Reply to Anonymous Referee #2 for "Estimation of the Degree of Decomposition of Peat and Past Net Primary Production from Mid-Infrared Spectra"

Henning Teickner[1,2,*] Julien Arsenault[3] Mariusz Gałka[4]
Klaus-Holger Knorr[1]

09 May, 2026

[1] ILÖK, Ecohydrology & Biogeochemistry Group, Institute of Landscape Ecology, University of Münster, 48149, Germany
[2] IfGI, Spatiotemporal Modelling Lab, Institute for Geoinformatics, University of Münster, 48149, Germany
[3] Département des Sciences biologiques, Université du Québec à Montréal, Montréal, H2X 1Y4, Canada
[4] University of Lodz, Faculty of Biology and Environmental Protection, Department of Biogeography, Paleoecology and Nature Conservation, Banacha 1/3, 90-237 Łodz, Poland

[*] Correspondence: Henning Teickner <henning.teickner@uni-muenster.de>

We thank the reviewer for the helpful comments and hope to have addressed them appropriately. Comments made by the reviewer start with a bold **Q** while our reply starts with a bold **A**. In section "Additional changes" we list additional changes we would like to incorporate in an updated version of the manuscript.

1 Reply to comments

1. **Q:** The training dataset appears to include only three species. However, *Phragmites australis* and *Typha latifolia* are not typically dominant species in northern peatlands. This is also consistent with the macrofossil data from the three peat cores analyzed, which are dominated by *Sphagnum* and *Eriophorum*. This raises concerns about the representativeness of the training data. Will the limit number of species affect the robustness of the prediction of the model?

   A: We thank the reviewer for the comment. Unfortunately, we do not have the data to test whether the selection of litter types affects the robustness of the predictions by the models. We agree that more tests are needed here, as mentioned in the Discussion

(section 4.2) and Conclusions (old: l. 610 to 615), and we agree that the variability in predictions of the three models indicates that differences in litter chemistry not covered by the models may lead to prediction errors, as discussed in our manuscript. We think that a more detailed discussion of this issue than provided in our manuscript is useful only with additional data and we welcome that others test our models, publish the results of these tests, and ideally also improved models.
Currently, we see our method as a promising first step. Our study contributes the general idea and a series of tests, attempts to shift the research focus from untestable decomposition indicators to measuring γ (any measurement is just an estimate), and illustrates how this idea could be useful to peatland research.
To mention more explicitly that also additional data from *Sphagnum* and graminoids would be useful to better test and improve the models, we suggest to change l. 510 to 513 (old) from:

"It would be useful to include material from more litter types (in particular woody plant organs) and samples with larger γ than is currently the case in order to reduce extrapolation errors of the models for decomposed peat samples."

to (new: l. 525 to 527):

"It would be useful to include material from more litter types (in particular woody plant organs but also other samples from already considered species) and samples with larger γ than is currently the case in order to reduce extrapolation errors of the models for decomposed peat samples."

2. **Q:** The reconstruction of NPP relies on a multi-step chain: γ–MIRS $\rightarrow \gamma \rightarrow$ initial mass $\rightarrow$ NPP, combining MIRS-predicted γ, MIRS-derived bulk density, and the age-depth model. This layered approach may introduce multiple sources of uncertainty. Why are MIRS-predicted bulk densities used instead of directly measured bulk densities, which could reduce uncertainty in the final estimates? The reliability of age-depth models is strongly dependent on the number and distribution of radiocarbon (14C) dates. I hope the manuscript can explicitly describe how uncertainties from the age–depth model are quantified and propagated into the reconstructed NPP.

   A: We use MIRS-predicted bulk densities instead of directly measured bulk densities not because we think that the MIRS predictions are more accurate — they are certainly not — but simply because we do not have access to peat cores that have all three, MIRS, macrofossil data, and bulk density measurements. It is clear that with bulk density measurements, prediction errors for the reconstructed NPP would become much smaller, as mentioned in (old) l. 596 to 597.
   We agree that this is not yet mentioned explicitly. To do so, we suggest to add at al. 276 (old; new: l. 284):

   "Clearly, prediction of bulk density and saturated hydraulic conductivity from MIRS introduces larger errors in our analysis than necessary. We used predictions because we do not have access to peat samples with estimates for peat ages, macrofossils, bulk density, and saturated hydraulic conductivity. We present these analyses as illustrations for the suggested approach but severe tests may require more accurate estimates

and measurements than are currently available. Errors can also be reduced further by measuring mass fractions of macrofossils and by measuring MIRS on individual macrofossils than on bulk samples."

Errors from the age-depth models are propagated into the reconstructed NPP and this is another large error source considered here that is usually neglected in other studies. Since we use a Bayesian age-depth modeling approach and our mixing model to estimate γ is also Bayesian, we could extract individual draws from the posterior distribution for the age and the γ estimates and propagate errors by calculating NPP for each individual draw. In the same way, errors from bulk density predictions are propagated into NPP estimates.
Error propagation is mentioned at l. 275 to 276 (old):

"Our analysis considers errors in estimates for γ_{MIRS}, bulk density, peat ages, and β_0."

To clarify this part, we suggest to modify it to (new: l. 280 to 283):

"Our analysis considers errors in estimates for γ_{MIRS}, bulk density, peat ages, and β_0. Errors from γ, bulk density predicted from MIRS and peat ages are propagated to NPP estimates by computing NPP with individual draws from the posterior distributions of the respective Bayesian models. Errors for γ_{MIRS} and β_0 are considered by using prior distributions within the mixing model with parameters chosen based on the prediction models for γ_{MIRS}."

