Gottschalk et al. present a high-resolution investigation of the vegetation indices against the carbon fluxes from a tower in the cropland. They processed the flux observations thoroughly to ensure the data quality and estimated the fluxes based on linear models using VIs, which were then compared with crop models. They showed that observed fluxes correlate well with the field management actions, phenological development of crops, and some Sentinel-2-derived VIs (especially for GPP). They concluded with discussions of the limitations of their analyses (e.g., adopting a simple linear model).

Major comment:

In general, the data processing is thorough, and the results are well presented. Leveraging the high-resolution feature from Sentinel-2 data, their EC-based carbon fluxes and Vis are all matched to the plot scale, which addressed the common spatial mismatch issue in prior studies.

Despite the promising flux-VI relationship presented, my major concern lies in the novelty of this study and the representativeness of such flux-VI relationships across crop traits and climates (despite that the authors acknowledged till the very end). The authors suggested that their analyses indicate "the suitability and developability of the proposed approach to monitor cropland C exchange with satellite derived Vis*"* (L25-26) and acknowledged the current challenges in flux estimates being the low number of EC towers in croplands (L79-81). When considering croplands over different climates/regions, how robust is the use of a linear model to predict carbon fluxes? Would involving additional meteorological variables be more helpful (which is relative to statements on L490-491)?

Concerning the novelty of the study: many prior studies have indicated the potential of satellite images for carbon flux estimation (Sims et al. 2006, Huang et al. 2019b) but mostly at coarse spatial resolutions of MODIS-like 1km. Some recent studies explore the potential of finer-resolution satellite images such as Landsat or Sentinel-2 (Fu et al., 2014, Madugundu et al., 2017, Chen et al., 2010, Wolanin et al., 2019, Bazzi et al., 2024) but mostly only investigate one or two spectral indices while focusing on GPP over vegetated lands in general rather than dedicated cropland nor NEE. While EC is the most defensible approach to measuring ecosystem-scale carbon fluxes, many prior studies do not appropriately account for the spatial mismatch between EC measurement footprints and satellite pixel footprints. Tramontana et al. 2015 & 2016 also highlighted the need to address such a spatial mismatch. A lack in our understanding of satellite imageries' capabilities for cropland carbon flux estimation and monitoring is a rigorous evaluation of comprehensive spectral indices from finer-resolution satellite imageries by appropriately leveraging EC-based carbon flux measurements. Our study fills in this knowledge gap through a detailed and rigorous analysis into the relationship between a large list of satellite-imagery VIs and EC-measured carbon fluxes. In doing so, our study also paves a way to exemplar protocols to appropriately use EC time series and satellite-imagery VI time series together for investigating and monitoring crop carbon fluxes in future studies.

Concerning "suitability and developability": Following the scientific principle of occam's razor and in respect of the complexity of mechanistic crop models not outperforming clearly the simple linear regressions we argue that this study has its value by showing just that. The complex and heavy on input parameters and often assumptions models are not necessarily better than simpler models. By stating "suitability and developability" we acknowledge the unknown if different crop types/traits/environmental conditions require a higher resolution of linear models. But this was not the aim of this study. At the same time this means as well that

we acknowledge the research demand to address just these questions. We found one reference addressing the question of generalizability of the linear models (Juszczak et al. 2018) which argues that a generic single relationship between VI and C flux *can* be valid for a range of different crops.

Furthermore, we acknowledge the research demand for finding potential additional meteorological variables (in the conclusions) once this approach is applied at more sites and results indicate the necessity of including more explanatory variables. However, we do like to mention that additional variables also introduce additional sources of error/uncertainty at spatial scales because e.g. temperature at spatial scale will always be uncertain as well. However, this would have exceeded the scope of this manuscript and is topic of a follow-up study.

We will elaborate on these points in a revised manuscript.

The results/conclusions presented from this study did not seem to necessarily address the limitation of low data density of the flux observations, as the results were derived from one crop flux tower in Germany with seemingly low data density (compensating for better data quality).

The intention was not to address the low availability of flux observations but – amongst others – to leverage a cropland site with detailed flux measurements and meter-scale satellite images to achieve a detailed and rigorous analysis into the relationship between satellite-imagery VI and EC-measured carbon fluxes. In doing so, we also evaluate whether large-scale statistical approaches (as used for FLUXCOM) are actually justified when evaluated at plot scale.

Furthermore, this case study not only shines light on the great potential of satellite image-based VI for carbon flux estimation but also serves as a template how to appropriately use EC data stream and imagery time series jointly in later studies over more sites in different climate regions.

Moreover, regarding the high-resolution capability, are there any prior studies that utilized Vis from Landsat and Sentinel-2 for estimating carbon fluxes over croplands at meter scales?

Some prior studies use Landsat and/or Sentinel-2 images for estimating carbon fluxes over vegetated lands, mostly looking at GPP (e.g. Pabon-Moreno et al., 2022, Spinosa et al. 2023, Fu et al., 2014, Madugundu et al., 2017, Chen et al., 2010, Wolanin et al., 2019, Bazzi et al., 2024) but very few at NEE specifically over croplands. We will cite these prior studies in the discussion section. Our study uses meter-scale satellite images to also estimate NEE, which is of higher necessity and importance to calculate C budgets but is more challenging to estimate due to its indirect correlation with VIs.

