Reply to Anonymous Referee #1 for “Underestimation of Anaerobic Decomposition Rates in Sphagnum Litterbag Experiments by the Holocene Peatland Model Depends on Initial Leaching Losses”

Henning Teickner[1,2,*] Edzer Pebesma[2] Klaus-Holger Knorr[1]

19 December, 2024

[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

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

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 manuscript entitled “Underestimation of Anaerobic Decomposition Rates in Sphagnum Litterbag Experiments by the Holocene Peatland Model Depends on Initial Leaching Losses” by Teickner et al tested decomposition rates from the Holocene Peat Model against litterbags decomposition experiments.

   The approach is interesting as it combines litterbag data and a Bayesian approach to improve the parametrisation of a model. Improving the use of litterbag data into peat decomposition models is timely and useful to further advance peat decomposition models. The manuscript is well structured. Unfortunately, major concerns regarding the scope, methods and some assumptions should be addressed before considering this manuscript for publication.

   A: We thank the reviewer for their comments and questions that are useful to clarify some points we make in the manuscript. We hope that our comments below and the suggested changes in the manuscript address the concerns.
   Please note that due to the extnsive reviewer comments, we made extensive changes

to the manuscript. Instead of listing all these changes out of context, we therefore attached a version of the manuscript that includes all suggested changes at the end of this document.

1.1 General comments:

2. **Q:** The motivation, overall approach and scope of the study are not clearly presented. Was the purpose of the study to identify an improved range of values for Wopt and c2?

   A: The aims of the study are twofold (compare with our manuscript, l. 77 to 82):

   1. We wanted to test whether the HPM decomposition module can predict decomposition rates estimated from available *Sphagnum* litterbag experiments using the standard parameter values reported in Tab. 1 and Tab. 2 in the original publication (Frolking et al., 2010).
   2. We wanted to estimate parameters of the HPM decomposition module from available litterbag data and analyze whether they differ from the standard parameter values defined in Frolking et al. (2010), what implication this has for the decomposition process, and whether they are negligible in terms of long-term peat accumulation as predicted by the HPM.

   Due to limitations of available litterbag experiments and limited information reported in these experiments discussed in our manuscript, it is not clear yet whether our estimates are an improvement and therefore we do not mention this as aim, even though it is of course a long-term goal.

   To better describe our aims, we suggest to change l. 79 to 82 to:

   1. "Test whether the HPM decomposition module can predict litterbag decomposition rates for different *Sphagnum* species along the gradient from oxic to anoxic conditions.
   2. Estimate HPM decomposition module parameters from litterbag data and compare them to the originally suggested values (standard parameter values) (Frolking et al., 2010) that are often used when applying the HPM (Tab. 1).
   3. If some of the parameter values differ, identify the possible causes why parameter estimates from litterbag data differ to provide guidance for future litterbag experiments.
   4. Analyze whether estimated differences in HPM parameter values could imply significant differences in decomposition rates and long-term peat accumulation."

   We are not sure whether the comment of the reviewer is due to a misunderstanding of our aims (i.e., different expectations of what is a good test of a model, please see also comments 8 and 13 by reviewer 1) or our failure to communicate our aims appropriately. We would be thankful if the reviewer could clarify whether he or she still thinks that the aims of the study are not appropriately communicated, so that we can try to further improve our manuscript if necessary.

3. **Q:** It is obvious that some interesting experiment has been done. Unfortunately, it is extremely difficult to judge from the current version of the paper if it is sound and whether is constitutes a significant contribution to the field.

   A: We thank the reviewer for the comment and apologize for the missing information in our manuscript. We suggest to make extensive changes to the methods section to describe in detail the formulas of the HPM decomposition module, of the litterbag decomposition model, and of the link between both.
   In particular:

   1. Reviewer 3 suggested to include a conceptual figure of our modeling approach and we will do so in the next version (Fig. 1 in the attached manuscript).
   2. We suggest to update the methods section to include the missing information. Please see our replies to comments 13, 14, 18 - 20, 22, 23, 30, 31 of reviewer 1.

4. **Q:** Also, it would be helpful to clarify what the difference is, in terms of contribution, between this paper and "A Synthesis of Sphagnum Litterbag Experiments: Initial Leaching Losses Bias Decomposition Rate Estimates" by Teickner et al., as the conclusions seem quite similar.

   A: The only similarity in the conclusions of both papers is that errors in initial leaching loss (l_0) estimates are an important control of bias and errors in decomposition rate (k_0) estimates. Even though our manuscript here builds on top of the *Sphagnum* decomposition model in Teickner et al. (2024a), both manuscripts have different scopes, address different questions and perform completely different analyses.
   Teickner et al. (2024a) tries to understand how initial leaching losses control bias and errors in decomposition rate estimates directly derived from *Sphagnum* litterbag experiments.
   In our manuscript here, we test to what extent the HPM decomposition module can predict decomposition rates in the same litterbag experiments from plant functional type identity, water table depth and degree of saturation. Moreover, we estimate parameters of the HPM decomposition module from the litterbag data and how they differ from the parameter values defined in the original publication of the HPM (Frolking et al., 2010). We also discuss what these discrepancies imply for the peat decomposition process modeled by the HPM and for long-term peat accumulation modeled by the HPM (please also see our reply to comment 2 of reivewer 1). We suggest to re-write large parts of the Introduction section to reduce dependency of this manuscript to our previous study and to better explain how both are connected.
   We apologize that we did not provide sufficient details on the HPM decomposition module here and how we estimated its parameter values from litterbag data. We think that this may have caused the confusion and we hope that our changes suggested in comments 13, 14, 18 - 20, 22, 23, 30, 31 of reviewer 1 address this issue.

5. **Q:** The experimental design and methods need to be well justified. The experimental design includes a series of assumptions that need to be justified in more detail, so that the reviewer can assess of the quality and soundness of the study. Please see specific comments for details.

   A: We thank the reviewer for pointing out several ambiguous sections of our manuscript. We hope that our replies to comments 14, 18 - 20, 22, 23, 30, 31 of reviewer 1 address this issue.

6. **Q:** Methods should include more detail to allow the reader to reproduce the experiment. Some areas of the methods are well covered. However, it remains unclear what has been done with the litterbag data and with the model. More detail on the models, including HPM, as well as the Bayesian tools used are needed.

   A: We hope that our replies to comments 14 and 22 of reviewer 1 address this issue. The reviewer does not mention what details on the Bayesian tools used are needed and we think that we provide detailed information on the Bayesian data analysis in section 2.4 (section 2.3.2 in the updated version of the manuscript) and the supporting information. We also emphasize that all code to reproduce our analyses are published in Teickner et al. (2024b). If the reviewer has specific comments on details missing about the Bayesian data analysis, we are thankful for further explanations.

7. **Q:** The experiment should be presented as a stand-alone in the paper. Information from Teickner et al. 2024 and its supplementary material was useful to get some sense of what has been done here but it should not be necessary to read another paper to understand this study. If the paper draws from the other paper's conclusions, then maybe summarize the previous conclusions and explain the link and contribution of this paper.

   A: We suggest extensive updates of the Introduction and Methods section, where we de-emphasize the role of our previous publication that describes the litterbag decomposition model, and describe in detail our modeling approach. We hope that our suggested changes to the manuscript that describe the modeling approach in more detail (please see our reply to comments 14 and 22 by reviewer 1) clarify how the litterbag decomposition model from our previous manuscript is included in our model here and provide sufficient information that allows to understand our current study without the need to read our previous manuscript.

8. **Q:** Assessing the validity of the decomposition functions of the model is interesting. It does not constitute a test of the validity of the whole model. There are therefore limitations around the conclusions that can be drawn from this experiment. The limitations of the study should be clearly presented.

   A: We agree with the reviewer that testing a module of a model is no test of the whole model. Our aim was not to test the HPM but the decomposition module of the HPM, as stated in the introduction. We agree that our wording was ambiguous in some places and we suggest to correct this (please see our reply to comment 9 of reviewer 1).
   However, we do not think that it is a limitation that our study only focuses on the HPM

decomposition module. Instead, we consider it a strength that our focus on only one aspect of the HPM allows us to derive explicit aspects that, when considered in future tests, will give novel insights. Please also see our reply to comment 13 of reviewer 1 where we provide additional detailed justification of our approach to test the HPM decomposition module. We suggest to include a separate subsection at the end of the discussion section (section 4.4 in the attached manuscript with the suggested changes) where we summarize limitations of our test and what experiments and data are needed to improve our approach to test decomposition modules of peatland models.

9. **Q:** Terminology: some technical terms are not consistently or not appropriately used throughout the paper, which brings confusion: e.g. model, prediction, equation, confidence level. Please check usage throughout the text. Also, some statements mention that the decomposition module was tested in this study and some statements indicate that this is not what was done here.

   A: We suggest to correct and clarify all instances where our statement may imply that we tested the whole HPM or a modified version of the HPM by explicitly stating that we tested the HPM decomposition module or a modified version of the HPM decomposition module.
   We suggest to correct one instance where "confidence interval" instead of "confidence level" may be used (see our reply to comment 42 of reviewer 1).
   We suggest that our usage of "model equation" questioned in comment 44 of reviewer 1 is not inappropriate (see our reply to comment 44 of reviewer 1).
   We are not aware of instances where we used the terms "model" or "prediction" inconsistently or inappropriately or of specific comments by reviewer 1 where such misuses are mentioned. In case the reviewer disagrees, we would be happy to see specific parts/sentences where this applies and would then include changes or clarification.

10. **Q:** Readability: In many places, the statements are difficult to follow, which makes the paper hard to read. For example, sentence L4-6 should be rephrased to improve clarity.

    A: We thank the reviewer for this suggestion. We suggest to simplify statements where they were mentioned in specific comments (please see our reply to comment 36 by reviewer 1). In addition, we extensively reviewed the text and identified parts of the text that can be simplified and suggest to do so in the updated version of the manuscript.

