Author comment for RC1

Dear Anonymous Referee #1,

We sincerely thank you for the time and effort you devoted to reviewing our manuscript. We greatly appreciate your positive and constructive comments, which have helped us improve the manuscript's scientific clarity, logical flow, and overall structure. We have carefully considered each of your comments and revised the manuscript accordingly. Our detailed responses to your comments are provided below.

Black- Reviewer comments

Green -Author comments

Underlined text - new text/ changed text to add to the manuscript

RC1:

In this manuscript, the authors used a boundary line approach to define attainable soil organic carbon (SOC) content based on several explanatory variables that could limit SOC storage. They developed boundary line models in southern Australia and applied them at the farm level to illustrate their use. The authors then produced an interesting map of limiting factors for the farm. The methodology is sound, and the manuscript deserves to be published with a few adaptations. Below are my comments and recommendations for improving the manuscript.

The progression is very slow in the Materials and Methods and Results sections. The authors are encouraged to adopt a sharper writing style by presenting the methodology synthetically at the beginning of the Materials and Methods section then detailing the different steps, making better use of Fig. 3, and avoiding repetitions. This would also solve the problem of information in the Materials and Methods section being mixed (for example, BLA and DSM are briefly presented in Section 2.1 Study site and dataset). The description of the subsoil model results is especially lengthy (BLA figures, attainable topsoil SOC, and limiting factor map). Adopting a broader perspective would be more efficient, by highlighting the differences compared to the topsoil models. Additionally, many sentences in the Results section read more like discussions.

Response: Thank you for your positive assessment and constructive comments. We appreciate your recognition of the methodological strength and practical relevance of our Boundary line analysis (BLA) framework. We agree that the Materials and Methods section did not present the analytical workflow clearly and concisely. We will provide a brief overview of the methodological framework at the beginning, followed by a step-by-step description of each step. Figure 3 has been revised to clearly demonstrate the sequence of the study (as noted in the later comment). Figure 3 will be moved to the front of the Method section, and the description of the methodology will reflect Figure 3. We also agree that some information was mixed in Section 2.1. We recognised this and will revise the text accordingly.

We also agree that the results section is overly lengthy, particularly the description of the subsoil findings. We will merge the topsoil and subsoil BLA results (Lines 318–462) and

substantially condense this section by focusing on the key findings and differences between topsoil and subsoil, rather than describing every detail of the BLA model results. We will also shorten the description of the attainable SOC and limiting-factor maps (Lines 523–634), highlighting only the most important spatial patterns and dominant constraints.

The discussion section could be expanded:

- Attainable SOC is often calculated from mineral-associated organic carbon content. I think you are right to proceed this way but please discuss this difference with the literature.
- The delineation of attainable SOC is based on the area with the highest density on the plot SOC = f(soil property). Therefore, the results highly depend on the dataset and sampling design used, maybe more than the quantile method.
- I think it is necessary to discuss the values of attainable SOC found, at least using https://doi.org/10.1071/SR23096 and https://doi.org/10.1111/gcb.17053 also developed in Australia.

Response: Thank you very much for these important suggestions. We will expand the Discussion to clarify that this study estimates attainable total SOC concentration, whereas previous Australian studies, Karunaratne et al. (2024) and Viscarra Rossel et al. (2023) have focused on mineral-associated SOC (MAOC), also referred to as Humus Organic Carbon (HOC). Our study focuses on total SOC, which is inherently more variable than the stable MAOC pool. However, the use of total SOC is important from a practical perspective because growers routinely measure total SOC rather than MAOC.

We will also include statements that the BLA estimates (density-based selection procedure) depend on the sampling design, dataset, and distribution of observations near the upper boundary. We will also clarify that the estimated values represent attainable SOC under the existing conditions of the study area, rather than universal maximum values. Therefore, they reflect attainable SOC under the conditions represented in this study rather than universal maximum values. Finally, we will compare our findings with those of Karunaratne et al. (2024) and Viscarra Rossel et al. (2023), while recognising differences in carbon fractions, units, spatial scales, soil depths, and modelling approaches. It should also be noted that the BLA estimates can be updated as more data becomes available, which will improve their generalisability.

Because of the LOESS method used to adjust boundary lines, no equations are shared with the readers. How do you plan to disseminate the results?

Response: Thank you for highlighting this issue. The underlying data is private grower data and cannot be shared. We will publish the code for performing the BLA with a simplified and shareable dataset for reproducibility purposes. In the future, our intention is to create a RShinyApp (or something similar) where people can enter their data and location to get an attainable SOC estimate for their location.

