REVIEWER 1

# Recommendation

Reject due severe flaws in the methodological approach

# Comments to the Authors of the article: *Soil organic carbon stocks as driven by land use in Mato Grosso State: The Brazilian Cerrado agricultural frontier*

## **Overall comments/suggestions**

The manuscript discusses an important topic regarding the dynamics of soil organic carbon (SOC) in relation to land use changes in Mato Grosso State, Brazil. Initially, this article proposes an interesting approach combining modeling approach from Machine Learning and InVEST. Also, the focus on soil carbon is highly relevant, given its importance in carbon sequestration and climate change mitigation. However, it fails to deliver an analysis with coherent methods. The major flaws are in the lack of appropriated land change modeling approach that is presented in the Table 2 and Fig 3, as well as incomplete description of the Machine Learning and other aspects that may lead to issues in the interpretation of results. Therefore, the goal in the abstract isn’t grounded in the paper analyses but in common knowledge available in the literature, i.e. “work aimed to unveil the nexus between long term changes in land use and changes in carbon stocks” is mostly supported by references in the results such as:

*“In a global scale, it has been reported that if the SOC content in the topsoil in cropland increased from 0.27% to 0.54%, a stock ranging from 0.56 to 1.15 t C.ha.yr-1 could be sequestered , and this could represent 0.90 to 1.85 Pg C yr-1 for at least a continuous 20 years of the sequestration [65]. However, in contrast to the findings in our study, a recent study in China concluded that conversion from grassland to cropland produced a negative SOC stock [66].”* (page 22, line 443-447).

Moreover, I’d strongly suggest a clear guiding question and hypothesis in the introduction or methods to the reader grasp on the purpose and contributions of the paper.

See below comments breaking down more issues in the modeling, and other minor suggestions.

## **Major comments**

First, the **Machine Learning approach**. You need provide more details to the ML such as the turning process, the regulation method to optimize the loss function (e.g. lasso, ridge), and the results comparing MSE, etc. I didn’t see any aspect of the ML being explored in the results, or how it could inform the “nexus between long term changes in land use and changes in carbon stocks”.

The Table 2 show the results of a **land use change modeling to 2050**. There are many questions here about the method and a non-cited literature. Still, looking only to the results, the reader can empirically question “How can you model the expansion of settlements, if this is defined by land policy? i.e. settlements is not a continuous and stationary data to be easily modeled. Diving into the approach, there is an extensive literature on land change model that is ignored. In a land change model, you must define proximate and underlying drivers (e.g. see Geist & Lambin 2002) while the environmental variables are endowments to the spatial distribution. Moreover, you must describe how the rates of land change were defined (e.g. historical, scenario description, econometric approach, etc). In contrast, your methods confusedly mentioned the InVEST to model land use change amid the SOC stocks, but this software is not fitted to this type of model. Besides, any modeling requires an accuracy test that is not explored by the paper.

The Table 3 present an **OLS regression.** You must describe it in the methods, including the population model and the relationship for each variable. For instance, the NDVI is explained by SOC not the other way around, hence, what is your argument to use NDVI as explanatory variable to the SOC?

Lastly, there are many studies about the Land use change in this region, as well as recent SOC estimates for long periods, such as Mapbiomas. A quick search using Google Scholar will give you an extensive list. The reader will question why you didn’t cite or compare your results with the literature.

## **Minor comments**

Overall, **I strongly recommend review the writing to check typos and ensure a plain English style.** For example, the introduction is informative but overly verbose, so you could reduce redundancy to give space to the goal and question that you’ll address in this paper. Furthermore, there are minor typographical errors, such as "builds the largest terrestrial carbon storage" instead of "represents the largest terrestrial carbon storage". Ensure consistent verb tenses throughout the paper. Some sentences switch between past and present tenses, which can confuse readers.

In the content, **please consider that most readers are non-experts in the Brazilian context**, thus, you must explain the relevance of the cited programs Renovabio and ABC Plan to the topic of SOC.

**Regarding the article structure**, I suggest you re-organize it. For instance, the final paragraph of the introduction could be in the methods. Also, a good portion of the results could be moved to a discussion section.

In the Table 1, **google earth is not a data source**, it is a platform. Thus, what is the exact source? e.g. TerraClimate, ERA5, etc. Similarly, what are your specific variables for climate? Include units and maybe a description of the mean and variance in the area of interest.

Finally, readers honestly expect a **decent cartography** work for a paper, instead of a screenshot such as Fig 4.