11. **Q:** Conclusions: The gradient from oxic to anoxic conditions in HPM is designed for long-term peat decomposition within a dynamic simulation. Is it therefore possible to draw conclusions on its suitability from this experiment?

    A: We agree that litterbag experiments cover shorter time periods than are typically modelled with the HPM. It is also correct that changes in litter chemistry over longer periods of decomposition change the decomposition rate and this effect is not covered by litterbag experiments. Apart from this, long-term decomposition is the outcome of the same decomposition process happening on short time scales after a long time. Thus, even though litterbag experiments do not cover all aspects of decomposition, a fact

highlighted both by Frolking et al. (2010) and in the introduction of our manuscript (ll. 41 to 43), there currently is no better option to directly test how accurate the decomposition process is represented in long-term peatland models other than litterbag experiments.
Due to the assumption that the decomposition process is essentially the same, the HPM was explicitly developed based on litterbag studies, and in particular how the effect of degree of saturation (oxygen availability) and depth below the water table are modeled and how parameter values are defined (Frolking et al., 2010). Moreover, the HPM has an annual time step and also models decomposition of fresh litter over time periods covered by litterbag experiments. We therefore think that litterbag experiments are suitable to test the HPM decomposition module and are not aware of other options that have fewer limitations.
Since we mention this limitation in the introduction, we do not suggest to include changes in the manuscript that address this comment, unless the reviewer thinks this point is still not sufficiently addressed.

1.2 Specific comments:

12. **Q:** L1: The title needs to be revised as the litterbag experiments were not conducted with the HPM per se. Also, it might be worth considering a title in line with the general scope of the study.

    A: We agree that the title is currently ambiguous as we focus on the decomposition module of the HPM. We also agree that the title can be improved to better summarize all important aspects of our study and shorten it. We therefore suggest to change the title of the manuscript to:

    "Peat Oxic and Anoxic Controls of *Sphagnum* Decomposition Rates in the Holocene Peatland Model Decomposition Module Estimated from Litterbag Data"

13. **Q:** L35: Could you please justify and explain this approach? Considering only a section of the model does not seem to be an adequate method to test the validity of a model.

    A: We agree that considering only a part of a model is not an adequate method to test the validity of a model and we did not want to imply that we tested the validity of the entire HPM.
    Previous tests of the HPM and other peatland models typically compare predictions for properties of peat cores after hundreds or thousand years of past or simulated peat accumulation (e.g. Frolking et al. (2010), Tuittila et al. (2013), Quillet et al. (2015), Zhao et al. (2022)). These models test predictions conditional on the complete HPM, but they are of limited use when it comes to detecting what process causes discrepancies and what concrete actions are necessary to improve the HPM because of parameter unidentifiability, many omitted processes, and the practical limitation that peat cores with millenia of peat accumulation cannot be produced in experiments. In philosophy of science, the problem that many possible alternative hypotheses (including sets of parameter values) can equally well explain (fit) a phenomenon is known as Duhem-Quine problem and has been widely discussed (see e.g., Mayo (1996)). A necessary

condition to improve peatland models is to identify what process causes predictions to differ from measured peat properties (e.g. accumulated peat mass) and tests of complete peatland models are not apt for this job because of the Duhem-Quine problem.
To overcome the Duhem-Quine problem, scientists conduct piece-meal tests of complex models or research problems (Mayo, 1996). Our study here can be seen as application of some part of this approach to the decomposition module of the HPM. Thus, when we split off the decomposition module from the rest of the HPM, we can test whether the HPM decomposition module correctly predicts decomposition rates for specified plant functional types (PFT), degrees of saturation, and water table depths (WTD), while we do not also have to worry whether the HPM describes litter production, peatland hydrology, root production, changes in bulk density with decomposition, etc. correctly or whether our weather data are correct for the past millenia. This test is possible because suitable experiments exist where we can control (experimentally or statistically) PFT, degree of saturation, and water table depth and from which we can estimate decomposition rates. When the HPM decomposition module fails to predict decomposition rates in litterbag experiments, this certainly indicates a discrepancy between the HPM and a specific aspect of the peat accumulation process and identifies a concrete opportunity to improve one specific aspect of the HPM, instead of just identifying a general problem of the HPM, where the error source is not clearly identified.
We tried to explain this rationale in the introduction while keeping the focus on the problem we try to address (see ll. 25 to 43). We also note that our study has many limitations and because of these our test applies only a fraction of the severe testing approach developed in Mayo (1996) (typically, a series of well designed experiments is necessary to test some subject matter hypothesis severely). However, our study is useful in identifying — at no additional experimental costs — what differences in parameter values may be expected, whether these would have important consequences, and what kind of new litterbag experiments and data are needed to improve our test and estimates.
We therefore think that tests of individual modules typically give more targeted information on what aspects of a model are inadequate and how to improve modules and tests of modules; tests of whole ecosystem models provide insights equivalent to these piece-meal tests only in rare occasions.
We currently think that lines 25 to 43 (old manuscript) justify our approach (also since the other reviewers did not make similar comments), but if the reviewer has specific comments on how our approach may not be well justified, we are interested to hear their suggestions.

14. **Q:** L73: Could you please give more detail for clarification: Are the decomposition rates predicted by the HPM obtained from HPM simulations?

    A: We thank the reviewer for this very helpful comment and apologize that the first version of the manuscript did not contain a more detailed description of our model. Although these information are also given in supporting information S1, we agree that it is necessary to include a more detailed description of the actual formulas from the HPM we consider as decomposition module and how we link *Sphagnum* litterbag experiments with this decomposition module (compare also with comment 22 of reviewer

1) to avoid any ambiguities.
We suggest a complete re-write of the Methods section to describe our modeling approach in detail, to describe how the HPM decomposition module is used as prior for the litterbag decompostion model, and what model versions we computed. These changes are too extensive to list each of them individually, therefore we attached the manuscript with the planned changes included. Please see this attachment for the suggested changes to the methods section.

15. **Q:** L81: What kind of information does that deliver? Identifying new estimates to be used in future HPM versions would seem a better aim?

    A: The long-term goal certainly is to improve the accuracy of parameter estimates to be used in future HPM versions. However, currently there is no information on how accurate the parameter values of the HPM decomposition module are when estimated from data and the first steps (our study provides) therefore are to estimate parameters from data, test whether and how they differ from previous values, and analyze what could cause differences in parameter values and whether there is sufficient evidence that these differences actually are an improvment (or rather require to improve the test).
    Our analysis indicates some differences, but we discuss that these differences could either be caused by deficiencies of current litterbag experiments or indeed represent improved parameter estimates, and our analysis suggests how to improve litterbag experiments to rule out the former. Thus, identifying new estimates to be used in future HPM versions would be a better aim, but first one has to develop a strategy that ensures that parameter estimates derived from litterbag experiments indeed are improvements. This is what our study does.
    We agree that our study does not merely compare parameter values, but also tries to identify possible causes and consequences for peat accumulation. We therefore suggest to modify the description of the second aim and include a third and fourth aim to amphasize this (see our reply to comment 2 of reviewer 1).

16. **Q:** L92: Could you be more specific? It seems that HPM's decomposition module was not tested here.

    A: We are not sure why the reviewer has the impression that the HPM decomposition module was not tested here, even though we mention this as explicit aims in previous parts of the introduction.
    We agree that in some places we ambiguously wrote something like "predicted by the HPM" or "HPM modification" throughout the text and we will change these instances to make clear that we refer to the HPM decomposition module.
    We also agree that we do not test the HPM decomposition module alone, but that we need auxiliary models to link it to the litterbag data and to estimate the degree of saturation. However, these limitations are clearly stated in our manuscript (following the reviewer comments, some of these are clarified), and in addition, we suggest to summarize these limitations in a separate subsection in the Discussion section (please see our reply to comment 8 of reviewer 1). Despite these auxiliary models, we consider

our study a test of the HPM decomposition module because it isolates this module from the HPM and tests it against data.

17. **Q:** L115: It is unclear how this can be done while excluding the interactions with the other modules of the model. Could you please give more details?

    A: We assume that the reviewer asks how "… to link decomposition rates estimated from litterbag data to the decomposition rates predicted by the HPM …" (l. 115) is possible (or useful?) while excluding the interactions with the other modules of the model. We hope that our reply to comment 14 of reviewer 1 addresses this question and that our modifications to the Methods section mentioned there provide the necessary changes to the manuscript.

18. **Q:** L139: Would it not be more appropriate to reject this study as it brings a lot of uncertainty in the results?

    A: The reviewer refers to (l. 139) "Litterbag data from Prevost et al. (1997) are incubations of peat samples where the species is unknown."
    It is true that this study can be be assigned to one of the three *Sphagnum* PFT in the HPM only ambiguously and this potentially introduces errors when testing the HPM. In particular, this mainly introduces errors when estimating $k_{0,i}$.
    However, we note that assigning litterbag experiments where the species was known to one of the three *Sphagnum* PFT in the HPM also is ambiguous (as described in section 2.2.2 of our manuscript) and that all samples required estimating $k_{0,i}$ to make the decomposition rates from the litterbag experiments compatible with predictions of the HPM decomposition module. In particular, our analyses do not indicate that data from this study introduce additional variation which would be different from other studies. At the same time, there are few experiments where WTD were reported and we would therefore lose useful information if these samples were dropped. We agree that there is a need to improve estimates for $k_{0,i}$ in general and this is discussed in section 4.6 of our manuscript (section 4.4 in the updated version).

19. **Q:** L146: Could you please explain why this is needed and how this differs from the approach taken in HPM?

