General comments: This is a well-organized study that addresses a genuine gap: whether in-situ Vis–NIR spectroscopy can predict SOC fractions (POC/MAOC) in forest soils, and how spectral correction (EPO) and model transfer strategies affect that feasibility. The four-workflow comparison is a clean way to isolate the effects of spectral correction versus direct in-situ calibration, and the addition of SHAP based interpretation and independent cross-year validation strengthens the manuscript considerably. The writing is generally clear and the figures are informative. That said, several points need clarification or stronger support before the manuscript can be fully accepted. Most of these points are addressable through clarification and revision.

General response: We sincerely thank the reviewer for the constructive and well-encouraging comments. We have carefully considered all of the comments and have revised the manuscript to address the points requiring clarification and stronger support. The corresponding revisions include clarifying the workflow design and model comparisons, strengthening the discussion of POC and MAOC prediction performance, improving the interpretation and presentation of the results, and carefully checking the manuscript for citation, formatting, and consistency issues. Detailed responses to each specific comment and the corresponding revisions are provided below.

Specific comments

1. **L23–24 vs. L345–347: The workflow A performance values reported in the Abstract (POC: R^2 = 0.80; MAOC: R^2 = 0.59) do not match the range reported for workflow A in the Results (POC: R^2 = 0.04–0.70; MAOC: mostly below 0.43). Please clarify which specific model/preprocessing combination the abstract values correspond to, and ensure consistency between the abstract and the results text.**

   Response: Thank you for pointing this out. We agree that the original presentation of the Workflow A results in the Abstract was unclear, as the reported R^2 values were not explicitly linked to the corresponding model and preprocessing combinations. We have revised this part of the manuscript. Specifically, we now explicitly report the best-performing model and preprocessing combination for each algorithm/target variable and clearly indicate the corresponding R^2 values. The revised presentation also distinguishes the performance of individual model configurations from the overall performance ranges across all tested configurations. This makes the source of the reported values transparent and ensures consistency between the Abstract and the Results section. The revised abstract part is as follows:

   'Direct transfer of the optimal laboratory-calibrated models (AB_SNV + MBL for POC; AB_SNV + Cubist for MAOC) to raw in-situ spectra substantially reduced prediction accuracy, with R^2 decreasing from 0.96 to 0.57 for POC and from 0.74 to 0.24 for MAOC. EPO correction showed model-dependent effects, improving prediction performance for some model configurations but reducing or negligibly affecting performance for others. By contrast, direct in-situ modelling achieved higher predictive accuracy than laboratory-to-in-situ transfer, with POC performing best

using EPO-corrected spectra with AB + Cubist (R^2 = 0.90), whereas MAOC performed best using raw in-situ spectra with AB_SNV + MBL (R^2 = 0.71).'

2. **L89-90: for four modeling strategies, spectral-correction-assisted in situ modeling is not included. It would better to mention both direct in situ modeling (direct and EPO corrected spectra).**

   Response: We have revised the Introduction as follows:

   'Specifically, we compare laboratory calibration, laboratory-to-field model transfer with and without EPO correction, and direct in-situ modelling using raw and EPO-corrected spectra under varying environmental conditions.'

3. **Fig. 2: "Rebuilding models" branch, the Yes/No decision box in the flowchart is not explained in the text, what criterion triggers rebuilding the KS train/test split?**

   Response: Thank you for pointing this out. We agree that the previous flowchart contained some ambiguity regarding the workflow for in-situ model development. We have therefore revised the flowchart by replacing "Rebuilding models" with "Building in-situ models" and adding the step "Inherit laboratory KS assignment" following the "Yes" branch. This clarifies that the in-situ models were developed using the same samples and corresponding training/test partition as the laboratory models, rather than independently partitioning the in-situ dataset. We have also added a clearer description of this procedure in Section 2.4.1 as follows:

   'The corresponding in-situ spectra from the same sampling locations were then assigned to the same training or test subset, rather than being independently partitioned. This common location-based partitioning ensured that the laboratory and in-situ datasets shared identical training and test locations, with corresponding samples assigned consistently to the same data subset.'

