Dear Raphael

Many thanks for the comments. Some good suggestions that we think have improved the paper. We have also made a few minor adjustments of our own to improve the clarity of the paper.

Out responses to your comments are below.

With best wishes

John Quinton on behalf of the authors.

Dear Authors,

Thank you for submitting your manuscript to SOIL. I am pleased to inform you that your paper is accepted pending minor technical corrections. The study addresses an important gap in the literature on grassland soil degradation in sub-Saharan Africa, and the combination of multi-year remote sensing (RS) classification with extensive soil sampling across two ecologically distinct districts is a genuine strength. The honest reporting of limited RS–soil agreement adds to the findings.

Please address the following points before the manuscript proceeds to production.

1. Temporal mismatch and transient variables. The multi-year RS classification (2013–2018) is conceptually well-suited to validation against stable soil variables and this is defensible. However, for transient variables (microbial biomass C, enzyme activities, mineralisation rates), the one-year offset between the RS period and soil sampling (2019) weakens their value as primary validation metrics. Please add an explicit discussion of this limitation, clarifying that transient variables reflect the biogeochemical state at the time of sampling and that their alignment/non-alignment with RS classes carries inherent uncertainty. This does not require reanalysis, but the current text does not acknowledge this issue sufficiently.

We agree that transient soil variables did not show any agreement with the remotely sensed data and have now made this explicit. However, we would argue that their inclusion is important since they are often seen as indicators of soil health (see multiple papers on this topic). We think it is a valid hypothesis to expect to see greater enzyme activity at those sites which have historically higher biomass production as indicated by the remotely sensed data. Although, enzymes are expected to react rapidly to changes in environmental conditions, for the reasons we set out in the paper, work by (Cordero et al., 2023) does demonstrate that environmental effects of drought can be persistent with soil enzyme activity continuing to be depressed six months after the drought has been alleviated. Given that the differences are taking place over a five-year period we do not think that sampling a year after the remote sensing period is a problem (see comments to this effect is earlier responses to reviewers).

2. Stable vs. transietn variables in the validation framework. The stable/transient distinction is conceptually valuable but is not used consistently to frame the validation discussion. Please revise the Discussion to foreground stable variables (total C, N, P, pH, bulk density, texture, aggregate stability) as the primary validators of the long-term RS classification, and to present transient variables as complementary, process-oriented indicators of current biogeochemical activity rather than primary validators of degradation state. This reframing strengthens the contribution by making explicit that productive sites are associated with both higher stable nutrient stocks and higher short-term biological activity.

There are essentially two parts of the paper. The first compares the values of the soil variables measured against the remote sensing classification. We discuss the best relationships first and the weaker ones after this. We think this is the logical way to proceed. The second component was to examine whether stable or transient soil variables could be related to degradation status (now made explicit in the sub-section header, note also that we have added these headings to the discussion). Here we discuss the relationship of the stable variables first and the transient second, as the editor suggests.

You will note, in response to the other points raised , that we have made some important changes to the discussion and conclusions (see below) . In addition, we have moved the critique of remotely sensed soil degradation classification towards the end of the discussion where it helps to draw the discussion to a conclusion.

3. Microbial biomass C. The argument linking microbial biomass C to aboveground biomass dynamics via organic matter inputs is reasonable. However, because microbial C is classified as a transient variable, a single sampling occasion cannot rule out the possibility that measured values partly reflect recent short-term conditions (e.g. rainfall events, grazing pulses) rather than chronic degradation. Please revise the relevant sections of the Discussion and Conclusion to state that microbial C is a useful short-term integrator of vegetation and organic inputs in these systems, while being explicit that it should not be interpreted as a stable long-term degradation proxy.
While we have some sympathy with the EE's position here, we respectfully disagree that microbial biomass C only reflects short term conditions. While some variability in the microbial biomass C may be due to short-term environmental changes, microbial biomass C is known to relate strongly to soil organic carbon content and water availability (Patoine et al., 2022). Thus, we would argue that microbial biomass carbon is a useful integrator of both long term inputs of carbon, reflected

Commented [QJ1]: @Bardgett, Richard Final comments from the Editor on the ReDeal paper. Could you check that I have it right concerning the soil biological parameters please. Still checking some statistical stuff with Mengyi on some of the other points. I will send you a link to the main documnet.

in the soil organic matter and shorter term inputs and may sit on the cusp of transient and . We have now edited lines 545-554 to try to stress this.