    A: We think the reviewer refers to why the degree of saturation had to be estimated and why we did not use the equation for the degree of saturation used in the HPM (equations (16) to (18) in Frolking et al. (2010)), but the ModGBerg model (Granberg et al., 1999; Kettridge and Baird, 2007).
    We indeed do think that a large limitation of available litterbag data — and therefore also our study — is that measurements of the degree of saturation (ideally with high temporal resolution) are lacking. In line with our approach to test the decomposition module of the HPM in isolation of all other modules of the HPM, it makes sense to use an estimate for the average annual degree of saturation as accurate as possible because then discrepancies the test detects cannot be caused by an incorrect representation of the degree of saturation.
    As mentioned in the manuscript (ll. 202 to 206, old version), an improved test of the

HPM decomposition module would use direct measurements of the degree of saturation. Since these data are not available, it makes sense to use a model to estimate the degree of saturation that is accurate and uses only the information available from litterbag data (often, this is only WTD and depth of the sample below the peat surface). The ModGBerg model (Granberg et al., 1999; Kettridge and Baird, 2007) is derived from comparatively accurate laboratory measurements, whereas it is less clear how the equations used in the HPM are exactly derived. We are aware that more sophisticated models exist to model peat water content, but these require additional data not available from available litterbag studies.
Therefore, we do not consider as limitation *which* of the available approximate models for the degree of saturation applicable to available litterbag data to choose, but *that* accurate estimates/measurements for the degree of saturation are currently not available for litterbag experiments. Of course, this limits the severity of our test which is why we emphasize the need to do additional experiments that rule out these limitations in our manuscript (see for example the conclusion section) to replicate our results and to improve our test (ll. 202 to 206, old version).
To even more explicitly state this limitation, we suggest to add in l. 150 (old ersion): "An improved test of the HPM decomposition module would require litterbag experiments with direct measurements of the degree of saturation at sufficient temporal resolution."

20. **Q:** L158: It is unclear why 2 PFTs were assigned. Could you please give more details on your approach?

    A: We thank the reviewer for pointing out that we were not explicit enough here. The reason for assigning two dummy PFT to samples from Prevost et al. (1997) is that sample material of two different origins were used in this study and we can therefore assume that at most two different values for $k_{0,i}$ are necessary to predict decomposition rates with the HPM decomposition module.
    To improve our description, we suggest to change ll. 155 to 156 from

    "Prevost et al. (1997) incubated *Sphagnum* peat collected from different depths below the surface and these samples probably have already experienced some decomposition, however it is difficult to estimate how much"

    to

    "Prevost et al. (1997) incubated *Sphagnum* peat collected from two different depth levels from the same location and these samples probably had already experienced some decomposition, however it is difficult to estimate how much."

21. **Q:** L160: Could you please specify what is included or modified in the model versions? How are these model versions linked to the decomposition equations from Frolking et al. 2010?

    A: We thank the reviewer for pointing out that we omitted important information here. We hope that our reply to comment 14 and our changes to the manuscript suggested there address this issue.

22. **Q:** L169: Could you please explain how they are combined?

    A: We thank the reviewer for this helpful suggestion. We hope that our reply to comment comment 14 by reviewer 1 as well as the suggested changes to the manuscript address this question.

23. **Q:** L203: Could you please explain why uncertainties in water table depths can be considered negligible?

    A: The sentence starting in l. 203 is "Strictly, we do not test the decomposition module in the HPM, but the combination of the decomposition model in the HPM and the modified Granberg model, assuming that uncertainties in water table depths are negligible and that we accounted sufficiently for uncertainties in total porosity."
    This sentence does not state that uncertainties in water table depths can be considered negligible but that our analysis assumes this for simplicity because no sufficient information to consider this uncertainty are available. We also want to emphasize that we estimated a standard deviation for each annual WTD estimate and, where multiple WTD measurements were available in a study, we estimated standard deviations for the average annual WTD from these values to consider errors in average annual WTD estimates. However, often WTD values were aggregates or not reported or measured with sufficient temporal resolution and therefore all of these estimates are only approximate.
    To reduce any ambiguity, we suggest to change the sentence starting in l. 205 from

    "Litterbag experiments where the degree of saturation is measured would be needed to avoid this ambiguity."

    to

    "Litterbag experiments where water table depths and the degree of saturation are measured at sufficient temporal resolution are needed to avoid this ambiguity in future studies and to improve any test of the HPM decomposition module."

24. **Q:** L208: Model descriptions need more details. A separate section could be appropriate.

    A: We hope that our reply to comments 14 and 22, the additions to the manuscript suggested there, and the information provided in Tab. 2 (a new table that summarizes general information on the littebag studies per species) address this comment.

25. **Q:** L254: Table 3, inaccuracy: different versions of the decomposition module combined with other tools. Also, could you give details as to where details can be found on the 53 litterbag experiments?

    A: We suggest to change the first sentence of the caption of Tab. 3 from "Training and testing RMSE for decomposition rates as predicted by different versions of the Holocene Peatland Model ..." to "Training and testing RMSE for decomposition rates as predicted by different versions of the decomposition module of the Holocene Peatland Model ..." to avoid this ambiguity.
    Details on the 53 litterbag experiments are available from the original publications

explicitly mentioned and cited in l. 103 to 105 (old version) of our manuscript. Moreover, we suggest to include an additional table in the Methods section that lists basic information on the litterbag experiments available for the different *Sphagnum* species (see our reply to comment 24 of reviewer 1).

26. **Q:** Figure 1: 0.3 -0.4 values seem high. Could they be related to S. angustifolium or other species that are seen in a wide range of habitats and might not always be typical of hummocks?

    A: As described in section 2.2.2 of our manuscript, we agree that the assignment of individual *Sphagnum* species to *Sphagnum* PFT as defined in the HPM is ambiguous. Future studies should ideally define these PFT and how to assign specific samples to these PFT based on values for the parameter estimates that distinguish the PFT (i.e., $k_{0,i}$, but also all parameters that control the net primary production, see Frolking et al. (2010)).
    We are not sure why the decomposition rates estimated for many *S. angustifolium* are rather large and whether this is a generalizable pattern. However, we want to emphasize that these estimates also depend on the estimated magnitude of initial leaching losses and how one interprets α in equation (2) (updated version) (see also Teickner et al. (2024a)). We state that our test is uncertain here and that errors in $k_{0,i}$ estimates are rather large and we think more targeted experiments are required to address such specific questions (section 4.6 — sections 4.3 and 4.4 in the updated manuscript).

27. **Q:** L289: Figure 2, water table depth are more than 25 cm above litterbag in some cases. Could you please explain how it can be the case?

    A: The x-axis in Fig. 2 is the depth of the water table below the litterbag. Negative values occur whenever the litterbag was buried below the average annual WTD (as estimated from the information given in the studies). Thus, for some studies some litterbags were buried more than 25 cm below the average annual WTD. We suggest to include the following note in the caption to clarify this:

    "(negative values represent litterbags placed below the water table, positive values represent litterbags placed above the water table in the unsaturated zone)"

28. **Q:** L350: Was this done within this study? If so, could you please add details in the methods section?

    A: Yes, we did these computations within this study. The computations are predictions of decomposition rates with different versions of the modified HPM decomposition module under different conditions. We suggest to move the parts of this section appropriate for the Methods section to the Methods section, as suggested.

29. **Q:** L396: Cannot be estimated: do you mean in this study?

    A: The sentence ending in l. 396 is: "A difficulty is that available litterbag experiments cover only a comparatively small depth below the water table level (at most around 30 cm, Fig. 2) and therefore gradients in anaerobic decomposition rates across larger

depths below the water table currently cannot be estimated."
Yes, we wanted to say that because available litterbag data comprise litterbags buried at most around 30 cm below the average annual WTD, it is currently not possible to estimate along gradients in anaerobic decomposition rates covering larger depths below the water table.
To avoid this ambiguity, we suggest to change "… currently cannot be estimated." to "… currently cannot be estimated with available litterbag data."

30. **Q:** L407 and L443: please refer to more recent versions of HPM

    A: We thank the reviewer for pointing out that we should mention here recent extensions of the HPM.
    We suggest to add at then end of l. 426: "It is also worth mentioning that a modification of the HPM, HPM-Arctic (Treat et al., 2021), has a seasonally dynamic WTD and this modification may account for at least a part of the discrepancies we observed here. Unfortunately, most available litterbag data do not report WTD at sufficient temporal resolution to test whether standard HPM parameter values are more compatible with litterbag data when such seasonal variations in WTD are considered."
    One reason why we did not consider basing our analysis on HPM-Arctic is therefore that available litterbag experiments do not report WTD at sufficient temporal resolution to test whether standard HPM parameter values are more compatible with litterbag data when such changes in WTD are considered. Additional reasons are discussed in our reply to the following comment.

31. **Q:** L443: HPM was first published (Frolking et al. 2010) with an annual water balance calculation. This has been modified to a monthly water balance calculation a few years later, e.g. Treat's HPM-Arctic . Please consider looking at the latest available HPM code to ensure your conclusions are in line with the current state of development of the model.

    A: In addition to the reason mentioned in our reply to the previous comment, we here focus on the HPM version described in Frolking et al. (2010) for two reasons: First, the HPM version described in Frolking et al. (2010) is better analyzed than HPM-Arctic. In particular, the two sensitivity analyses from Quillet et al. (2013a) and Quillet et al. (2013b) allow us to derive consequences for peat accumulation implied by the dscrepancies in parameter values of the decomposition module we identified here, whereas, to our knowledge, no such sensitivity analyses exist for HPM-Arctic. Second, whereas HPM-Arctic makes crucial modifications to describe peat accumulation under permafrost conditions, we do not think that there is sufficient evidence indicating that this model version generally makes more correct predictions.
    We therefore think that practical limitations that make it impossible to sensibly test relevant innovations in HPM-Arctic and less information available on how exactly the modifications in HPM-Arctic affect the sensitivity of peat accumulation make a test of the HPM version described in Frolking et al. (2010) a useful contribution. That said, we suggest that tests of decomposition modules of other peatland models would also be useful to compare and improve these models, including HPM-Arctic.
    As mentioned in our reply to comment 8 of reviewer 1, we suggest to include a new

subsection in the Discussion section, where we list limitaitions of our test and suggest how it can be improved. We suggest to mention there that litterbag data with WTD and degree of saturation reported at sufficient temporal resolution are needed to analyze how seasonal variations in these variables would change estimates for the HPM decomposition module parameters.

32. **Q:** L428: For each site-specific simulation, c2 is adjusted to better represent site-specific conditions. Which value of c2 do you refer to when making this statement?