Please, see also the following minor comments:

In the introduction, some of the phrasings are too abrupt: “boosts soil water holding capacity”, “stop soil loss”.

Response: Thank you for highlighting this. We agree that the expressions “boosts soil water-holding capacity” and “stops soil loss” were overly abrupt. These phrases have been revised to “improves soil water-holding capacity” and “reduces soil susceptibility to erosion and land degradation,” respectively.

Proposed text:*”Furthermore, it enhances the stability of soil aggregates (Wagner et al., 2007; Ayoubi et al., 2012; Bünemann et al., 2018; Wang et al., 2019), improves the availability and cycling of nutrients (Wang et al., 2019), and reduces susceptibility to soil erosion and land degradation (Mirchooli et al., 2020; Kiani-Harchegani and Sadeghi, 2020).”*

l. 54: Why do you add “although SOC dynamics rely on carbon inputs and losses from the system”.

Response: Thank you for your comment. We agree that the phrase “although SOC dynamics rely on carbon inputs and losses from the system” did not provide a clear contrast and was therefore unnecessary. We have revised the sentence.

Proposed text: *“SOC loss is common in agriculture, particularly in intensively managed landscapes where carbon inputs are insufficient to compensate for carbon losses through decomposition, erosion, crop removal, and other management-related processes”.*

l. 60: Remove “are highly varied.”

l. 61-62: Please rephrase or make a better link between site-specific factors and soil-forming factors.

Response: Thank you for your comment. We have removed “are highly varied” from the sentence. We have rephrased to provide better clarity on site-specific factors and soil-forming factors.

Proposed text: *“ Appropriate thresholds vary among sites because they are influenced by the combined effects of inherent soil properties, such as soil texture; soil-forming factors, including parent material, topography, climate, organisms, and time; and human interventions, particularly agricultural management practices (Jenny, 1994; Richter and Yaalon, 2012; Drexler et al., 2022; Prout et al., 2022).”*

l. 86-87: Please rephrase “determines the SOC for a given property”.

Response: Thank you for the comment. We have rephrased.

Proposed text: *"A regression-based model fitted through the scatter of observations, such as that used by Hassink (1997), estimates SOC in relation to the selected soil property but does not identify the upper limit of SOC or isolate observations representing non-limiting conditions."*

l. 106-107: Please rephrase "under a condition when a parameter".

Response: We have rephrased in the manuscript.

Proposed text: *"This boundary line indicates the maximum yield, often referred to as yield potential, for a given value of the parameter of interest."*

l. 117-125: I don't think it is useful for the understanding of the manuscript.

Response: Thank you for this comment. We agree that this paragraph was too detailed. We will therefore substantially shorten it and remove irrelevant text. However, we will retain a brief statement with the key references because these studies demonstrate previous applications of BLA for estimating crop yield, identifying yield-limiting factors, and assessing soil nutrient status (phosphorus and potassium). We think this provides relevant background for applying BLA to estimate attainable SOC in the present study.

l. 139: Avoid using abbreviations ESP and EC here.

Response: Thank you for your suggestion. The full forms of ESP and EC have been included.

l. 140-141: Rather group this with your objective (ii).

Response: Thank you for the suggestion. We reviewed this and found that this sentence is not needed and thereby removed the text.

l. 144: Add if SOC is a concentration or a stock.

Response: Thank you for the suggestion. It is SOC concentration. It is rewritten accordingly.