3. **Q:** Figure 4 requires improvements to enhance readability and interpretability. A y-axis should be added for the TA-reconstructed WTD (blue lines) to clearly indicate wet versus dry conditions. The macrofossil panel also lacks a y-axis and should be properly labeled. The color scheme should be revised, as *Eriophorum vaginatum* and *Sphagnum medium* are difficult to distinguish, which obscures interpretation of phase relationships discussed in Section 3.7.

   A:

   - "A y-axis should be added for the TA-reconstructed WTD (blue lines) to clearly indicate wet versus dry conditions.": We thank the reviewer for pointing out this omission. We do not want to clutter the already quite full figure with an additional axis which would contain only few information due to the normalization of WTD. Instead, we suggest to modify the figure caption to mention that positive values indicate drier conditions.

   - "The macrofossil panel also lacks a y-axis and should be properly labeled.": This seems to be a misunderstanding. The plot elements colored by macrofossil taxon are not macrofossil abundances or counts but the NPP estimates for these taxa. Therefore, the y-axis that is already there also applies to all plot elements colored by macrofossil taxon.

   - "The color scheme should be revised, as *Eriophorum vaginatum* and *Sphagnum medium* are difficult to distinguish, which obscures interpretation of phase relationships discussed in Section 3.7.": We thank the reviewer for this suggestion.

We changed the color palette to better distinguish taxa and provide color blind friendly colors.

4. **Q:** Several issues in Figure 4 require clarification. A large proportion of results from models 2 and 3 appear to fall outside the prediction domain. Will this issue affect the reliability of prediction? Besides, in peat core MK1, reconstructed γ and NPP values are shown for the deepest section (prior to ~5000 BCE), while macrofossil data appear absent. The authors should clarify how reconstructions are supported in intervals lacking macrofossil constraints.

   A:

   - "A large proportion of results from models 2 and 3 appear to fall outside the prediction domain. Will this issue affect the reliability of prediction?": This is explained in section 2.4.2. If a MIRS is outside the prediction domain, this means that the model extrapolates for this sample which, in absence of additional information, makes predictions less reliable. However, without additional data and a more detailed analysis of extrapolation properties, it is impossible to say how unreliable these predictions are — this is why we currently advocate to consider this variability when applying our models and why we advocate to better test the models (old: l. 510 to 512).

   - "Besides, in peat core MK1, reconstructed γ and NPP values are shown for the deepest section (prior to ~5000 BCE), while macrofossil data appear absent. The authors should clarify how reconstructions are supported in intervals lacking macrofossil constraints.": Sections of the cores where macrofossil data appear absent do indeed have macrofossil data, but, as the figure caption mentions, we only show estimates for NPP for individual taxa "with a volume fraction of at least 40% in at least 15 samples (colored points are medians and error bars 90% confidence intervals)". Such data are usually absent in sections of the cores with very low apparent depth accumulation rates (thus the data are stretched along the time axis) or sections where macrofossil anlysis could not identify taxa clearly due to a large degree of decomposition.

5. **Q:** It seems like the both K_{sat} and γ is predicted by the MIRS, this constitutes a comparison between two modeled variables rather than independent observations. I recommended interpreting these results more cautiously and avoid drawing strong mechanistic conclusions without independent validation.

   A: As mentioned in our reply to comment 2 by reviewer 2, we think that our analysis has an illustrative purpose. We hope that our modifications to the text suggested there emphasize more why we use MIRS predictions, that we consider prediction errors and therefore avoid narrow interpretatios, and that we do not consider our analyses as severe tests of hypotheses. Unfortunately, the reviewer does not mention where our interpretation is no cautious enough or where we would draw strong mechanistic conclusions. We do not think that we did any of this, but are happy to revise our text in light of more concrete criticism.

To test many of the hypotheses we analyzed here, we need some approach to estimate γ of peat and it will therefore be difficult to get independent validation. However, we agree that other measurements, especially those where the MIRS prediction models have large errors, such as for K_{sat}, should be replaced by measurements to provide better tests of the hypotheses analyzed here.

2 Additional changes

1. At l. 525 (old; new: l. 540 to 544), we suggest to add:

   "While not considered here, an additional limitation of the current litterbag data may be that they do not consider how litter preprocessing may bias mass losses from decomposition compared to decomposition under more natural conditions (e.g., Bärlocher, 1997; Teickner et al., 2025). This may confound relations between MIRS and γ under more natural conditions when making predictions with models trained on litterbag data. However, at the moment, it is not clear how serious this problem is."

   Initially, we did not want to mention this issue because we do not address it in the study here and we did not want to complicate matters more than necessary. However, in hindsight, we think that it is useful to at least mention this possible limitation since it may also be relevant for studies that want to test or improve our models.

2. We replaced Teickner and Knorr (2025) by Teickner and Knorr (2026) which is the final published version of the same reference (a preprint).

3. We corrected typos in the names of some taxa.

References

Bärlocher, F.: Pitfalls of traditional techniques when studying decomposition of vascular plant remains in aquatic habitats, Limnetica, 13, 1–11, 1997.
Teickner, H. and Knorr, K.-H.: Prediction of peat properties from transmission mid-infrared spectra, https://doi.org/10.5194/egusphere-2025-4955, 2025.
Teickner, H. and Knorr, K.-H.: Prediction of peat properties from transmission mid-infrared spectra, SOIL, 12, 497–519, https://doi.org/10.5194/soil-12-497-2026, 2026.
Teickner, H., Pebesma, E., and Knorr, K.-H.: A synthesis of *Sphagnum* litterbag experiments: Initial leaching losses bias decomposition rate estimates, Biogeosciences, 22, 417–433, https://doi.org/10.5194/bg-22-417-2025, 2025.