    A: We think the reviewer may refer to l. 420 and not l. 428, since l. 428 does not mention site-specific conditions. The part starting in l. 420 is: "What we have not considered due to limited data is that c_2 can be expected to depend on long-term changes in groundwater flow … or site-specific differences in hydrology and other factors … . Therefore, c_2 may differ between litterbag studies and our data only indicate that c_2 is larger on average, whereas more research is necessary to estimate and understand site-specific controls of c_2 and how a change in hydrology controls c_2."
    Here, we do not refer to any particular value for c_2. What we wanted to say is that c_2 is intended to describe a process that is assumed to be different between different sites because it is controlled by site-specific differences in hydrology, availability of terminal electron acceptors, etc. Therefore, it is not assumed that one value for c_2 is suitable for all sites, but rather that the value of c_2 differs between sites and this is one of the few parameters that is estimatedwith peat core data in studies applying the HPM (e.g. Quillet et al. (2015), Treat et al. (2021), Treat et al. (2022)).
    We are not sure whether this answers the question of the reviewer and would like to hear more detailed recommendations if this is not the case.

33. **Q:** L467: How does S12 support this statement?

    A: We think that the reviewer confused "supporting Fig. S12" (as stated in the manuscript) with supporting *section* S12 which is part of our supporting information, but not referenced here.

34. **Q:** L477: As mentioned above S. angustifolium might not be appropriately classified here. However, this brings to light that a more detailed classification could be useful to avoid misinterpretation of the HPM PFTs.

    A: We agree with the reviewer here and we discuss this issue in detail in section 4.6 (section 4.4 in the updated version of the manuscript). Since we estimate $k_{0,i}$ for each species separately in the other modifications of the HPM decomposition module and our discussion focuses on species-independent discrepancies implied by the litterbag data, this possible misclassification would not have an effect on our conclusions, considering the other limitations we identified, even if we changed our procedure to assign *Sphagnum* species to the HPM PFT.

35. **Q:** L494: Would some of the results not be beneficial to other users or help enhance wider knowledge?

    A: We are not sure what the reviewer would consider as "beenficial to other users" or as "enhance wider knowledge". Perhaps the reviewer expects concrete recommendations such as that future studies should definitely use the parameter estimates we suggest. As explained in our reply to comment 11 of reviewer 1, we do not think that such specific recommendations can be derived from our analysis and instead we suggest that the knowledge gaps for improved tests of the HPM decomposition module are an important contribution that is of practical relevance of users of the HPM and for improvements of the HPM, even if only in the long run.

1.3 Technical comments

35. **Q:** L3: missing blank space

    A: We thank the reviewer for pointing this out. We will correct this typo in the next version of the manuscript.

36. **Q:** L4-6: Long and complicated sentence, would benefit from being rephrased.

    A: The sentence is "Large uncertainties in available litterbag data allow predictions of the HPM to fit decomposition rates estimated from litterbags by adjusting initial leaching losses and decomposition rates estimated from the litterbag data within the range of their uncertainties."
    We agree that this sentence should be rephrased. Also based on comments by the othre reviewers, we suggest a complete re-write of the abstract. Please see the attached manuscript with the suggested changes for how the abstract section would look like.

37. **Q:** L174-182: Some sentences are unclear.

    A: We hope that our reply to comments 14 and 22 of reviewer 1 and the additions to the manuscript suggested there address this issue.

38. **Q:** L214: typo that

    A: We will remove the redundant "that" as suggested.

39. **Q:** L223: typo from

    A: We will correct "form" to "from" as suggested.

40. **Q:** L254: model behaviour of HPM: unclear, could you rephrase?

    A: Based on comments by the other reviewers, we suggest a complete re-write of the Results section and we will also change section titles here. Please see the attached manuscript with the suggested changes.

41. **Q:** L370: references seem to be in the wrong place

    A: We suggest to change

    "The less pronounced gradient in measured decomposition rates above the water table depth is, however, also visible for *S. fuscum* replicates within the same study (Johnson and Damman, 1991; Golovatskaya and Nikonova, 2017; Mäkilä et al., 2018) and in addition similar across these (independent) studies (supporting information S8), indicating that this pattern cannot be explained in all cases by differences between studies."

to

"The less pronounced gradient in measured decomposition rates above the water table depth is, however, also visible for *S. fuscum* replicates within the same study and in addition similar across these (independent) studies (Fig. 4, supporting information S8) (Johnson and Damman, 1991; Golovatskaya and Nikonova, 2017; Mäkilä et al., 2018), indicating that this pattern cannot be explained in all cases by differences between studies."

42. **Q:** L397: Figure7, confidence interval

    A: We will change the legend label as suggested.

43. **Q:** L504: The last sentence is unclear. Could you please rephrase it?

    A: The sentence is "Future litterbag experiments should improve the accuracy of initial leaching loss and decomposition rate estimates and then test whether the identified parameter discrepancies are reproducible and whether they can be described by known, but not yet fully quantified, controls of decomposition rates in dependency of water table fluctuations."
    We suggest to completely re-write the Conclusions section and hope that, by simplifying some sentences, all parts of this section are now understandable. Please see the attached manuscript with the suggested changes.

44. **Q:** S1: model equations: are they not rather sample distributions?

    A: We think that "model equations" comprises all mathmeatical formulas required to describe a mathematical model and this also comprises probability distributions for parameters. The model formulas contain both prior and sampling distributions (likelihood) and also mathematical equations and we therefore think that "model equations" is a good summary for the presented information.
    Please note that we suggest to remove supporting section S1 because we now will include all model formulas in the Methods section of the main text. Please see the attached manuscript with the suggested changes.

2 Additional changes

1. We suggest a complete re-write of large parts of the manuscript to address the reviewer comments. Specific aspects of this re-write are listed in the comments of the reviewers, others are too numerous for a list of them to be useful without knowing the context of these changes. Please see the attached manuscript with the suggested changes.

2. We suggest to include Quillet et al. (2015) as reference for studies estimating c_2 from peat cores. We suggest the following changes:

   We suggest to change l. 392 to 393 from

"Larger and smaller c_2 than the standard value have been estimated for several permafrost peatland cores with a modified version of the HPM with monthly time step (Treat et al., 2021, 2022)."

to

"Larger and smaller c_2 than the standard value have been estimated for several peatland cores with the HPM and a modified version with monthly time step (Quillet et al., 2015; Treat et al., 2021, 2022)."

3. Frolking et al. (2010) also mention that peat accumulation as predicted by the HPM is sensitive to c_2 and a site-specific parameter. We therefore add Frolking et al. (2010) as reference at l. 47 and 422.

4. In l. 77 we will correct "decmposition" to "decomposition".

References

Frolking, S., Roulet, N. T., Tuittila, E., Bubier, J. L., Quillet, A., Talbot, J., and Richard, P. J. H.: A new model of Holocene peatland net primary production, decomposition, water balance, and peat accumulation, Earth System Dynamics, 1, 1–21, https://doi.org/10.5194/esd-1-1-2010, 2010.

Granberg, G., Grip, H., Löfvenius, M. O., Sundh, I., Svensson, B. H., and Nilsson, M.: A simple model for simulation of water content, soil frost, and soil temperatures in boreal mixed mires, Water Resources Research, 35, 3771–3782, https://doi.org/10.1029/1999WR900216, 1999.

Kettridge, N. and Baird, A.: In situ measurements of the thermal properties of a northern peatland: Implications for peatland temperature models, Journal of Geophysical Research: Earth Surface, 112, https://doi.org/10.1029/2006JF000655, 2007.

Mayo, D. G.: Error and the growth of experimental knowledge, University of Chicago Press, Chicago, 1996.

Prevost, M., Belleau, P., and Plamondon, A. P.: Substrate conditions in a treed peatland: Responses to drainage, Écoscience, 4, 543–554, https://doi.org/10.1080/11956860.1997.11682434, 1997.

Quillet, A., Frolking, S., Garneau, M., Talbot, J., and Peng, C.: Assessing the role of parameter interactions in the sensitivity analysis of a model of peatland dynamics, Ecological Modelling, 248, 30–40, https://doi.org/10.1016/j.ecolmodel.2012.08.023, 2013a.

Quillet, A., Garneau, M., and Frolking, S.: Sobol' sensitivity analysis of the Holocene Peat Model: What drives carbon accumulation in peatlands?, Journal of Geophysical Research: Biogeosciences, 118, 203–214, https://doi.org/10.1029/2012JG002092, 2013b.

Quillet, A., Garneau, M., Van Bellen, S., Frolking, S., and Tuittila, E.-S.: Integration of palaeohydrological proxies into a peatland model: A new tool for palaeoecological studies, Ecohydrology, 8, 214–229, https://doi.org/10.1002/eco.1501, 2015.

Teickner, H., Pebesma, E., and Knorr, K.-H.: A synthesis of *Sphagnum* litterbag experiments: Initial leaching losses bias decomposition rate estimates, https://doi.org/10.5194/egusphere-2024-1686, 2024a.

Teickner, H., Pebesma, E., and Knorr, K.-H.: Compendium of R code and data for ”A Synthesis of *Sphagnum* Litterbag Experiments: Initial Leaching Losses Bias Decomposition Rate Estimates” and ”Underestimation of Anaerobic Decomposition Rates in *Sphagnum* Litterbag Experiments by the Holocene Peatland Model Depends on Initial Leaching Losses”, https://doi.org/10.5281/ZENODO.11472955, 2024b.

Treat, C. C., Jones, M. C., Alder, J., Sannel, A. B. K., Camill, P., and Frolking, S.: Predicted vulnerability of carbon in permafrost peatlands with future climate change and permafrost thaw in western Canada, Journal of Geophysical Research: Biogeosciences, 126, e2020JG005872, https://doi.org/10.1029/2020JG005872, 2021.

Treat, C. C., Jones, M. C., Alder, J., and Frolking, S.: Hydrologic controls on peat permafrost and carbon processes: New insights from past and future modeling, Frontiers in Environmental Science, 10, 892925, https://doi.org/10.3389/fenvs.2022.892925, 2022.