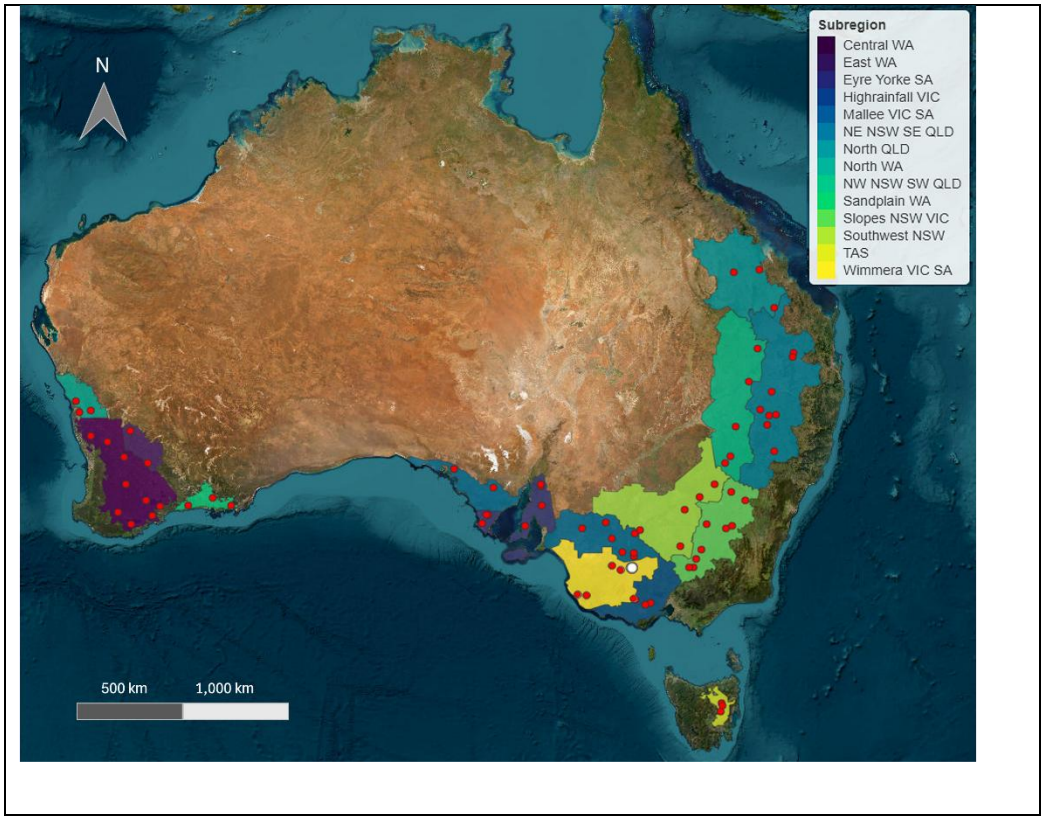
Proposed text: "This study focuses on attainable soil organic carbon (SOC) concentration in the 0–30 cm soil layer and the soil properties that limit its attainment."

Fig. 1: Maybe you need to add the source of the image and of the agroecological zones in the figure caption. Add the meaning of the red and white symbols in the legend or figure caption. You can also consider removing this figure because the different regions are not used then, and

only showing the outline of the entire grain growing zone in Figure 2 (then it will be quicker to say that you have stratified your dataset with the rainfall classification).

Response: Thank you for this suggestion. We have retained Figure 1 because it provides useful geographical context by showing the locations of the sampled farms across Australia's major GRDC grain-growing regions. Although these regions were not directly used in the analysis, the figure helps readers understand the spatial coverage of the dataset and the distribution of the farms within the main grain-growing areas. We will revise the caption to include the sources of the image and agroecological zones. We have also clarified the meaning of the red and white symbols in the caption. Figure 2 is presented separately to show the classification of farms based on BOM seasonal rainfall regions.

The revised Figure 1 is as follows:



l. 160-162: Rephrase to better fit a Material and Methods section (not a Discussion section).

l. 161: "and the same commercial laboratory was used", add that it is for soil analysis.

Response: Thank you for the suggestion. The sentence has been rephrased to make a better fit for the method section. And also added "and the same commercial laboratory was used"

Proposed text: *"All farms were sampled using the same protocol over a 26-month period (2023–2025), with sampling focused on grain-growing farms. All soil samples were analysed by the same commercial laboratory using consistent analytical procedures. This reduced variation caused by differences in laboratory methods and land-use differences."*

Title 2.2 could be simplified.

Response: Thank you for the suggestion. The title of Section 2.2 has been simplified accordingly.

Proposed title: "Distribution of farms across seasonal rainfall classes."

Caption of Fig. 2: Remove or move "Sites span six rainfall regimes, demonstrating the wide climatic range represented in the dataset" to Section 2.2.

Response: Thank you for the comments. This has been removed from the caption.

Table 1: Move the column "Rainfall region" to the left. Column "Classification ID" is not useful.

Response: Thank you for the suggestion. Rainfall region column has been moved to the left, and the Classification ID column has been removed.

The revised Table 1 is as follows:

Rainfall region	***Soil sample sites***	***No. of farms***
Summer 350-650 mm	*324*	*13*
Uniform 500-800 mm	*183*	*8*
Uniform 250-500 mm	*75*	*3*
Winter 250-500 mm	*225*	*9*
Winter dominant 500-800 mm	*300*	*12*
Winter dominant 250-500 mm	*675*	*27*

l. 212: Rephrase "Subsoil is about processes that control plant growth and so impact on biomass production, which relates to biomass inputs into the topsoil".