> Across the two studied districts (i.e., Nyando and Kuresoi) we observed consistent alignment between remote sensing classification of degradation and microbial biomass C. Microbial biomass C is a key soil biological parameter related to nutrient and C cycling processes in soil (Tate, 2017) that is tightly linked to soil carbon and water availability (Patoine et al., 2022), as well as plant diversity and productivity (Chen et al., 2019; Winterfeldt et al., 2026). It is also known to respond quickly (c.100 days) to inputs of fresh organic matter to soil, including plant litter and animal wastes (Dai et al., 2021). The strength of this relationship across both districts suggests that microbial biomass C is a useful integrator of both long term carbon cycling, reflected in the soil organic carbon content, and shorter-term turnover of carbon inputs, such as litter, root exudates, and dung from grazing animals.

4. Statistical power and interpretations of non-significant results. The productive class at Kuresoi (n ≈ 10–11) and the degraded class at Nyando (n ≈ 10–11) are modest in size, and within-class variability is large throughout, with coefficients of variation frequently exceeding 50–80%. Non-significant results for many soil variables should therefore not be interpreted as evidence of no relationship with RS class. Please add an explicit acknowledgement in the limitations section that low statistical power may contribute to the non-significance observed, and that this constrains the strength of conclusions regarding RS–soil disagreement.

We have added the following discussion to the end of the subsection "Relation of remote sensing classification to measured soil parameters" on page 13. We also added related comments in the Discussion section.

> It should be noted that the productive class at Kuresoi and the degraded class at Nyando both have relatively small sample sizes (around 10 to 11 observations), which, given the relatively high coefficients of variation (i.e., the ratio between the standard deviation the mean of a variable), results in a low statistical power in the class mean difference test. The small sample size is expected to contribute to the insignificant differences between classes in some variables. We can reason that, had we had a larger sample size , that more variables may have shown a significant difference between the degradation classes. Therefore, the disagreement between the degradation classification based on the remote-sensing data and the soil sampling data does not necessarily negate the relationship between the two. Instead, it suggests that more field sampling data may be needed to reach a more conclusive result.

Commented [QJ2]: @Gong, Mengyi I have deleted to the bit about Anova. I don't think it is necessary

In addition, we have added a line to the conclusions:

> However, we cannot discount that the low statistical power may have contributed to the non-significant agreement with the remotely sensed data.

5. RS classification thresholds. The two models were selected from five candidates on the basis of visual consistency of classification outputs. Please report the spatial agreement between the two selected models (e.g. percentage of pixels classified identically) as a basic measure of classification robustness, and briefly describe why the other three models were rejected. This should not require additional analysis beyond what has already been performed.

We have now improved the clarity of this section. Three models were rejected because it was obvious that they could not decern forest from pasture. This is now written as :

> Five models were tested and three were rejected as they could not differentiate between pasture and forest cover leaving two models.

We have added the following to the end of the paragraph on degradation unit classification detailing the relationship between the two models:

> For the analysis of Kuresoi, we estimated that 49% of all pixels were identically classified in model 1 and 2. This represented 6,330,389 pixels (or For Kuresoi, 49% of all pixels were identically classified in model 1 and 2. This represented 6,330,389 pixels (or 633 km^2) classified similarly as either productive, transient or degraded. In Nyando, 82% of pixels were identically classified in model 1 and 2. Representing 1,329,048 pixels or (133 km^2) of the landscape classified similarly as either productive, transient or degraded. Higher agreement in model classifications for Nyando might be due to a sharper contrast between degraded and live vegetation. Whereas in Kuresoi, the landscape is relatively heterogeneous and harder to separate transient pixels from degraded pixels

6. Transient variable clustering. The failure of transient-variable clustering to align with RS degradation classes is noted briefly but is not carried through adequately to the abstract or conclusion. This is an informative negative result that speaks directly to the limitations of transient variables for degradation assessment. This finding should be reflected in the conclusion (and abstract).