Tuittila, E.-S., Juutinen, S., Frolking, S., Väliranta, M., Laine, A. M., Miettinen, A., Seväkivi, M.-L., Quillet, A., and Merilä, P.: Wetland chronosequence as a model of peatland development: Vegetation succession, peat and carbon accumulation, The Holocene, 23, 25–35, https://doi.org/10.1177/0959683612450197, 2013.

Zhao, B., Zhuang, Q., Treat, C., and Frolking, S.: A model intercomparison analysis for controls on C accumulation in North American peatlands, Journal of Geophysical Research: Biogeosciences, https://doi.org/10.1029/2021JG006762, 2022.

Peat Oxic and Anoxic Controls of *Sphagnum* Decomposition Rates in the Holocene Peatland Model Decomposition Module Estimated from Litterbag Data

Henning Teickner[1, 2], Edzer Pebesma[2], and Klaus-Holger Knorr[1]

[1]Ecohydrology & Biogeochemistry Group, Institute of Landscape Ecology, University of Münster, 48149, Germany
[2]Spatiotemporal Modelling Lab, Institute for Geoinformatics, University of Münster, 48149, Germany

Correspondence: Henning Teickner (henning.teickner@uni-muenster.de)

Abstract. The Holocene Peatland Model (HPM) is a widely applied model to understand and predict long-term peat accumulation, but it is difficult to test due to its complexity, measurement errors, and lack of data. Instead of testing the complete model, tests of individual modules may avoid some of these problems. In particular, the HPM decomposition module can be tested with litterbag data, but no such test has been conducted yet.

Here, we estimate parameter values of the HPM decomposition module from available *Sphagnum* litterbag experiments included in the Peatland Decomosition Database and with a litterbag decomposition model that considers initial leaching losses. Using either these estimates or the standard parameter values, we test whether the HPM decomposition module fits decomposition rates (k_0) in *Sphagnum* litterbag experiments along a gradient from oxic to anoxic conditions.

Both Litterbag data and model versions where HPM decomposition module parameters were estimated suggest a less steep gradient of decomposition rates from oxic to anoxic conditions and larger anaerobic decomposition rates for several species than the standard parameter values. This discrepancy may be caused by ignoring effects of water table fluctuations on aerobic and anaerobic decomposition rates. Moreover, our analysis suggests that maximum possible decomposition rates of individual species ($k_{0,i}$) vary more than suggested by the standard parameter values of the HPM plant functional types. Based on previous sensitivity analyses of the HPM, the estimated differences to the standard parameter values can cause differences in predicted 5000 year C accumulation up to 100 kg m^{-2}.

The HPM decomposition module with standard parameter values fits k_0 estimated from *Sphagnum* litterbag data, but model versions where HPM decomposition module parameters were estimated and differ significantly have an equivalent fit. The reason why models with different parameter values have equivalent fit is that errors in remaining masses and the design of available litterbag experiments support a range of initial leaching loss and k_0 estimates. Consequently, applications of the HPM and any other peatland model should consider that a broad range of decomposition module parameter values is compatible with available litterbag experiments.

Improved litterbag experiments are needed for more accurate tests of any peatland decomposition module and for obtaining parameter estimates accurate enough to allow even only approximate predictions of long-term peat accumulation. The modeling approach used here can be combined with different data sources (for example measured degree of saturation) and decomposition modules. In light of the large differences in long-term peat accumulation suggested by the parameter estimates, we conclude

that it is worth to conduct such experiments, not only to improve the decomposition module of the HPM, but to improve peatland models in general.

1 Introduction

Decomposition is one of the major controls of how much carbon (C) peatlands can store. Compared to other ecosystems, northern peatlands usually have small decomposition rates because of cold temperatures, high water table levels, acidic pH value, and litter that does not decompose fast even under environmental conditions favorable for decomposition (van Breemen, 1995; Rydin et al., 2013). These slow decomposition rates caused northern peatlands to accumulate at least 400 Gt C (Yu, 2012; Nichols and Peteet, 2019) during the Holocene and changes in the controls of decomposition rates may cause them to loose considerable amounts of C to the atmosphere under climate and land use changes (Frolking et al., 2011; Loisel et al., 2017).

Peatland models are used to better understand past C accumulation and to predict future changes in peat C stocks, but because of the long time scales which have to be considered, they are difficult to test. Past studies have compared site-adapted simulations of peat height, age, C and N stocks, macrofossil composition, and water table level predicted by peatland models against peat core data (e.g., Frolking et al., 2010; Tuittila et al., 2013; Treat et al., 2021; Zhao et al., 2022), and have shown that existing peatland models can reproduce observed patterns to some extent. However, these tests suffer from two problems. First, they cannot reliably identify the parameter values or model equations that cause discrepancies between model predictions and measurements because they test entire peatland models against observed data. Second, there often are large uncertainties on both sides of the test; peatland models have large uncertainties in parameter values and model structure and these may produce a range of predictions as illustrated by uncertainty analyses (e.g., Quillet et al., 2013a, Quillet et al. (2013b)) and model intercomparisons (e.g., Zhao et al., 2022). Observed data also has uncertainty from measurements, peat dating, or simply missing data, for example for past precipitation. Large uncertainties can make tests inconclusive, no matter how much data we use. As a consequence, there remains large and often not quantified uncertainty about parameter values that control decomposition rates.

An alternative that avoids some of these problems is to test only some part of a model while taking into account relevant uncertainty sources. To estimate uncertainties in and test values of parameters that directly control decomposition rates, such a test could address the decomposition module of a peatland model. For example, in the Holocene Peatland Model (HPM) (Frolking et al., 2010), we only need to know litter species, peat degree of saturation, the depth of the litter below the peat surface, water table depth, and only five parameters to predict decomposition rates. The predictions can be compared to decomposition rates estimated from litterbag data and therefore future litterbag studies can directly test whether discrepancies identified in such a test are replicable. Admittedly, such a test is restricted to short time ranges and not representative for long-term decomposition rates which may differ from that of fresh litter (e.g., Frolking et al., 2001), but future tests with different scope and applications of the model will benefit from the reduced parameter uncertainties and can consider where the model fails already on short time scales.

A test of decomposition modules is relevant because of the importance of decomposition for long-term C accumulation in peatlands. Previous sensitivity analyses of the HPM and applications to peat cores suggest that the anoxia scale length (c_2), the parameter controlling how anaerobic decomposition rates are limited by electron acceptor depletion and accumulation of decomposition products, can result in a doubling of accumulated C, depending on climate conditions (Frolking et al., 2010; Quillet et al., 2013b; Kurnianto et al., 2015). These sensitivity analyses used assumed parameter ranges that are not informed by litterbag experiments. A test of only the HPM decomposition module can provide better estimates for c_2 and may therefore help to reduce uncertainties in predicted C accumulation rates.

Currently, litterbag experiments are not as extensively used for testing peatland models as they could and only a fraction of the information available from litterbag experiments is used to develop models. The HPM derives initial decomposition rates of moss plant functional types from litterbag data, but parameters for environmental controls of decomposition are assumptions which appear to be informed at most qualitatively by litterbag experiments, and it is not tested whether the HPM decomposition module successfully fits available litterbag data (Frolking et al., 2010). This is also the case for other dynamic peatland models, e.g. Frolking et al. (2001), Bauer (2004), Heijmans et al. (2008), Heinemeyer et al. (2010), Morris et al. (2012), Chaudhary et al. (2018), Bona et al. (2020).

One reason why such tests have been difficult is that suitable litterbag raw data to test peatland models are scarce. Bona et al. (2018) developed a Peatland Productivity and Decomposition Parameter Database, but it contains only data from studies older than 2010 and no error estimates for remaining masses in litterbag data. Since decomposition rates have been estimated with different litterbag decomposition models in previous studies, their values are not directly comparable. Moreover, initial leaching losses (losses of soluble compounds, which do not originate from microbial depolymerization, due to leaching during the first days to weeks of incubation) can bias decomposition rate estimates if they are not explicitly considered and can vary between species and experiments (Yu et al., 2001; Teickner et al., 2024b). Therefore, raw data (remaining masses) are necessary for any meaningful test of decomposition modules with litterbag data. The recently published Peatland Decomposition Database (Teickner and Knorr, 2024b) contains raw data from available *Sphagnum* litterbag experiments and therefore allows to estimate parameters with any mass loss-based decomposition model and therefore also allows to consider initial leaching losses.

Even though tests of only a part of a model are less uncertain than tests of whole models, there still is a risk that they are dominated by uncertainties. Remaining masses in litterbag experiments are often very variable, even under controlled environmental conditions (e.g., Bengtsson et al., 2018), and for many litterbag experiments, a range of decomposition rates may produce similar predictions for remaining masses (e.g., Yu et al., 2001), also if a litterbag decomposition model compatible with the HPM is used (Teickner et al., 2024b). Finally, also only five model parameters, as in the case of the HPM decomposition module, can make predictions uncertain. These uncertainties have to be taken into account to check whether litterbag data are compatible with the peatland model. A possible way to do this is to combine the HPM decomposition module, a litterbag decomposition model compatible with this module, and available litterbag experiments into one model and use Bayesian data analysis (Gelman et al., 2014) to estimate uncertainties of data and parameters.

If such a test suggests that decomposition rates predicted by the HPM decomposition module do not fit estimates from litterbag experiments, or only if parameter estimates of the decomposition module differ from the parameter values originally suggested, even if main uncertainty sources are considered, the test has identified a discrepancy worth considering in more detail. We can then analyze whether previous sensitivity analyses of the HPM suggest that these discrepancies may have larger effects on the predicted C accumulation, and if this is the case, the discrepancies are worth testing in future litterbag experiments.

Our aim is to test the HPM decomposition module against decomposition rates estimated from available *Sphagnum* litterbag experiments. Specifically, we want to:

1. Test whether the HPM decomposition module can predict litterbag decomposition rates for different *Sphagnum* species along the gradient from oxic to anoxic conditions.