Response: Thank you for this comment. The sentence has been rephrased.

Proposed text: *"Subsoil properties influence plant root growth and, therefore, biomass production, which can affect organic carbon inputs to the topsoil."*

l. 220: End with ":".

Response: Corrected accordingly.

l. 228-230: Theoretically, your kernel density approach should be enough to remove outliers. Why not here? (It is just curiosity, no modification required)

Response: Thank you for this comment. We agree that KDE can theoretically be used as a standalone density-based method for identifying outliers. However, in our framework, Mahalanobis distance and KDE were applied for different purposes.

Mahalanobis distance was used as an initial global screening step to identify extreme observations in the bivariate predictor–yield space. In contrast, KDE was applied locally within predictor intervals to identify a density-supported upper yield envelope. Observations with low KDE density were not removed from the dataset; they were excluded only from the points used to fit the boundary line.

The preliminary Mahalanobis filtering also reduces the potential influence of extreme observations on KDE bandwidth estimation and the resulting density surface.

l. 240: Not useful.

Response: Thank you for your comment. We have removed this equation.

l. 247: and is time-consuming.

Response: Thank you. The text has been revised

Proposed text: *"Nevertheless, this approach is often time-consuming, particularly when working with large datasets".*

l. 256: "but retained in the data cloud". Remove, "data cloud" not used then.

Response: Thank you for this comment. The phrase "but retained in the data cloud" has been removed, and the sentence has been revised.

Proposed text: *"The low-density data points were not used to fit the boundary-line model but were retained in the dataset."*

l. 253-263: Could you please explain this more directly?

Response: Thank you for this comment. We will revise Lines 253–263 to explain the methodological steps more directly.

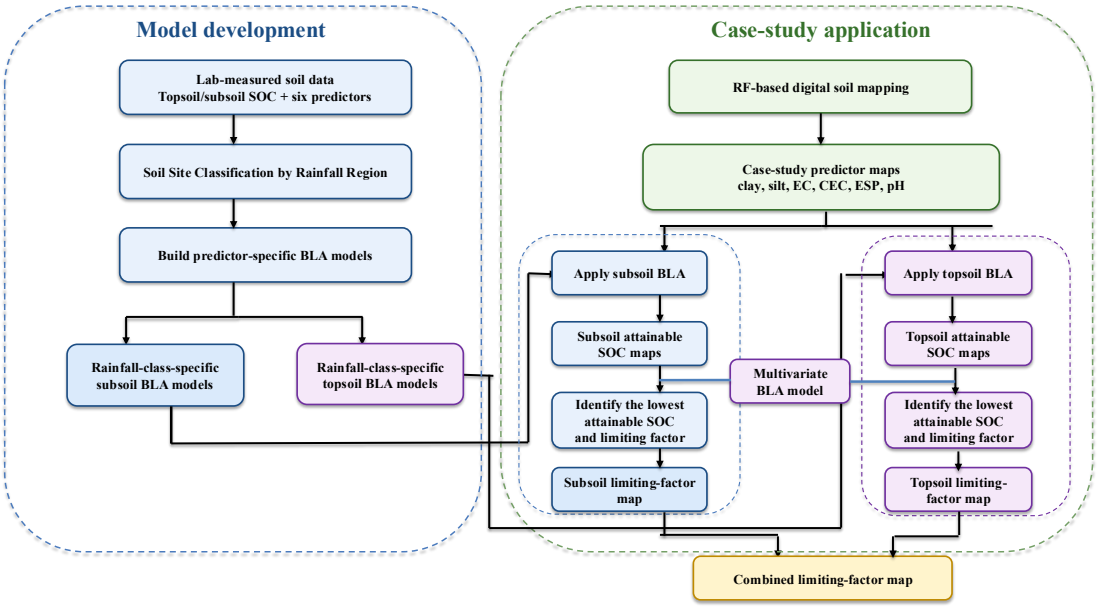
l. 302-303: “The above-mentioned steps are significant as they involve identifying the most limiting factors influencing the SOC accumulation and stabilisation for the case study farm”. Remove?

Response: Thank you for this suggestion. We agree that the sentence was unnecessary and have removed it.

Fig. 3 is crowded. You could reduce the information to a minimum and split it between a development and test part. Refer to Fig. 3 in the text.