   We also add detail explanations in the Figure 2 caption:

   'Flowchart of the modeling framework. The train and test sets were partitioned using spectral-based Kennard–Stone (KS) sampling with location-level grouping. For in-situ applications, in-situ spectra inherited the same training and test assignments as the corresponding laboratory samples based on sampling location, rather than being independently partitioned. "Building in-situ models" decision separates laboratory-to-field model transfer from direct in-situ model development. When "Yes," the in-situ spectra inherit the laboratory KS training–test assignment based on sampling location and are used to develop direct in-situ models. When "No," the laboratory-calibrated models are transferred to the corresponding in-situ spectra. Gray arrows indicate shared steps; red dashed arrows indicate in situ-specific steps; red arrows indicate laboratory-to-field model transfer rather than direct in situ model development; and blue arrows indicate laboratory-specific steps.'

4. **L288–289: for claiming "stable spectral importance patterns.", please describe some indication of stability, e.g., variance across the ten local models, it would help.**

Response: We agree that the original terminology "stable spectral importance patterns" was ambiguous and imprecise. By "stable," we did not mean that feature importance weights were static or identical across individual test samples (which would contradict the fundamental mechanism of Memory-Based Learning). Rather, we intended to express that certain key diagnostic spectral regions (e.g., SWIR 2200–2400 nm for POC and VIS 500–600 nm for MAOC) consistently recurred as dominant contributors across representative local models, despite sample-level variability. To eliminate confusion, we have removed the imprecise term "stable" and replaced it with "recurring spectral contribution patterns" and "sample-level variability."

Specific Manuscript Revisions Made:

- In Section 2.6.2 (Methodology), we updated the SHAP calculation pipeline as follows:

'Ten test samples were selected using stratified sampling across the full target-value range to represent different positions along the target concentration gradient while maintaining computational feasibility. SHAP values were calculated separately for each local model, and absolute SHAP values were averaged across the ten local models to characterize recurring spectral contribution patterns. The consistency and variability of feature attribution across local models were further assessed using pairwise Spearman rank correlations and coefficients of variation (CV).'

- In Section 3.5 (Results), we reported the exact quantitative metrics (Pairwise Spearman correlation ranges, mean CV%) for POC and MAOC as follows:

'SHAP analysis was first conducted to interpret the best-performing laboratory models using spectra preprocessed with the AB_SNV transformation (Fig. 7a). Because MBL is a local modeling approach, SHAP contributions were evaluated across ten representative local models selected by stratified sampling across the target-value range. The feature-attribution patterns showed substantial sample-level variability. Pairwise Spearman correlations among the ten local models ranged from −0.51 to 0.68 for laboratory POC and from −0.61 to 0.79 for in-situ MAOC, while the mean CV of absolute SHAP contributions was 258.3% for laboratory POC and 126.1% for in-situ MAOC, respectively. These results indicate that the magnitude and ranking of spectral contributions varied among local models, reflecting the sample-customized nature of MBL. Nevertheless, several spectral regions showed recurring contributions across the representative local models, which were used to identify the dominant spectral regions associated with POC and MAOC.'

- In Figure 7 Caption, we revised the figure legend to explicitly label which optimal model configuration corresponds to each heatmap panel (Lab POC: AB_SNV + MBL, Lab MAOC: AB_SNV + Cubist, In-situ POC: AB + Cubist, In-situ MAOC: AB_SNV + MBL) as follows:

'Figure 7. Feature importance for soil organic carbon fractions (POC and MAOC) derived from SHAP values of the optimal models under laboratory and in-situ conditions. (a) Laboratory models: AB_SNV + MBL for POC, AB_SNV + Cubist for MAOC; (b) In-situ models: AB + Cubist for POC,

and AB_SNV + MBL for MAOC. Heatmaps display SHAP-derived feature importance for ten stratified test samples, representing local model-specific spectral contributions across the target-value concentration gradient. The bars above each heatmap represent the mean absolute SHAP values across the ten selected samples.'

5. **L328: On selection of the "three best-performing models", It would help to state explicitly that these three models were selected and carried forward (presumably PLSR, Cubist, MBL) rather than requiring the reader to infer this from Fig. 5/6.**

Response: We have revised Section 3.3 to explicitly state that PLSR, Cubist, and MBL were selected as the top-performing algorithms and carried forward for all subsequent in-situ application workflows (Workflows A - D).

Revised text (Section 3.3): *'Among all evaluated algorithms, PLSR, Cubist, and MBL consistently achieved the highest prediction accuracy. Consequently, these three top-performing models were selected and carried forward for all subsequent in-situ application workflows and model transfer analyses (Workflows A - D).'*

6. **L339: Fig.S4 was not included in supplementary data, please ensure it is available. It could be Fig.?**

Response: Thank you for pointing this out. During the revision, the original Fig. S4 was moved from the Supplementary Materials to the main text and is now presented as Fig. 5. Due to an oversight, the corresponding Fig. S4 reference was not updated in the Supplementary Materials and manuscript text. We have now corrected the figure numbering and all related cross-references.