2. Estimate HPM decomposition module parameters from litterbag data and compare them to the originally suggested values (standard parameter values) (Frolking et al., 2010) that are often used when applying the HPM (Tab. 1).

3. If some of the parameter values differ, identify the possible causes why parameter estimates from litterbag data differ to provide guidance for future litterbag experiments.

4. Analyze whether estimated differences in HPM parameter values could imply significant differences in decomposition rates and long-term peat accumulation.

To address these aims, we used the HPM decomposition module to predict decomposition rates in available litterbag experiments and compared these to decomposition rates estimated for the same litterbag experiments with a litterbag decomposition model that considers initial leaching losses (Teickner et al., 2024b) (Fig. 1). These predictions require the peat degree of saturation, which we estimate with the modified Granberg model (Granberg et al., 1999; Kettridge and Baird, 2007) from water table depth data reported in these studies. Furthermore, some *Sphagnum* litterbag experiments do not report water table depths and therefore cannot be used to test the HPM, but they still provide information on initial leaching losses and decomposition rates and therefore help to constrain parameter estimates. We therefore include these data via Bayesian hierarchical modeling in the litterbag decomposition model. In summary, our approach combines the HPM decomposition module, the modified Granberg model, and a *Sphagnum* litterbag decomposition model that allows to consider intitial leaching losses and to pool information across litterbag experiments (Teickner et al., 2024b). While this approach has its limitations, it exploits available data as far as possible, while considering known confounders and propagating relevant uncertainties.

We only test the decomposition module of the HPM, but the decomposition modules of many other peatland models are also parameterized based on litterbag experiments and our modeling approach is flexible enough to be combined with other decomposition modules. Therefore, our test could serve as a blueprint for similar tests of other peatland model decomposition modules. Similarly, the parameter discrepancies identified here suggest future litterbag experiments that would provide novel insights into oxic and anoxic controls of *Sphagnum* decomposition rates and our study therefore suggests a strategy to improve decomposition modules in general.

Table 1. Standard values of parameters of the decomposition module in the Holocene Peatland Model (Frolking et al., 2010).

HPM parameter	Standard value	Description
W_{opt} ($\mathrm{L_{water}\ L_{pores}^{-1}}$)	0.450	Optimum degree of saturation for aerobic decomposition.
c_1 (-)	2.310	Curvature of the relation of the aerobic decomposition rate to the degree of saturation (larger values imply a steeper decrease of decomposition rates for degrees of saturation diverging from W_{opt}).
f_{min} (yr^{-1})	0.001	Minimum anaerobic decomposition rate.
c_2 (m)	0.300	Anoxia scale length. Represents limitation of anaerobic decomposition rates with increasing distance below the annual average water table depth due to end product accumulation and limitation of available electron acceptors. Larger values mean that anaerobic decomposition rates decrease less strongly with depth below the average annual water table level.
$k_{0,\mathrm{hollow}}$ (yr^{-1})	0.130	Maximum possible decomposition rate for hollow $Sphagnum$ species.
$k_{0,\mathrm{lawn}}$ (yr^{-1})	0.080	Maximum possible decomposition rate for lawn $Sphagnum$ species.
$k_{0,\mathrm{hummock}}$ (yr^{-1})	0.060	Maximum possible decomposition rate for hummock $Sphagnum$ species.

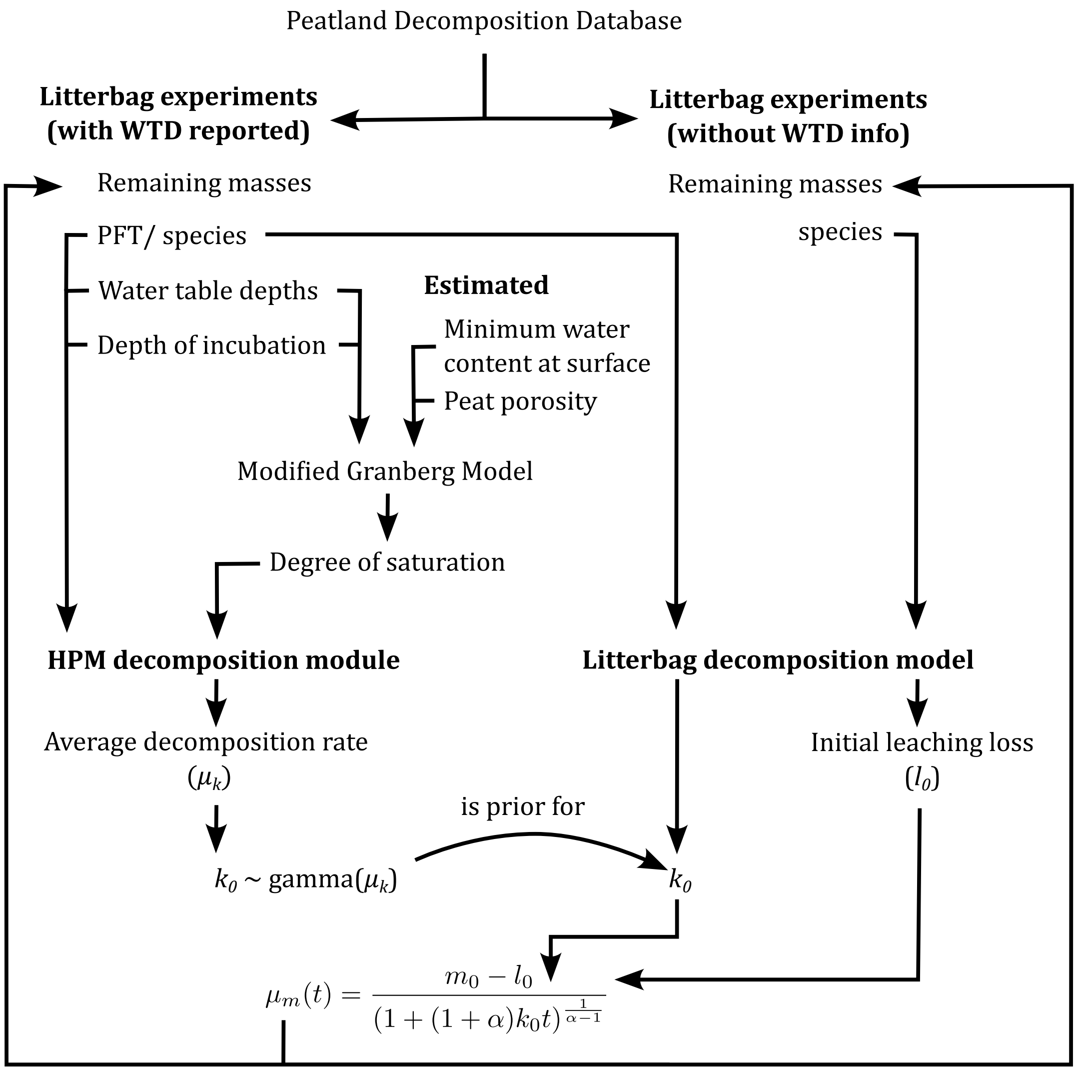

Figure 1. Conceptual representation of the modeling approach. Arrows represent flows of information. Litterbag data that have information on water table depths (WTD) and incubation depths are used to estimate average decomposition rates with the HPM decomposition module (μ_k). The HPM decomposition module needs plant functional type identity, peat degree of saturation, WTD, and incubation depth to predict decomposition rates. The modified Granberg model is used to estimate peat degree of saturation at incubation depths from WTD, minimum water content at the surface, and porosity, of which the latter two are estimated from the remaining masses. The litterbag decomposition model is used to estimate decomposition rates (k_0) for all litterbag studies, including those that have information on WTD and those that have not. A gamma distribution with μ_k as average is used as prior distribution for k_0 for the litterbag experiments that have information on WTD (curved arrow). This helps to constrain initial leaching loss and decomposition rate estimates for studies that can be predicted with the HPM decomposition module. The Litterbag decomposition module also estimates initial leaching losses (l_0) for all litterbag experiments. The equation at the bottom uses these to estimate remaining masses in the litterbag experiments. The litterbag decomposition model is described in more detail in section 2.2.1. See the text for further details.

2 Methods

2.1 *Sphagnum* litterbag data

To test the HPM decomposition module against litterbag data, we used the Peatland Decomposition Database (Teickner and Knorr, 2024b). In this study, we use data from Bartsch and Moore (1985), Vitt (1990), Johnson and Damman (1991), Szumigalski and Bayley (1996), Prevost et al. (1997), Scheffer et al. (2001), Thormann et al. (2001), Asada and Warner (2005), Trinder et al. (2008), Breeuwer et al. (2008), Straková et al. (2010), Hagemann and Moroni (2015), Golovatskaya and Nikonova (2017), and Mäkilä et al. (2018) to estimate k_0 using the litterbag decomposition model. Data from Johnson and Damman (1991), Szumigalski and Bayley (1996), Prevost et al. (1997), Straková et al. (2010), Golovatskaya and Nikonova (2017), and Mäkilä et al. (2018) reported WTD and therefore only these data were used to predict k_0 also with the HPM decomposition module. Samples originally classified as *Sphagnum magellanicum* are here classified as *Sphagnum magellanicum aggr.* (Hassel et al., 2018).

Table 2. Overview on litterbag experiments included for each *Sphagnum* taxon in this study. "HPM microhabitat" is the HPM microhabitat assigned to each taxon. Taxa without value are not considered in Johnson et al. (2015) (see section 2.2.2). "Number of experiments" is the number of litterbag experiments available from the Peatland Decomposition Database (these are either individual replicates or average values of replicates, depending on what data were reported in the studies). "Number of experiments with WTD data" is the number of litterbag experiments that also report water table depths and for which we therefore could make predictions with the HPM decomposition module. "Depth range" are the maximum and minimum depth below the peat surface at which litterbags were placed [cm]. Missing values mean that no study reported depths.