Response: Thank you for this comment. We agree that Figure 3 (work flowchart) seems too crowded. So, we simplified and substantially revised Figure 3, splitting model development and application. This provides a clear sequence of the methodological steps and their relationships. We will also refer to Figure 3 in the corresponding text.

The proposed Figure 3 is given below



l. 364: Better explain how you computed a boxplot from the BLA models (actually, I don't understand how “the measured SOC demonstrated narrower variability than attainable SOC in all soil properties” in l. 367-368).

Response: Thank you for this comment. We have added the text to clarify how the boxplots were generated from the BLA model. We have also revised the statement concerning variability. Rather than referring generally to “narrower variability,” we now specify that the measured SOC distributions showed greater interquartile ranges than the attainable SOC estimates derived from the predictor-specific BLA models

Proposed text: *“Figure 5 compares measured topsoil SOC with attainable SOC estimated from*

the rainfall-region-specific BLA models using data from the 72 grain-growing farms. For each sampling point, the measured value of each soil property was entered into the corresponding BLA model for that rainfall region. The SOC value predicted from the fitted upper boundary was considered the attainable SOC for that predictor. These estimates were then grouped by rainfall region and soil property to construct the boxplots. Across the rainfall regions, attainable SOC was generally higher than measured SOC, indicating a consistent gap between current and attainable SOC levels. However, the magnitude of this gap varied among rainfall regions and soil properties. The measured SOC distributions also generally showed greater interquartile ranges than the corresponding attainable SOC estimates."

l. 405-407: And also due to correlations between topsoil and subsoil properties.

Response: Thank you for this suggestion. We agree that correlations between topsoil and subsoil properties may also contribute to the observed relationships. The sentence has been revised accordingly.

Proposed text: *"This is likely to be due to the indirect effects of their impact on plant productivity and therefore biomass inputs, as well as correlation between topsoil and subsoil properties rather than direct impacts of stabilisation and decomposition, which are more relevant for the previous BLA based on topsoil properties."*

l. 413-414: "The summer 350-650 mm region showed soil containing 40-50 % clay produced a moderate level of attainable SOC (1.25 %)". Two verbs, please rephrase.

Response: Thank you for your comment. The sentence has been rephrased

Proposed text: *"In the summer 350–650 mm rainfall region, soils with 40–50% clay were associated with a moderate attainable SOC concentration of 1.25%."*

l. 484-485: "From Figure 8, distinct spatial variability is evident throughout the case study farm, emphasising regions with differing soil properties". Please rephrase.

Response: Thank you for the comment. The sentence has been rephrased.

Proposed text: *"Figure 8 shows substantial spatial variability across the case-study farm, reflecting differences in soil properties among locations."*

l. 493: "Silt+clay represents the summation of clay and silt content". Can be removed.

Response: Thank you. This has been removed.

l. 495: Replace "firm" by "farm".

Response: Replaced.

l. 496: Vertosol and soil description can be moved to section 2.1.
Response: Thank you for the suggestion. This has been moved to section 2.1, added next to line 19.

l. 501: You are moving from the present to the past tense.

Response: Thank you for this comment. The sentences have been revised to ensure consistent tense throughout.

l. 505: Rephrase "Textural qualities".
Response: Thank you. Yes, rephrased.

Proposed text: *"Soil textural properties, such as clay and silt content, typically showed higher prediction errors and were predicted with low to moderate accuracy."*

Caption of Fig. 5: Add definitions for CCC, NNSE and RMSE to understand the table without reading the text.

Response: Thank you for this suggestion. Definitions of CCC, NNSE, and RMSE have been added to the Table 5 caption to ensure that the table can be interpreted independently of the main text.

Proposed text: *"Table 5. Model performance metrics for the validation dataset. R^2, coefficient of determination; CCC, concordance correlation coefficient; NNSE, normalised Nash–Sutcliffe efficiency; RMSE, root mean square error."*

l. 523-526: Can be removed.

Response: Thank you for your comment. Yes, we have removed the text you mentioned.

l. 530: Refer to Fig. 9.

Response: Thank you for this comment. Figure 9 is now referred to in the corresponding text.

Fig. 10: Homogenize "Min SOCatt" and "SOCatt" in the figure and caption (I guess it is the same).

Response: Thank you for this comment. The terminology has been standardised between the figure and its caption.

l. 614-617: Please shorten.

Response: Thank you for this comment. The description has been shortened accordingly.