7. **The POC:MAOC ratio deserves comment. In most mineral soils MAOC dominates SOC; here POC is almost 4x larger than MAOC. This is plausible for these albic forest soils but is unusual enough relative to the global literature that a sentence explaining why (parent material, low clay content, litter inputs) would help readers judge generalizability.**

Response: We thank the reviewer for this insightful comment, and for recognizing the pedogenic hypothesis of the POC-dominated carbon composition in our forest soils. As suggested, we have expanded the discussion in Section 4.1 by explicitly addressing the combined effects of organic matter inputs, soil texture, and albic pedogenesis:

'The relative difference in prediction accuracy between POC and MAOC may partly reflect the differences in organic matter inputs and decomposition, as well as pedogenic and mineralogical characteristics of the forest soils investigated here. Unlike the low-altitude agricultural systems examined in previous studies (e.g., wheat–rice rotations on Alisols or Fluvisols derived from lacustrine and marine deposits, where MAOC dominated the SOC pool at an average of 76% [14.6

g kg⁻¹ MAOC vs. 5.5 g kg⁻¹ POC out of 20.0 g kg⁻¹ SOC]; Dai et al., 2024), our soils developed under a forest environment characterized by albic/podzolization-related features (Albic Luvisols or Podzols). In our forest soils, POC dominated the SOC pool, accounting for more than 80% of total measured SOC (mean POC of 19.04 g kg⁻¹ vs. MAOC of 4.56 g kg⁻¹; Fig. 3). The soils in the present study were dominated by silt-rich textures, comprising approximately 74% silt, 13% sand, and 14% clay on average across the profile (Table S1). In such albic forest soils, prolonged eluviation associated with podzolization severely reduces the abundance of fine clay and reactive Fe/Al mineral complexes in upper horizons, directly attenuating the diagnostic mineral absorption signatures (Fe–O charge transfer < 550 nm and Al–OH lattice vibrations at 2200 nm) that normally drive MAOC quantification (Hartemink et al., 2020; Lundström et al., 2000). At the same time, heavy inputs of forest litter exceed the sluggish microbial activity constrained by low temperatures (Tian et al., 2016), favoring sustain a large POC pool (Lavallee et al., 2020). Together, these pedogenic conditions alter the balance between particulate and mineral-associated carbon by modulating both organic matter input regimes and the availability of reactive mineral surfaces for organo-mineral association (Kleber et al., 2015). Therefore, the POC-dominated SOC composition and its superior predictive performance ($R^2 = 0.96$) reflect the combined influence of forest organic inputs and quartz-rich albic pedogenesis, aligning cleanly with the global synthesis of Ding et al. (2026), which identified soil texture, mineralogical context, and concentration dynamic ranges as dominant environmental drivers governing SOC-fraction prediction accuracy, particularly for MAOC.'

By adding this explanation, we clarify that the observed POC dominance is a distinctive feature of quartz-rich, silt-dominated albic forest ecosystems, thus helping readers contextualize model generalizability across contrasting land-use and pedogenic settings.

8. **Fig. 5: It can be helpful to remind the preprocessing description in the figure description.**

   Response: Thank you, done. We have expanded the caption of Figure 5 to define all preprocessing acronyms.

9. **Fig. 6: It has REF acronym, while it is not one of the optimal preprocessing methods in the figure.**

   Response: Thanks for your comment. Technically, REF is the raw reflectance, not an actual preprocessing method. We have updated the description of Figure 6 to harmonize preprocessing notation.

10. **Fig. 6 "optimal R^2 obtained under laboratory conditions", since laboratory R^2 spans a wide range across algorithms/preprocessing (POC: 0.85–0.96; MAOC: 0.30–0.74), it isn't clear which single value the dashed line represents. Please specify If AB_SNV is intended for all**

three models, as in Fig. 5, SNV looks better in PLSR and MBL for MAOC.

Response: Thank you for this insightful observation. The horizontal dashed lines in Figure 6 do not represent a single fixed preprocessing method applied universally across all algorithms; rather, they denote the algorithm-specific optimal laboratory benchmark achieved by each model under its respective best-performing preprocessing strategy. We also add the results table in the supplement as Table S2.