On rereading the discussion, the point that the transient variables did not perform well is missing. It has now been added and a sentence explaining it moved from the results to the discussion (where it should have been!).

> The transient variables aligned less well with the remote sensing classification presumably as the transient variables are highly variable and can change substantially in a short period of time.

In the abstract we have inserted:

> For the transient variables agreement between the clusters and the remote sensing classification was poor indicating a lack of utility for degradation assessment.

In the conclusion we have added:

> In the case of the transient variables, agreement between the cluster labels and the remote sensing classification was poor, indicating that the transient variables are less suitable for degradation assessment.

In addition, on rereading this session the repeated use of cluster 1 and 2 is confusing. Therefore, we have relabelled the clusters as Kuresoi stable and transient 1 and 2 (KS1, KS2, KT, KT2) and Nyando stable and transient 1 and 2 (NS1, NS2, NT, NT2). Hopefully this will improve the reader's experience.

7. Bulk density. Bulk density shows significant differences between degradation classes at both sites but with inconsistent rankings. The Discussion notes this briefly but does not explain it. Given the contrasting areas, a brief discussion of why bulk density may respond differently to degradation in each context would strengthen the paper.

We have added an explanation of why there may be differences in bulk density between the two regions, but given the small differences and the inconsistent patterns in each of the regions we do not think there is any benefit in postulating on potential reasons for the differences between degradation classes. We now state this clearly in the paper:

> Apart from microbial biomass C, only bulk density showed differences related to degradation class in both study areas. Bulk densities were higher in Nyando, probably due to the coarser textured soils, but the within region rankings for degradation classes were inconsistent and differences small. This makes drawing conclusions as to why we see these patterns difficult, but the small differences suggest that they are unlikely to make a functional difference.

Please note that none of the above points require new data collection or substantial reanalysis. They concern framing, interpretation, and presentation.

We thank you for choosing SOIL as the venue for your work and look forward to seeing your article published once these minor corrections are made.

Sincerely,

Raphael VISCARRA ROSSEL
EE, SOIL.

Chen, C., Chen, H. Y. H., Chen, X., and Huang, Z.: Meta-analysis shows positive effects of plant diversity on microbial biomass and respiration, Nature Communications, 10, 1332, 10.1038/s41467-019-09258-y, 2019.
Cordero, I., Leizeaga, A., Hicks, L. C., Rousk, J., and Bardgett, R. D.: High intensity perturbations induce an abrupt shift in soil microbial state, The ISME Journal, 17, 2190-2199, 10.1038/s41396-023-01512-y, 2023.
Dai, W., Peng, B., Liu, J., Wang, C., Wang, X., Jiang, P., and Bai, E.: Four years of litter input manipulation changes soil microbial characteristics in a temperate mixed forest, Biogeochemistry, 154, 371-383, 10.1007/s10533-021-00792-w, 2021.
Patoine, G., Eisenhauer, N., Cesarz, S., Phillips, H. R. P., Xu, X., Zhang, L., and Guerra, C. A.: Drivers and trends of global soil microbial carbon over two decades, Nature Communications, 13, 4195, 10.1038/s41467-022-31833-z, 2022.
Tate, K. R.: Microbial Biomass: A Paradigm Shift in Terrestrial Biogeochemistry, WORLD SCIENTIFIC (EUROPE), 348 pp., doi:10.1142/q0038, 2017.
Winterfeldt, S., Bardgett, R. D., Brangarí, A. C., Eisenhauer, N., Hicks, L. C., Liu, S., and Rousk, J.: Plant diversity increases microbial resistance to drought and soil carbon accumulation, Journal of Ecology, 114, e70250, https://doi.org/10.1111/1365-2745.70250, 2026.