Proposed text: *"A subsoil limiting-factor map was generated using a multivariate BLA model (Fig. 12a). Following the principle of the law of the minimum, the soil property associated with the lowest predicted SOC at each location was identified as the dominant limiting factor."*

Fig. 12. Title of panel a, above the map: is it topsoil or subsoil?
Response: Thank you for pointing this out. This is subsoil. It has been corrected.

l. 655-656: "Jien et al. (2025) also demonstrated that the BLA estimated SOC sequestration potential was 2.1 times higher than that of quantile regression, reinforcing the suitability of the BLA model in determining SOC". I would not give so much importance to this 2.1 ratio because this ratio highly depends on the arbitrary choices Jien et al. (2025) made for both methods (i.e., which quantile is used). You would find another ratio with your dataset. The main advantage of your approach is rather the ability to adjust nonlinear functions.

Response: Thank you for this comment. We agree that the 2.1-fold higher SOC estimate reported by Jien et al. (2025) does not necessarily mean that BLA is better than other methods. We have therefore removed the emphasis on this ratio and revised the text to focus on the main strength of BLA: its ability to model nonlinear boundary relationships.

Proposed text: *"Jien et al. (2025) reported higher SOC sequestration potential using BLA than quantile regression. However, the difference depended on the selected upper-boundary threshold, particularly the use of the 90th quantile. Using a different quantile could change the fitted boundary and the estimated SOC sequestration potential. The main strength of BLA is its ability to capture nonlinear boundary relationships between SOC and potentially limiting soil properties."*

l. 660-664: "In contrast, the attainable topsoil SOC derived from the subsoil-based BLA model was calculated to be 1.01 % (mean minimum) -1.39% (mean maximum), indicating a further sequestration potential of 0.33-0.71%. These findings highlight…". It is not a "further sequestration potential", it is another way to calculate it, isn't it? Please, clarify.

Response: Thank you for pointing this out. We agree that the term "further sequestration potential" was misleading because the subsoil-based BLA provides an alternative estimate rather than additional sequestration potential. We have revised the text to clarify that the reported SOC gap was estimated using subsoil properties.

Proposed text: *"In contrast, the attainable topsoil SOC derived from the subsoil-based BLA model was calculated to be 1.01 % (mean minimum) -1.39% (mean maximum), which represents an alternative estimate of SOC sequestration potential based on subsoil constraints."*

l. 689-688: “Across all rainfall regions, attainable SOC is found to be higher in the topsoil compared to the model built on the subsurface soil in all soil properties. This is expected as topsoil receives direct organic matter such as plant residue, litterfall, organic compost and microbial activity”. You did not use subsoil for this reason, I don't understand why you give such an explanation.

Response: Thank you for highlighting this. We agree that the original explanation was unclear, as both model sets estimated attainable topsoil SOC and differed only in whether topsoil or subsoil properties were used as predictors. We have revised the text to clarify this distinction. The discussion of organic matter inputs and depth-related microbial activity is to explain why topsoil properties may have a more direct relationship with topsoil SOC, whereas subsoil properties are more likely to influence it indirectly through plant growth, water and nutrient availability, biomass production, and the relationship between topsoil and subsoil conditions.

Proposed text: *“Across all rainfall regions, the topsoil-based BLA models generally estimated higher attainable topsoil SOC than the subsoil-based models. This may partly reflect the more direct relationship between topsoil properties and the processes governing topsoil SOC accumulation and stabilisation. The topsoil receives substantial organic inputs from crop residues, roots, microbial residues, and organic amendments where applied (Clara et al., 2017). In contrast, microbial activity generally declines with soil depth (Preusser et al., 2017), while subsoil properties may influence topsoil SOC indirectly through their effects on plant growth, water and nutrient availability, biomass production, and correlations between topsoil and subsoil conditions.”*

Quadratic curve: when this term is employed alone, we cannot know whether the parabola opens upward or downward.

Response: Thank you for this observation. We agree that the term “quadratic curve” does not clearly describe the direction of the relationship. The text has therefore been revised to indicate whether the fitted parabola opens upward or downward.

Homogenize the use of the terms SOC storage, sequestration, accumulation and buildup throughout the manuscript.

Response: Thank you for your suggestions. We will revise and homogenise the terminology throughout the paper.

Consider producing a single figure with panels for the different limiting factor maps.

Response: Thank you for the suggestion. However, we retained separate figures because the topsoil limiting-factor map is closely linked to the ESP amelioration analysis shown in Figure 10, whereas the subsoil and combined limiting-factor maps are part of a separate analysis.