniger, F., Rannik, U., 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, Global. Change. Biol., 10, 1707–1723, 2004.

Kowalski, S., Sartore, M., Burlett, R., Berbigier, P., and Loustau, D.: The annual carbon budget of a French pine forest (Pinus pinaster) following harvest, Global. Change. Biol., 9, 1051–1065, 2003.

7.- Carbon budgets usually account for terms like carbon export through harvest, carbon input through organic fertilizer, etc. The most important are typically NEE and carbon export. Please avoid the term "carbon budget" throughout the text if no other lateral C transport is considered. Additionally, if there are data on C export through harvest, they should be presented in the Results section, together with annual NEE sums, and then discussed appropriately later on. If no other terms are considered, then also the title of the manuscript should change to avoid confusion.

We thank Referee 3 for this important comment. We agree that the term 'carbon budget' may be misleading in the absence of a full accounting of all carbon lateral fluxes (e.g., harvest export, organic amendments). In the revised manuscript, we will therefore avoid using this term and instead refer more specifically to NEE-based carbon dynamics. We will also clarify in the Methods and Results sections what components are included in our carbon accounting.

8.- The results and method described in the paper could be put in context with more literature describing how to address heterogeneity effects over EC sites. There are plenty of studies using either footprint modeling (e.g. Chen et al., 2011) or multiple towers (e.g. Oren et al., 2006), or a distinct approach (e.g. Wang et al., 2016). Furthermore, related to a comment in the conclusions, (line 632), if the focus is the cost of the EC stations, there have been in recent years developments of lower-cost eddy covariance systems; some studies validated the systems (Hill et al., 2017; van Ramshorst et al., 2024; Rannik et al., 2026) and some focused on their application to improve spatial representativity of measurements (Cunliffe et al., 2022; Callejas-Rodelas et al., 2025). The text could put this also in context.

We thank Referee 3 for this helpful suggestion. We agree that the manuscript would benefit from better positioning within the existing literature on addressing spatial heterogeneity in EC measurements (e.g., footprint modelling, multi-tower approaches, and alternative methods), as well as recent developments in lower-cost EC systems. We will revise the Introduction and Discussion sections of the revised MS accordingly to better contextualize our approach within these studies. ***[New citations added: Chen et al. 2011, Oren et al. 2006, Wang et al. 2016 (footprint/multi-tower context); Cunliffe et al. 2022, Callejas-Rodelas et al. 2025, van Ramshorst et al. 2024, Rannik et al. 2026 (low-cost EC) - see Introduction and Discussion, and References list]***

8.- The manuscript would benefit from re-wording many sentences and sections for technical completeness and coherence, particularly in the Methods section. Some of these comments are included in what follows, but they do not cover the whole text.

We will take into account Referee 3 minor comments below in the revised version of the manuscript being more clear, complete and coherent in the different sections especially in the M&M as mentioned too in a previous comment (3.) above.

Minor comments

1.- Section 2.2.2 would benefit from a more clear structure and presenting more details about the EC data processing; particularly, as this is a paper presenting a novel method, all important details in data processing are necessary. For instance, which method did you apply for high- and low-frequency spectral losses?

We believe this 2.2.2 EC data processing and filtering method section is already clear and detailed in the submitted manuscript as based on previous published papers we did with an associated validated method now also used in the present study (see Polsenaere et al., 2012, bg-9-249-2012; Mayen et al., bg-21-993-2024; Mayen et al., bg-22-5387-2025). We agree that our present study shows a novel method with modelling RF approaches but our EC data processing/filtering procedure remains the same as before; in our EC studies, we always justify our procedure that remains specific to the corresponding study site. Our EC data are always carefully checked with regards to the 10 min. averaging time period choice (see comment 2 response) that can potentially lead to (low) frequency losses indeed. According to Referees comment and as requested, we will add these information

and further computations (comment 2 response) in the revised manuscript to more strongly specify and justify from this point (flux losses) our EC data processing adapted to the present studied site.

2.- Line 154: Please try to write the description of EC method in a more correct way, such as "This method measured vertical wind speed and molar density or fraction of the gas of interest at high frequency. The flux density is then derived by calculating the covariance between w and s, and multiplied by air density".

We thank Referee 3 for this precision, we will modify the sentence as proposed in the revised manuscript.

3.- Line 161: Please rewrite - Fluxes are not measured directly, but calculated from high-frequency measurements.

We will modify accordingly this sentence in the revised manuscript as: “In this study, calculated fluxes from high-frequency measurements were averaged every 10 minutes to maximize the number of observations”.

4.- Line 183: "to ensure accurate CO2 flux measurements" - this sentence is not necessary. Other software (EddiSoft, EddyUH, etc.) also produce standardized and similarly accepted calculations; furthermore, the fluxes are not measurements, but calculations from raw measurements.

We will modify accordingly this sentence suppressing this part.

5.- Line 194: "calculated using the linear detrending method" - this is not a method for calculating fluxes, but just one of the steps along the processing routine. Please rewrite this.

We will rewrite accordingly this sentence as “Linear detrending was used to calculate fluctuations as the deviations from a linear trend on a chosen 10-minute averaging time period.

