

Interactive comment on “Cereal-legume mixtures increase net CO₂ uptake in a forage system of the Eastern Pyrenees” by Mercedes Ibañez et al.

Investigations aimed at improving C-sequestration in soil through adapted land use are of great importance because they could make an important contribution to the short- and medium-term mitigation of anthropogenic climate change. At present there are still too many knowledge deficits to fully exploit the potential of this approach. In view of this, the authors' intention to contribute to the solution of this problem is very reasonable and logical (introduction, lines 52-77). This also applies to the selection of forage cultivation systems and the use of the eddy-covariance technique for conducting the investigations.

Unfortunately, however, the manuscript is characterized by two serious deficiencies. These are so fundamental and at the same time irreparable that publication of the study in *Biogosciences* cannot be recommended.

- The chosen experimental approach does not reflect the current state of the art. This is comprehensively described and discussed e.g. in the publications of Smith et al. 2010, *Agriculture, Ecosystems & Environment*, 139, 302–315; Soussana et al. 2010, *Animal*, 4, 334–350; Chenu et al. 2018, *Soil Tillage Res.* 188, 41–52; Smith et al. 2020, *Global Change Biology*, 26, 219–241. If the authors had followed this approach, they would never have come up with the idea to characterize the climate impact of crop species solely on the basis of several months of NEE fluxes and seasonal NBP budgets. In order to determine the influence of crop species on the context-relevant C sequestration (longer-term storage of CO₂-C in the soil's C stock), annual NBP budgets would have had to be determined over a period of several years. Only then is it possible to avoid bias of the results due to the different temporal dynamics of plant C-input and C-output via soil C mineralization and interannual weathering variability. To be on the safe side, the CO₂ flux-based approach is now also combined with direct measurements of changes in the soil C stock. Since this did not happen, the authors have missed the self-set goal of their investigations. This is also indirectly admitted at the end of the discussion (lines 405-414).
- Contrary to the authors' assertions, the experimental approach used is only suitable to a very limited extent for clarifying the question of whether cereal-legume mixtures represent a stronger CO₂ sink than grain monocultures. Clear statements on this would have required the simultaneous investigation of cereal monocultures and cereal-legume mixtures. Since the authors have only examined the different cultivation variants one after the other in a crop rotation, they are not able to separate the direct effect of the respective crop on the CO₂ source function from the indirect preceding crop effect and the influence of the current annual weather. In addition, the form and amount of fertilizer applied varied between years, even with the same crop. This is a clear violation of the *ceteris paribus* principle, one of the most important prerequisites for obtaining clear results in experimental research. With the help of the diversity interaction model used, it is only partially possible to compensate for this deficit. This is because, when determining the so-called species-specific effects, the effect of the current random variation of the other factors is inevitably included. Finally, only the expected but trivial statement remains that the prolonged presence of photosynthetically active plants during the vegetation period can lead to a temporary improvement of the CO₂ sink function.

Minor deficits

- A lot of information and data that are important for the interpretation of the results are missing. This is especially true for

- Description of the study site: physical and chemical properties of the entire soil profile, cultivation history
 - Type and timing of tillage
 - Dealing with the above-ground phytomass: what is behind the yield and the C export? Only the amount of grain harvested or always the total above-ground phytomass? What and how much remained on the field in the form of harvest residues?
- In the case of triticale, results do not seem to be consistent. It is not plausible that GPP (Figure 3b) should be somewhat lower than the C yield (Table 1).