Revised Figure 6 Caption: *'The red horizontal dashed lines represent the optimal laboratory R^2 benchmark for each algorithm under its model-specific best-performing preprocessing scheme (Table S2).'*

As the reviewer correctly noted, AB_SNV is not universally optimal for all three algorithms. For MAOC prediction under laboratory conditions (as shown in Figure 5):

- **PLSR**: Achieved its optimal laboratory performance with **SNV** preprocessing ($R^2 = 0.65$).
- **Cubist**: Achieved its optimal laboratory performance with **AB_SNV** preprocessing ($R^2 = 0.74$).
- **MBL**: Achieved its optimal laboratory performance with **SNV** preprocessing ($R^2 = 0.69$).

11. **L351: "under SNV + PLSR, R^2 for POC decreased compared with Workflow A." this is not shown in the figure 6, supplementary data or a table. Please present the mentioned result somewhere.**

    Response: Thank you for pointing this out. To address this issue, we have added the complete model performance results for the relevant workflows and preprocessing–algorithm combinations to Supplementary Table S2. The corresponding statement in the main text has also been revised to ensure that the reported comparison can be directly verified from the supplementary results.

12. **Fig. 8b/8d labeling. Panel (b) is labeled "REF_AB + Cubist" — this preprocessing tag doesn't correspond to any of the six preprocessing options defined before. Is this a typo for "AB"? Also, please use a bigger font to improve readability of numbers in plots. It's better if the label describes each panel rather than left and right columns.**

    Response: Corrected and updated. In Figure 8, we have: (1) corrected the tag typo "REF_AB" in panel (b) to "AB"; (2) increased the font size of all numbers and axis labels for improved legibility; and (3) updated the figure caption to replace generic column descriptions with explicit, non-redundant definitions for each subfigure panel (a)–(d) specifying the target carbon fraction and modeling strategy, as follow: '***Figure 8.*** *Comparison of environmental transferability between optimal laboratory and in-situ models across external validation datasets: (a) optimal laboratory-based model for POC; (b) optimal in-situ modeling strategy for POC; (c) optimal laboratory-based* *model for MAOC; and (d) optimal in-situ modeling strategy for MAOC. Marginal density distributions of the test (2025) and external (2024) datasets are shown along the top and right axes.'* *Thank* you for your comments.

13. **The weaker spectral signal of MAOC (mediated by iron oxides/mineral surfaces) may be related to albic characteristic of these soil. I suggest the authors discuss this pedogenic connection explicitly.**

    Response: We sincerely thank the reviewer for this constructive suggestion. We completely agree that the weaker spectral signal of MAOC is directly tied to the albic pedogenic characteristics of these forest soils. In response to your suggestion, we have explicitly incorporated this pedogenic connection into Section 4.1 and Section 4.3 of the revised manuscript.

    First, in Section 4.1, we updated the pedogenic mechanism explanation as follows:

    'In such albic forest soils, prolonged eluviation associated with podzolization severely reduces the abundance of fine clay and reactive Fe/Al mineral complexes in upper horizons, directly attenuating the diagnostic mineral absorption signatures (Fe–O charge transfer < 550 nm and Al–OH lattice vibrations at 2200 nm) that normally drive MAOC quantification (Hartemink et al., 2020; Lundström et al., 2000).'

    Second, in Section 4.3, we expanded the mechanistic interpretation of why mineral depletion restricts Vis–NIR spectroscopic sensitivity for MAOC:

    'In contrast, MAOC showed stronger contributions in the visible region, where spectral variability is largely influenced by mineralogical features (Fig. 7a). This is consistent with the nature of MAOC as organic matter associated with mineral surfaces rather than as a spectrally distinct organic phase (Kleber et al., 2015). Consequently, its spectral representation is mediated through mineral-related signals and therefore influenced by broader soil mineralogical characteristics (Schwertmann and Taylor, 1989; Chen et al., 2014; Kleber et al., 2015). This indirect and context-dependent spectral representation may help explain the comparatively lower prediction accuracy of MAOC observed in this study.'

Technical corrections

1. **L22 in the abstract, please define EPO at its first occurrence.**

   Response: Added the full term '*External Parameter Orthogonalization (EPO)*' in the abstract.