In sum, the knowledge/technical gaps from prior studies and the novelty of this study can be better phrased. I would suggest conducting a more thorough literature review and rewording relevant text to highlight the advances of this study from prior studies (e.g., the examination of various VIs, the estimation of the carbon budget, the alignment of footprint between VIs and flux signals, and investigation of crop management signals from EC and VI data, from my understanding?).

Good point. We will elucidate and phrase better the novel aspects of this study in relation to more dedicated literature. To our knowledge there is only one study though dedicated at

assessing the impact of the alignment or not-alignment of EC with satellite footprints when estimating C fluxes (Kong et al. 2022). Furthermore, we will highlight better that our study fills the gap in exemplar protocols to appropriately use EC time series and satellite-imagery VI time series together for investigating and monitoring crop carbon fluxes in future studies.

Here are some minor comments:

Figures 2, 3: the time format on the x-axis is not super intuitive. Either change the date/time format or add a caption. What do the vertical lines in the dashed or dotted lines indicate in Figure 3? Add a few text labels to facilitate the comprehension of the discussion on page 18 (e.g., the timing for flowering on L273 and senescence on L276 - 278).

Yes, agreed. Same has been mentioned by RC1 and this formatting issue will be addressed as stated in the response to RC1: changing x-axis labels, add explanation of dashed and dotted lines to Fig. 3, adding horizonal segments indicating the timing of flowering and senescence.

L294-296: The poorer correlation between R_{eco} and VIs seems as expected. Does that suggest additional meteorological variables describing the air and soil columns are needed?

This is probably true when focusing at Reco. We will add relevant words in the discussion.

There are a lot of numbers going into the result sections and tables. There could be better ways to present the data more intuitively, e.g., replacing Table 6 with a bar plot.

Yes, we will provide the information also as a figure. We would still like to add the table in the appendix as it can be frustrating when results are only shown as plots when actual numbers are needed for comparison or reference.

References

Bazzi, H., et al. (2024). "Assimilating Sentinel-2 data in a modified vegetation photosynthesis and respiration model (VPRM) to improve the simulation of croplands CO2 fluxes in Europe." International Journal of Applied Earth Observation and Geoinformation 127: 103666.

Chen, B., et al. (2010). "A data-model fusion approach for upscaling gross ecosystem productivity to the landscape scale based on remote sensing and flux footprint modelling." Biogeosciences 7(9): 2943-2958.

Fu, D., et al. (2014). "Estimating landscape net ecosystem exchange at high spatial–temporal resolution based on Landsat data, an improved upscaling model framework, and eddy covariance flux measurements." Remote Sensing of Environment 141: 90-104.

Huang, X., et al. (2019). "Evaluating the Performance of Satellite-Derived Vegetation Indices for Estimating Gross Primary Productivity Using FLUXNET Observations across the Globe." Remote Sensing 11(15): 1823.

Juszczak, R., Uździcka, B., Stróżecki, M., and Sakowska, K.: Improving remote estimation of winter crops gross ecosystem production by inclusion of leaf area index in a spectral model, PeerJ, 6, e5613, 10.7717/peerj.5613, 2018.

Kong, J., Ryu, Y., Liu, J., Dechant, B., Rey-Sanchez, C., Shortt, R., Szutu, D., Verfaillie, J., Houborg, R., and Baldocchi, D. D.: Matching high resolution satellite data and flux tower footprints improves their agreement in photosynthesis estimates, Agricultural and Forest Meteorology, 316, 108878, https://doi.org/10.1016/j.agrformet.2022.108878, 2022

Madugundu, R., et al. (2017). "Estimation of gross primary production of irrigated maize using Landsat-8 imagery and Eddy Covariance data." Saudi Journal of Biological Sciences 24(2): 410-420.

Pabon-Moreno, D. E., et al. (2022). "On the Potential of Sentinel-2 for Estimating Gross Primary Production." IEEE Transactions on Geoscience and Remote Sensing 60: 1-12.

Sims, D. A., et al. (2006). "On the use of MODIS EVI to assess gross primary productivity of North American ecosystems." Journal of Geophysical Research: Biogeosciences 111(G4).

Spinosa, A., et al. (2023). "Assessing the Use of Sentinel-2 Data for Spatio-Temporal Upscaling of Flux Tower Gross Primary Productivity Measurements." Remote Sensing 15(3): 562.

Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E., and Papale, D.: Uncertainty analysis of gross primary production upscaling using Random Forests, remote sensing and eddy covariance data, Remote Sensing of Environment, 168, 360-373, https://doi.org/10.1016/j.rse.2015.07.015, 2015.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291-4313, 10.5194/bg-13-4291-2016, 2016.

Wolanin, A., et al. (2019). "Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations." Remote Sensing of Environment 225: 441-457.