6.- Line 196 to line 204: Please try to re-organize this section; it is not conceptually wrong, but maybe it would help for clarity to organize the different filters that were applied. "After flux processing, data were filtered to remove outliers using MAD (explain), to exclude periods with weak or non-existent turbulence (ustar filter, explain) and to ()".

We thank Referee 3 for her/his help clarifying this part and modify it in the revised manuscript as proposed and as also explained in your comment 5 above.

7.- Section 2.3.1: Please include information on the footprint model. How were footprints calculated? Did you parametrize something specific, like aerodynamic canopy height (as in Chu et al., 2018)?

We apologize for not mentioning any information on our footprint calculation in the submitted manuscript. The footprint 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 (u_mean, m s^{-1}), standard deviations of the lateral velocity fluctuations after rotation (sigma_v, m s^{-1}), the Obukhov length (L), friction velocities (u_star, m s^{-1}) and wind directions (°), measured by the station from February, 1st 2023 to April, 3rd 2024. We did not parametrize something specific like aerodynamic canopy height though being definitely a good approach Referee 3 mentioned and as done in Chu et al., 2018. In the present study, we only present an averaged footprint computed on the whole time period as mentioned above that was enough adapted to our modelling approach we propose in the present study. All these information according to footprint calculation will be added in the revised manuscript.

8.- Line 227: Do you have an explanation why this happens? Is this the case for all these anemometers? Please justify it.

It is not surprising as due to the sonic anemometer position, actually the whole IRGASON sensor (CSAT3A and EC150 Campbell IRGA) with regard to wind directions coming right behind the sensor in particular in the opposite direction (i.e., Southern directions) of the sensor transducer head (133°-283°) (see attached photo below). It could be specific to the IRGASON Campbell sensor (EC150 parallel arms very close to the CSAT3A transducers) though this effect could also exist for other high frequency IRGA/sonic set-up according to associated sensor geometry. Very few studies tested such potential bias through u* versus wind direction plots as we did in the

present study, so it is complicated to know if it is the case for all EC sensor set-up and/or more pronounced for some models than others. It is then something that could be checked in EC studies; again, this effect has no influence on CO_2 flux calculations directly but needs to be taken into account carefully when u* filtering procedure are done on post-processing EC data.



(April, 9th 2026)

9.- Lines 248 to 252: this should be discussed in Discussion; other methods could complement this, like footprint modeling with land use/cover attribution.

It will be discussed in the Discussion section of the revised MS along with other methods as suggested.

10.- Lines 254-257: this paragraph belongs to discussion or conclusions.

This paragraph will be displaced in the right (discussion/conclusion) section in the revised MS.

11.- Line 266: "key temporal subdivision" - do you mean the separate models were applied to wind sector and specific periods? Please make this more clear in this sentence.

The sentence will be rephrased accordingly in the revised MS.

12.- Line 267: "Specifically, for each crop parcel ..." -- this should be commented in Discussion in the context of whether this is accurate also for several years and how the transition from senescence or dormancy to growing season... by daily sum of NEE, by meteorological inputs, by phenology, etc. Generally, crops' carbon sink during growing season starts to decline before harvest, as crop ripes, therefore it could have an impact on when and how the growing season is defined.

It will be commented in the Discussion section of the revised MS in that context as rightly suggested by Referee 3.

13.- Line 280: How was the qualitative day-night indicator defined? And why was it qualitative? Couldn´t it be better to define day or night based on, e.g., radiation? Please clarify this.

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}. We will add this clarification in the revised MS.

14.- Line 306: "This evaluation framework..." - this sentence is used in similar forms in other subsections in Methods; maybe it could be avoided or just write it once in discussion to summarize why this approach was adopted.

It is true and will be modified accordingly in the revised MS.

15.- Line 310: "Predictions were made ..." - this is not necessary, as it was clear from before that the time resolution is 10 min, and the units of CO_2 flux densities.

In the revised manuscript, this sentence will be removed as this information was already given in the manuscript.

16.- Line 311: The factor 0.0072 comes from the number of seconds in a year, multiplied by molar mass of carbon and the conversion from umol to mol. However, either this should be clarified for unexperienced readers or be left out as it is common knowledge in EC studies and it is just a simple calculation.

"using a factor of 0.0072" will be removed in the revised manuscript as obvious/common calculation in carbon flux studies.

17.- Line 314: Please rewrite the whole paragraph for clarity.

Since we carried out the partitioning of simulated NEE into GPP and Reco, this paragraph will be completely rewritten in the revised MS.

18.- Line 326-329: This paragraph could be in Methods.

This information is already given in the M&M section but we would like to keep it here at the beginning of the results section for a reminder and a better understanding of the coming results and analysis next.

19.- Line 342: particularly at small flux values.

We will modify the sentence accordingly in the revised manuscript.

20.- Line 343: what is meant by unmodeled management practices? Please clarify. As I understand it, no specific management was introduced in the RF models, just the predictor variables and the classification in wind sectors and periods.

This will be clarified in the revised MS.