2. **L24: There is a phrase that seems to be incomplete or separated from the main sentence: "In contrast, direct in-situ modeling", it seems it could be "in contrast to …..".**

   Response: We thank the reviewer for pointing out this incomplete comparative phrase. The revised text in the Abstract now reads:

   'In contrast to laboratory-to-in-situ model transfer, direct in-situ modeling achieved higher predictive accuracy, with POC performing best using EPO-corrected field spectra with AB + Cubist (R^2 = 0.90), whereas MAOC performed best using raw in-situ spectra with AB_SNV + MBL (R^2 = 0.71).'

3. **L122-123 uses "POM"/"MAOM" (particulate/mineral-associated organic matter) while the rest of the manuscript uses POC/MAOC (carbon); a one-line clarification that carbon content was measured on the POM/MAOM fractions would avoid ambiguity.**

   Response: Thank you for pointing this out. We have added a sentence in Section 2.2 clarifying that the organic carbon contents measured in the POM and MAOM fractions were reported as POC and MAOC, respectively, to ensure consistent terminology throughout the manuscript. The specific text is as follows:

   'The carbon content for each isolated fraction was determined using an elemental analyzer (Vario MACRO cube), and the carbon contents of the POM and MAOM fractions were reported as POC and MAOC, respectively.'

4. **L194-196: Several ML algorithm abbreviations are not spelled, SVMR, GPR, ANN, RNN, CNN. RF and GBDT appear in Fig. 5 legend/axis labels, but the text only writes "Random Forest" and "Gradient Boosting Decision Trees" in full.**

   Response: We appreciate your careful review. We have defined all machine learning algorithm acronyms in full upon first mention in the Fig 5 caption.

5. **L213: the (c) seemed to be a leftover symbol from a mathematical description of EPO. Remove or explain please.**

   Response: The parameter *c* refers to the number of principal components retained in the EPO procedure and was not a leftover mathematical symbol. We have clarified its meaning and selection criterion in Section 2.4.4 as follow. Specifically, *c* was set to 1 because the first principal component of the difference spectra explained more than 95% of the total variance, and therefore one component was sufficient to characterize the dominant spectral variation between laboratory and in-situ measurements.

'The number of principal components retained (c = 1) was determined as the minimum number required to explain 95% of the variance in the difference. PCA spectra. This threshold was selected as a conservative compromise to remove dominant measurement-related variability while minimizing the risk of eliminating soil-intrinsic spectral information. The derived projection matrix was consistently applied in all subsequent analyses.'

6. **L251: please define the CV at first mention, for consistency.**
   Response: Thank you for this helpful suggestion. We agree that the abbreviation should be defined at its first occurrence for consistency. We have revised the terminology for consistency throughout the manuscript. Cross-validation (CV) is now defined at its first occurrence in Section **2.4.3**, and the abbreviation CV is used consistently thereafter. The corresponding occurrences in Sections **2.4.7** and **2.5** have also been revised accordingly.

7. **Eq. (2) RMSE and Eq. (3) Bias use opposite sign conventions for the residual ($\hat{y}_i - y_i$ vs. $y_i - \hat{y}_i$); harmless for RMSE (squared) but worth aligning for clarity.**
   Response: Thanks for this advice; this is helpful to make the equations clearer. We have aligned the residual sign conventions across Eq. (2) and Eq. (3) to $y_i - \hat{y}$ for consistency.

8. **L263: Citation formatting: "R.A. Viscarra Rossel et al., 2006" needs a spacing.**
   Response: Corrected spacing in the citation.

9. **L269: Space after "."**
   Response: Added the missing space after the period.

10. **L297–299: "training and test sets showed comparable ranges" — the POC training max (134.88) is roughly double the test max (67.89); consider softening "comparable" or reporting IQR instead of full range for this comparison.**
    Response: We thank the reviewer for this sharp observation. As suggested, we have softened the wording in Section 3.1 to remove "comparable" and explicitly state that the training set covered a broader concentration range with a higher maximum value than the test set. We also updated the description to provide full transparency regarding concentration ranges.
    The revised text in Section 3.1 now reads:
    'The mean contents of POC and MAOC were 19.04 and 4.56 g kg^{-1}, respectively (Fig. 3). Both POC and MAOC exhibited positively skewed distributions, with skewness values of 1.87 and 1.70, respectively, indicating the presence of high-value observations. The training and test sets covered broad ranges of POC and MAOC concentrations, with the training set showing a broader POC range than the test set (Fig. 3). For POC, values ranged from 1.28 to 134.88 g kg^{-1} in the training set (IQR = 4.27–26.78 g kg^{-1}) and from 1.76 to 67.89 g kg^{-1} in the test set. For MAOC, the

