Referee comment on "Effects of climate change in the European croplands and grasslands: productivity, GHG balance and soil carbon storage" by Marco Carozzi et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-241-RC2, 2021

The manuscript describes the set-up for spatial runs of crop productivity and emissions from croplands and grasslands in Europe. This is analysed on simulation model results provided by two model simulation approaches for a historic (1985-2004) and a future period (2005-2999). The future scenario simulations include two scenarios, representing a moderate and an extreme climate development.

Spatial data about emissions of croplands and grasslands are very important and model approaches are most suitable to provide these spatial data. The data demand for models is high and the errors and uncertainties of used data as well as of the model results is high. This combination makes the judgement of the quality of spatial model studies on large scales, like this one here, very difficult. There is no doubt that the technical set up of the model is up-to-date and the used data represent the state of the art quality of available data. However, there are still a couple of genera aspects that let this manuscript down:
I mentioned the uncertainty and error affected data. I appreciate the open communication about the used data and the limitations in the approach. I see the extrapolation of the data for fertilizer application rates and crop rotation difficult. This extension introduces a different error and shifts the uncertainty compared to the period these data sets were originally aggregated for. Changed climate will affect the rotations and the fertilizer application rates and it is difficult to predict these changes. I do not mind that these data are used, but I would expect a more critical discussion on these points and a more careful conclusion.

I highly appreciate the detailed comparison of the different results with literature, but I see two problems with this. One is that some of the studies that are used for comparison are model approaches that used similar data and use similar model approaches (still worth to compare, but not necessarily completely independent). More relevant, I do not see the new or innovative contribution of the here presented data. There is a lot going on in the manuscript and it is not all linked with each other. Even though, croplands and grasslands are analysed, both are analysed parallel and not integrated. The most obvious connection (using manure from livestock production) is not included (I understand why and I agree to this point) and spatial emission budgets are not provided (as land use data are not used). Overall, the data are nicely presented, but more moderated than discussed. This is also reflected by the conclusion, which is more a summary of the manuscript and the conclusive sentences are not really convincing (lines 773-774: I do not think that the study proves this, even though it adds a strong indication and to conclude that C stock changes are variable over time does not require a complex study). I think that the errors and uncertainty and their impacts on the results should discussed stronger.

The presentation of the results for the historic period sounds partly like observed data or robust baseline. Considering the uncertainty of the use data (the crop rotation and fertilizer data are also simulation results, soil maps are developed over larger time periods and climate realisation are partly weak in the representation of extreme events) this is not a robust baseline. As mentioned above, the data that are used for the historic period are aggregated in a different way than the data for the future scenarios. While the trends, productivity and emissions during the historic period are driven by management changes (the farmers action is considered, as these are statistical data sets) the future scenario results are driven by climate impacts. Both can be presented and in the end they can be compared by a direct comparison including conclusions is difficult. Figure 5 shows the spatial distribution of the fluxes, but it would be great to see the emissions separated for croplands and grasslands. There are a couple of places where the N emissions surprise me. I would expect higher N emissions around Paris, less low emission spots (and more extreme emissions) in the Netherlands and I cannot follow the logic that North-West Germany shows lower emissions than the East of Germany (historically simulations). I am also critical about the results for Eastern Europe and, again, using these data for future extrapolation. Eastern Europe make these data doubtful as baseline for the future simulations. While management in historic data changed due to political changes or to adapt changing conditions, the management for the future simulations remains constant. I think simulations on the historic data are a good point to prove the functionality of the models, but I would only compare the future simulation scenarios with each other. Obviously, you can show general changes and trends, but I wouldn’t use the historic simulations or observations as baseline or reference. Errors and uncertainties are different in both data sets, which affects the results and the differences.

The spin-up for such a model approach is fairly short. Was the model tested for equilibrium in all grid cells? What were the criteria for equilibrium (no change over 10 years? Only 1 % change over a fixed period?). If the spin-up was only tested for an average, I wonder, if the authors think this is good enough? Where did the authors got the initial SOC values from or did the model used spin-up SOC?

I have some concerns about the management and the potential impact of the
management options on the results. I find the amount of 80% residues removed from the field very high, compared to other comparable model approaches. Do the authors think that this is a realistic number for Europe? I didn’t find any support for this number and would like to understand how the authors came to this result. Changing this number will change the NEP significantly. I also wonder, if the authors thought about introducing a second growing period. The shorter growing periods are only presented as a problem, but this can be used for double cropping seasons. Again, I do not mind limitation, by they should be discussed and in a best case the relevance (quantitative impact if possible) estimated.

In summary, I think this is an interesting study, but I do not see the clear objective. For me all the different points does not come together. There are plenty of options to improve the manuscript (focussing on single aspects; defining key/new aspects of the presented study; include a more critical discussion about the limitations, including estimates about their relevance; etc.). In the actual form I am not convinced by the manuscript, but there is definitely enough potential for improving it,

Line 51: become neutral

Line 179: species

Line 205-209: Was cutting and grazing on all grasslands?

Line 218: Do you mean 1901-1977?

Line 309: decreased (check writing).

Line 309: What is driving the yield reductions for maize: drought, heat, water or nutrient limitation? A shorter growing period or other weather impacts have a stronger impact for
spring crops, as the growing season is short and impacts show a stronger effect.

Line 360: Higher productions

Line 654: lower case 2

Line 686-689: Is this a realistic effect or a model artefact? I would expect increased emissions and leaching, as there is increased soil water content and water transport in the soil. Even though the nutrient uptake by plants is increased, I would still expect an increase of leaching and emissions. Only optimised systems would show a decrease.

Lines 707-710: Does the decline of NBP means stronger sink? Is this not affected by increased production due to rising temperature in an area with sufficient water availability?

Line 779: agro-ecosystem

Figure 1: please mention the spatial unit that is used here. Does each point represent a NUTS2 unit or a country?

Figure 5: Please indicate in the legend delta N2O.