Taxon	HPM microhabitat	Number of studies	Number of experiments	Number of experiments with WTD data	Depth range
Sphagnum spec.		2	16	10	10, 30
S. angustifolium	Hummock	4	14	8	1, 30
S. auriculatum		1	3	0	0, 6
S. balticum	Lawn	3	12	3	1, 30
S. cuspidatum	Hollow	1	5	5	10, 50
S. fallax	Lawn	1	4	1	1, 1
S. fuscum	Hummock	9	32	13	1, 50
S. lindbergii	Lawn	1	2	0	
S. magellanicum aggr.	Hummock	3	7	5	1, 50
S. majus	Hollow	1	2	2	10, 30
S. papillosum	Lawn	2	6	1	0, 1
S. rubellum	Hummock	1	2	2	10, 30
S. russowii	Hummock	1	3	2	1, 1
S. russowii and *capillifolium*		1	18	0	5, 5
S. squarrosum	Lawn	1	2	0	0, 0
S. teres	Lawn	1	1	1	2, 2

2.2 Prediction of litterbag decomposition rates with the Holocene Peatland Model

To predict decomposition rates, the HPM decomposition module needs as inputs the litter type in terms of the HPM plant functional types (PFT), the fraction of mass already lost due to previous decomposition, the depth of the litter below the peat surface, the water table depth, and the peat degree of saturation (Frolking et al., 2010).

Predicting decomposition rates for the available litterbag data is not straightforward because the HPM decomposition module does not consider specific features of available litterbag experiments. The HPM does not specify how to assign species to plant functional types. Moreover, none of the available litterbag studies reported the degree of saturation which therefore needs to be estimated in order to make predictions with the HPM decomposition module. The only variables that can be directly linked are the depth of the litter below the peat surface, and water table depths (both reported in litterbag experiments). All other variables can be estimated from available litterbag experiments only with additional assumptions that are described in the following subsections.

In the following subsection, we give a more detailed description of our modeling approach, in particular of the model used to estimate decomposition rates from litterbag data, of the HPM decomposition module and how it predicts decomposition rates, and how we link the decomposition rates estimated from litterbag data to those predicted by the HPM decomposition module. The remaining subsections discuss how we derived or estimated PFT, WTD, and degree of saturation for the litterbag data and additional steps to make the litterbag data compatible with the HPM decomposition module.

2.2.1 Remaining masses and decomposition rates

To estimate decomposition rates for available *Sphagnum* litterbag experiments we use the equation from the HPM that computes remaining masses from decomposition rates and decomposition time (Frolking et al., 2001, 2010), with three modifications. The original equation (equation (4) in Frolking et al. (2001)) is:

$$m(t) = \frac{m_0}{(1 + (1 + \alpha)k_0 t)^{\frac{1}{\alpha - 1}}}, \tag{1}$$

where $m(t)$ is the fraction of initial mass remaining at time t, m_0 is the fraction of initial mass remaining at time $t = 0$, k_0 is the decomposition rate, and α is a parameter that describes how decomposition slows down as mass is lost, where the HPM assumes $\alpha = 2$ for simplicity (Frolking et al., 2001, 2010).

The modified version we use here is:

$$\mu_m(t) = \begin{cases} m_0 & \text{if } t = 0 \\ \frac{m_0 - l_0}{(1+(\alpha-1)k_0 t)^{\frac{1}{\alpha-1}}} & \text{if } t > 0 \end{cases}, \tag{2}$$

where l_0 is the fraction of mass lost due to initial leaching. The HPM decomposition process does not assume that there are initial leaching losses, but these are commonly observed in litterbag experiments and bias decomposition rate estimates when they are ignored (Yu et al., 2001; Teickner et al., 2024b); therefore, the modification is necessary to allow a sensible test of the HPM decomposition module with litterbag data.

The second modification is that we do not assume $\alpha = 2$, but consider it as unknown parameter that is estimated from litterbag data. Since $\alpha = 2$ was chosen for simplicity and attempts to reliably estimate α have failed (e.g., Clymo et al., 1998; Frolking et al., 2001; Teickner et al., 2024b), we estimate α mainly to consider the possible error introduced by this parameter.

The third modification is that we change $m(t)$ to $\mu_m(t)$ because we assume that equation (2) describes only the average fraction of the initial mass remaining. For each retrieved litterbag, we assume that the remaining mass can be described with a beta distribution with precision parameter ϕ_m:

$$m(t) \sim \text{beta}(\mu_m(t)\phi_m, (1 - \mu_m(t))\phi_m), \tag{3}$$

Values for k_0 are estimated from remaining masses reported in available litterbag experiments conditional on equation (2) and a hierarchical prior structure (Teickner et al., 2024b):

$$k_0 = \exp(\beta_{k,1} + \beta_{k,2,\text{species}} + \beta_{k,3,\text{species x study}} + \beta_{k,4,\text{sample}}), \tag{4}$$

where $\beta_{k,1}$ is the estimated decomposition rate across all litterbag experiments, $\beta_{k,2,\text{species}}$ describes the difference of the average decomposition rate for the *Sphagnum* species, $\beta_{k,3,\text{species x study}}$ for the study (nested within species), and $\beta_{k,4,\text{sample}}$ for the sample (litterbag experiment). All these parameters have normal distributions as priors. Hierarchical models of the same structure are used to estimate l_0 and α from equation (2) and to estimate ϕ_m from equation (3).

These decomposition rates estimated from litterbag experiments are constrained by decomposition rates the HPM decomposition module (Frolking et al., 2010) predicts for the same litterbag experiments. The HPM decomposition module describes how decomposition rates depend on the *Sphagnum* PFT, the degree of saturation and the depth of a litter sample below the water table. Similar to the remaining mass, we here assume that the HPM decomposition module predicts an average decomposition rate, μ_k, instead of the decomposition rate of individual samples:

$$\mu_k = \begin{cases} k_{0,i} f_1(W) & \text{if } \hat{z} \leq 0 \\ k_{0,i} f_2(\hat{z}) & \text{if } \hat{z} > 0 \end{cases} \tag{5}$$

where $k_{0,i}$ is the PFT-specific maximum possible decomposition rate (Tab. 1), W is the degree of saturation ($\text{L}_\text{water}\ \text{L}_\text{sample}^{-1}$), $\hat{z}$ the depth of the sample below the average annual water table ($\hat{z} = z - z_\text{wt}$, where z_wt and z are the depth of the water table and litterbag below the peat surface), and f_1 and f_2 are modifiers due to W (under oxic conditions) and $\hat{z}$ (under anoxic conditions), respectively. These modifiers are described in equations (8) and (9) in Frolking et al. (2010):

$$f_1(W) = 1 - c_1(W - W_{opt})^2 \tag{6}$$

$$f_2(\hat{z}) = f_{min} + (f_1(1) - f_{min}) \exp\left(\frac{-\hat{z}}{c_2}\right), \quad (7)$$

where all not yet mentioned parameters are defined in Tab. 1.

In our model, k_0 estimated from the litterbag data for each litterbag experiment with reported WTD (sample) (equation (2)) is assumed to follow a gamma distribution with shape parameter α_{μ_k} (estimated) and average μ_k (predicted for each sample with equation (5)):

$$k_0 \sim \text{gamma}\left(\alpha_{\mu_k}, \frac{\alpha_{\mu_k}}{\mu_k}\right), \quad (8)$$

Thus, the decomposition rate predicted by the HPM decomposition module (equation (8)) is a prior for k_0 as estimated from the litterbag decomposition model (equation (4)). This forms the link between the litterbag decomposition model and the HPM decomposition module (Fig. 1) and also allows us to estimate parameters of the HPM decomposition module from the litterbag data. The advantage of this modeling approach is that we can consider litterbag experiments without water table depth to estimate l_0 and k_0 for individual *Sphagnum* species, which is additional information to constrain estimates of the HPM decomposition module parameters. Moreover, combining the litterbag decomposition model and the HPM decomposition module into one Bayesian model does not only estimate HPM decomposition module parameters from the litterbag data, but it also adjusts the decomposition rates estimated from litterbag data to the HPM decomposition module because the HPM decomposition module serves as prior in the combined model which therefore estimates what parameter values are compatible with the data and the combined model. This is exactly what we want because there is uncertainty both in the remaining masses reported in litterbag experiments and in HPM decomposition module parameters. If HPM decomposition module parameter estimates from the combined model are different from the standard values used in the original model (Tab. 1), even if we consider these uncertainties and use the HPM decomposition module as prior for the litterbag data, this is a discrepancy worth testing in future experiments.

2.2.2 Assignment of *Sphagnum* species to plant functional types

The HPM defines maximum possible decomposition rates ($k_{0,i}$) for three *Sphagnum* PFT (hollow, lawn, and hummock species), but not how to assign species to them. We assigned individual *Sphagnum* species to the three PFT by comparing their niche WTD with the optimal WTD for net primary production defined in the HPM. Specifically, we defined fixed average annual WTD intervals for the PFT: hollow (<5 cm), lawn (≥ 5 cm and < 15 cm), hummock (≥ 15 cm) based on the HPM (Frolking et al., 2010). Then, we used niche WTD and standard deviations from Johnson et al. (2015) to assign *Sphagnum* species to these three microhabitats. Using only average values and the microhabitat WTD thresholds resulted in unintuitive assignments, such as assigning *S. fallax* to hummocks. To avoid such obvious misclassifications, we defined rules to assign species to HPM microhabitats based on the probability a species would occur in the three niche WTD intervals. To compute the probabilities, we assumed a normal distribution (Johnson et al., 2015):

1. Species with a probability of occurrence $\geq 15\%$ in the intervals of all three PFT were classified as lawn species.

2. In all other cases, species were assigned to the PFT for which their probability of occurrence was largest.

Litterbag data from Prevost et al. (1997) are incubations of peat samples where the species is unknown. Based on descriptions in this study, it is likely that the peat was formed by hummock species. Hummock species are assumed to have the smallest decomposition rate among the three *Sphagnum* PFT in the HPM (Frolking et al., 2010) and this is in line with small decomposition rate estimates for these samples (Teickner et al., 2024b). For these reasons, we assigned these samples to the hummock PFT of the HPM.