corresponding ranges were 0.48–17.62 g kg^{-1} and 1.05–11.47 g kg^{-1}, respectively (Fig. 3a). The soils were generally silt-dominated in texture, comprising approximately 13% sand, 74% silt, and 14% clay on average (Table S1). POC exhibited greater variability than MAOC across samples and was the dominant SOC fraction, with a mean concentration approximately four times higher than that of MAOC (POC: MAOC ratio ≈ 4:1). Furthermore, both POC and MAOC concentrations displayed pronounced downward gradients across the four soil depth intervals (0–10, 10–20, 20–30, and 30–40 cm), with mean concentrations declining from 52.64 to 4.49 g kg^{-1} for POC and from 7.71 to 2.44 g kg^{-1} for MAOC, respectively, accompanied by a gradual increase in the relative contribution of MAOC to total measured SOC (Table S1).'

11. **L312: Principal component analysis to PCA.**
    Response: Thank you for the reviewer's careful and rigorous examination. We have defined principal component analysis (PCA) at its first occurrence in the manuscript and subsequently used the abbreviation PCA consistently throughout the manuscript.

12. **L317-318: In "while the surface layer exhibited lower reflectance and smoother spectral concavity.", lower reflectance is redundant as it is stated in the previous sentence.**
    Response: Thank you for pointing this out. We removed the repeated description and retained the distinct observation regarding the smoother spectral concavity. The revised text now reads:
    'Under laboratory conditions, subsurface layers (10–40 cm) generally showed higher reflectance than the surface layer (0–10 cm), while the surface layer exhibited a smoother spectral concavity.'

13. **L350: "R^2 for POC decreased compared with Workflow A", consider rewording for clarity.**
    Response: We have revised the statement for clarity by specifying the preprocessing and modeling combination and explicitly indicating Workflow A as the reference. The revised text reads: *'For example, under SNV + PLSR, POC prediction performance declined relative to Workflow A.'*

14. **L355-L362: Consider rephrasing this part to improve readability and make it easier to follow.**
    Response: Thank you for this helpful suggestion. We agree that the original presentation combined the comparison between Workflows C and D with the selection of the optimal in-situ modeling strategies, which could make the results difficult to follow. We have therefore reorganized and rephrased this section to first clarify the performance differences between Workflows C and D, followed by the target-specific selection of the optimal modeling strategies for POC and MAOC. We also clarified that the reported R^2 ranges refer specifically to the Cubist model under Workflow D. The revised text reads:
    'Workflow D achieved performance comparable to or slightly better than Workflow C, with improvements mainly observed for the Cubist model. For Cubist, R^2 ranged from 0.74 to 0.90 for POC and from 0.53 to 0.66 for MAOC under Workflow D. In contrast, EPO correction produced

only minor reductions or negligible changes in predictive accuracy for PLSR and MBL relative to Workflow C.

Overall, the comparison among the in-situ workflows indicated target-specific differences in the benefit of EPO correction. Workflow D provided the most robust performance for POC, with Cubist combined with AB preprocessing achieving the best balance of predictive accuracy and robustness. In contrast, MAOC achieved its best performance under Workflow C, using MBL with AB-SNV preprocessing, indicating that EPO correction did not provide additional predictive benefit for this fraction. Based on these comparisons, the optimal in-situ modeling strategy for each target variable was selected and subsequently evaluated using the independent 2024 dataset.'

15. **Fig 7. Please improve the quality of the figure.**

    Response: The quality of Fig 7 has improved.

16. **L556–558: "Providing new methodological insights..." is a sentence fragment; should read "This provides new methodological insights..." or similar.**

    Response: Corrected the sentence fragment as: '*This study provides new methodological insights into…'*

17. **L557: Hyphenation of "in situ" is inconsistent.**

    Response: Thank you for the advice on hyphenation; we have revised the 'in situ' to 'in-situ'. We also unified the hyphenation style of 'in situ' / 'in-situ' throughout the manuscript.

18. **In references: Kaiser & Guggenberger 2003 and Viscarra Rossel et al. 2016 appear in the reference list but not in-text.**

    Response: Thank you for your careful and rigorous examination. We acknowledge this oversight. Following the final revisions, we carefully rechecked the entire reference list and in-text citations to ensure that all references are appropriately cited and that there are no formatting or citation inconsistencies. The manuscript has been corrected accordingly. We will also pay closer attention to such details in our future work to ensure greater accuracy and consistency throughout the manuscript.