21.- Line 368: This is only a particular 30-min period, so it shouldn´t be taken as the example; furthermore, a flux density of -74 umol m-2 s-1 seems too large. Looking at the Figure, aren´t these values occurring in the hybrid parcel? Maybe it is my mistake, but I do not see these values happening in the wheat parcel, but in the hybrid.

It is true and we thank Referee 3 for her/his comment; please see again our response to your comment 5- above to get modifications done and presented in the revised version of the MS.

22.- Line 504: "This lower performance...". Why is this happening? It shouldn´t be the case.

Please see our response to your comment 5- and other Referees according to our modelling approach performance and associated modifications done in these revisions and that will be presented in the revised version of the MS.

23.- Line 512: if there was noise introduced by some fluxes that were not filtered, why not trying a different filtering approach, maybe more strict, or using e.g. hard limits?

In the present study (as our previous ones), we do our best to qualify all our EC data set, filtering bad data that do not fulfilled EC theory condition (stationary test) or bad weather conditions (stormy events, etc.) without being too abusive on the other hand removing potentially good data. After applying all our processing and filtering data

procedure, we already removed more than 30% of data (32%) (l. 204 in the submitted manuscript). For this particular stormy event in autumn 2023, we particularly checked corresponding EC data we kept based on our flux quality check procedure described in the submitted manuscript (M&M section) though a higher flux magnitude was logically recorded during these stormy conditions under high wind speeds in particular. However, owing to our further computations particularly described in our response to your comment 5- with regards to our flux data filtering, we will modify accordingly this part of the discussion in the revised version of the manuscript allowing to be more accurate in the description and discussion of the measured flux data in comparison with the modelling approaches we present.

24.- Line 516: from this sentence it is not clear whether the spurious values were included or not in the dataset that was used to train the model. Please clarify.

It will be clarified accordingly in the revised MS.

25.- Line 576: The annual NEE sums (not carbon budgets unless carbon export was considered, as commented previously) should be reported in results.

Referee 3 is right and NEE sums instead of carbon budgets will be reported in the Results section of the revised MS.

26.- Line 514: maybe this table could go in Appendix, as it is just a reference for literature values.

We would like to keep this table in the main text (discussion section) as it stands, instead of going in Appendix; indeed, this later gives relevant information on other EC fluxes in term of crops, methodology and flux values (range) allowing consistent comparisons with results we obtained and discussed in the present study.

27.- Generally, the past tense is used for Methods and Results, and partially on Discussion when referring to results. Although this is an election from the author, I would recommend to check this for consistency across the manuscript.

It is right and we will be more careful according to tenses used in the whole revised manuscript and corresponding sections.

Technical corrections

We thank Referee 3 for all her/his technical corrections we will address in the revised version of the manuscript.

Line 37: remove "as".
Line 42: suggestion to re-write as "as both sink or sources of atmospheric CO2".
Line 78: meteorological.
Line 93: write "a RF-based" instead of "an RF-based" as RF is read as "random forest".
Line 154: These eddies occur/are located.
Line 172: ... includes the following sensors positioned at the top of a 3-m tower mast (Fig. 2b): an ultrasonic anemometer...
Line 175: The station also included additional sensors...
Line 261: The first strategy (Scenario 1) involved...
Line 265: In the second strategy (Scenario 2) separate RF models were trained for...
Line 354: Box in red represents November's storm.
Line 400: Please avoid statements like "clear".

REFERENCES

We thank Referee 3 for references below we did browse and take into account in our revisions and the corresponding revised version of the manuscript.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Barr, A. G., Black, T. A., Arain, M. A., Bourque, C. P.-A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Assessing Eddy-Covariance Flux Tower Location

Bias across the Fluxnet-Canada Research Network Based on Remote Sensing and Footprint Modelling, Agricultural and Forest Meteorology, 151, 87–100, https://doi.org/10.1016/j.agrformet.2010.09.005, 2011.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the Separation of Net Ecosystem Exchange into Assimilation and Ecosystem Respiration: Review and Improved Algorithm, Global Change Biology, 11, 1424–1439, https://doi.org/10.1111/j.13652486.2005.001002.x, 2005.

Levy, P., Drewer, J., Jammet, M., Leeson, S., Friborg, T., Skiba, U., and Oijen, M. V.: Inference of Spatial Heterogeneity in Surface Fluxes from Eddy Covariance Data: A Case Study from a Subarctic Mire Ecosystem, Agr. Forest Meteorol., 280, 107783, https://doi.org/10.1016/j.agrformet.2019.107783, 2020.

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.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of Net Ecosystem Exchange into Assimilation and Respiration Using a Light Response Curve Approach: Critical Issues and Global Evaluation, Global Change Biology, 16, 187–208, https://doi.org/10.1111/j.1365-2486.2009.02041.x, 2010.

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.

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.-Biogeo., 127, e2021JG006240, https://doi.org/10.1029/2021JG006240, 2022.

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.

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.

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, Atmospheric Measurement Techniques, 17, 6047-6071, https://doi.org/10.5194/amt-17-6047-2024, 2024.

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

Rannik et al., Good performance of low-cost carbon dioxide sensor based on intercomparisons with the standard eddy-covariance system, preprint, https://doi.org/10.5194/egusphere-2026-144.