When estimating parameters of the HPM decomposition module from the litterbag data (see section 2.3.1), we also estimated the maximum possible decomposition rate ($k_{0,i}$). *Sphagnum* species differ in their decomposition rate and the PFT of the HPM are a simplification that may cause misfits of the HPM decomposition module to litterbag data. We therefore estimated $k_{0,i}$ for individual *Sphagnum* species in models HPM-all, HPM-leaching, and HPM-outlier (see section 2.3.1) and evaluated the variability of these species-specific estimates compared to the standard $k_{0,i}$ values of the HPM *Sphagnum* PFT.

2.2.3 Degree of saturation

We estimated the degree of saturation with the modified Granberg model (ModGberg model) (Granberg et al., 1999; Kettridge and Baird, 2007) from minimum water content at the surface ($\theta_{0,\text{min}}$), total porosity (P), the water table depth below the peat surface (z_{wt}), and the depth of the litterbags below the peat surface during the incubation (z):

$$\begin{aligned} \theta(z) = &\quad \min\left(P, \theta_0 + (P - \theta_0)\left(\frac{z}{z_{\text{wt}}}\right)^2\right) \\ \theta_0 = &\quad \max\left(\theta_{0,\text{min}}, 0.15 z_{\text{wt}}^{-0.28}\right), \end{aligned} \tag{9}$$

where θ_0 is the water content at the surface and $0.15 z_{\text{wt}}^{-0.28}$ is an empirical relation of θ_0 with the WTD (Kettridge and Baird, 2007).

The minimum water content at the surface was not reported in any study and we therefore assumed a minimum water content at the surface of 0.05 $\text{L}_{\text{water}}\ \text{L}_{\text{sample}}^{-1}$ with a standard deviation of 0.05 $\text{L}_{\text{water}}\ \text{L}_{\text{sample}}^{-1}$, based on measurements from Hayward and Clymo (1982). The total porosity was not reported in any study and therefore we assumed an average value of 80% with a standard deviation of 10%, roughly based on values reported for low-density *Sphagnum* peat (Liu and Lennartz, 2019). An improved test of the HPM decomposition module would require litterbag experiments with direct measurements of the degree of saturation at sufficient temporal resolution.

2.2.4 Fraction of mass lost during previous decomposition

The HPM decomposition module assumes that decomposition rates decrease the more of the initial mass has already been decomposed (Frolking et al., 2001, 2010). All litterbag data we use here, except samples from Prevost et al. (1997), are from

Sphagnum samples collected from the surface of peatlands and therefore can be expected to have not experienced mass loss due to decomposition at the start of the experiments ($m(t = 0) = 1$ in equation (2)). Prevost et al. (1997) incubated *Sphagnum* peat collected from two different depth levels from the same location and these samples probably had already experienced some decomposition, however it is difficult to estimate how much. To avoid this problem, we estimated $k_{0,i}$ separately for samples from different depths in Prevost et al. (1997), implicitly assuming that these are two different PFT with different maximum possible decomposition rate.

2.3 Testing the HPM decomposition module against litterbag data

2.3.1 Model versions

To test different aspects of the HPM decomposition module and the additional assumptions we make, we computed several models which differ in whether HPM decomposition module parameters were fixed to their standard values or estimated from data, whether peat properties (porosity, water table depth, water content, minimum water content at the surface) are estimated from data or not, and whether the HPM decomposition module was extended to also predict l_0 or not (Tab. 3).

The first model (HPM-standard) does not estimate any parameters of the HPM decomposition module (equations (5) to (7)) and does not estimate peat properties from the litterbag data and therefore corresponds to the HPM decomposition module with standard parameter values, while propagating prior uncertainties for peat properties. For this model, predictions of k_0 equal μ_k (equation (5)). This version of the HPM decomposition module is completely independent of the litterbag decomposition model, meaning that the HPM decomposition module is not used as prior for the litterbag decomposition model (Fig. 1). This also means that to compare k_0 predicted by HPM-standard to k_0 estimated from the litterbag decomposition model, we need to estimate the littebag decomposition model indpendently, without using the HPM decomposition module as prior. This independent litterbag decomposition model is called LDM-standard (Tab. 3). We use LDM-standard not only to compare k_0 estimates to k_0 predictions of HPM-standard, but also to analyze how k_0 estimates of the litterbag decomposition model changes when we use different versions of the HPM decomposition module as prior in the subsequent models.

Each subsequent model combines the HPM decomposition module and the litterbag decomposition model into one Bayesian model via equation (8). Each of these models estimates an additional set of parameters from the litterbag data relative to the previous model (Tab. 3). First, only the peat properties (HPM-peat) are estimated, and second all HPM parameters ($k_{0,i}$, c_1, W_{opt}, f_{min}, c_2) (HPM-all). Finally, HPM-leaching extends HPM-all by adding formulas to model how l_0 depends on the degree of saturation, similar to how the HPM decomposition module predicts k_0 with equation (6).

HPM-peat tested whether the HPM decomposition module can fit available litterbag data when the HPM decomposition module and the litterbag decomposition model are combined and when peat properties are estimated from data.

HPM-all estimates what HPM decomposition module parameter values are compatible with available litterbag data and therefore allows to test whether the standard parameter values are extreme relative to these estimates. Values of $k_{0,i}$ were estimated for each species separately, as described in section 2.2.2.

HPM-leaching was computed because decomposition rates estimated from available litterbag experiments are sensitive to initial leaching losses (Yu et al., 2001; Lind et al., 2022; Teickner et al., 2024b). It is therefore interesting to see whether litterbag decomposition rates are estimated differently in HPM-leaching — when initial leaching losses are constrained by adding formulas to model how l_0 depends on the degree of saturation — compared to HPM-all — when initial leaching loss estimates are constrained only by the litterbag decomposition model. Based on previous experiments with tea bags it is reasonable to assume that there is some relation between initial leaching losses and the degree of saturation (Lind et al., 2022). Specifically, we use the following logistic regression model to describe an average initial leaching loss per sample, in dependency of the degree of saturation:

$$\begin{aligned} \mu_l = & \quad \mathrm{logit}^{-1}(\beta_{l,1} + \beta_{l,2} W) \\ l_0 \sim & \quad \mathrm{beta}(\mu_l \phi_l, (1 - \mu_l)\phi_l), \end{aligned} \tag{10}$$

where μ_l is the average initial leaching loss for a sample, $\beta_{l,1}$ is the (hypothetical) average initial leaching loss at a degree of saturation 0 for each taxon, $\beta_{l,2}$ is the coefficient that describes the relation to the degree of saturation (W), and ϕ_l transforms μ_l and $(1 - \mu_l)$ into the shape and rate parameters of a beta distribution. This beta distribution has the same function as the gamma distribution (equation (8)) for k_0 (compare also with Fig. 1): it is a prior for l_0 estimated with the litterbag decomposition model, where the average of this prior is μ_l.

To check whether outliers in the litterbag data could influence our results, we computed one additional model, HPM-outlier, with the same structure as HPM-leaching, but estimated without litterbag experiments identified as outliers. Litterbag experiments were defined as outliers if the reported average remaining mass of any litterbag (batch) during the experiment had a posterior probability $> 99\%$ to be different from the remaining mass predicted by the litterbag decomposition model alone. This procedure identified experiments as outliers where remaining masses increased over time, where litterbags collected at intermediate time points had unexpectedly low remaining masses, or where initial leaching losses were retarded to later time points, presumably because of freezing after the start of the experiment (Teickner et al., 2024b). In total, 5 litterbag experiments were identified as outliers. Results for HPM-outlier are shown in supporting information S8 and HPM decomposition module parameter estimates agree with estimates of HPM-leaching and HPM-all.

Strictly, we do not test the decomposition module in the HPM, but the combination of the HPM decomposition module and the modified Granberg model, assuming that uncertainties in water table depths are negligible and that we accounted sufficiently for uncertainties in total porosity. This ambiguity has to be accepted when combining heterogeneous litterbag data where some variables have to be estimated. Litterbag experiments where water table depths and the degree of saturation are measured at sufficient temporal resolution are needed to avoid this ambiguity in future studies and to improve any test of the HPM decomposition module.

Table 3. Overview of HPM decomposition module modifications computed in this study.

Model	Description
LDM-standard	The litterbag decomposition model without the HPM decomposition module as prior. This is model 1-4 from Teickner et al. (2024b).
HPM-standard	The Holocene Peatland Model decomposition module with standard parameter values (Frolking et al., 2010). The model is run with peat water contents estimated with the modified Granberg model, using water table depths and litterbag depths reported from the litterbag studies, and assuming a fixed peat porosity, and minimum peat water content at the surface.
HPM-peat	The same as HPM-standard, but combined with LDM-standard into one Bayesian model, where the HPM decomposition module is a prior for the litterbag decomposition model (Fig. 1). Water table depths, peat porosity, and minimum peat water content at the surface are estimated from data.
HPM-all	The same as HPM-peat, but now also parameters from the HPM decomposition module ($k_{0,i}$, W_{opt}, f_{min}, c_1, c_2) are estimated from the litterbag data.
HPM-leaching	The same as HPM-all, but now also an average initial leaching loss for each species and, across all species, a factor, by which this average leaching loss increases or decreases as the peat degree of saturation increases, are estimated (equation (10)).
HPM-outlier	The same as HPM-leaching, but computed without litterbag experiments that were identified as outliers (see the text for details).

2.3.2 Bayesian data analysis

All models listed in Tab. 3 were computed with Bayesian statistics to account for relevant uncertainty sources and include relevant prior knowledge (for example that *Sphagnum* decomposition rates are unlikely to be larger than 0.5 yr^{-1}). Bayesian computations were performed using Markov Chain Monte Carlo (MCMC) sampling with Stan (2.32.2) (Stan Development Team, 2021a) in R (4.2.0) (R Core Team, 2022) via the rstan package (2.32.5) (Stan Development Team, 2021b) using the NUTS sampler (Hoffman and Gelman, 2014), with four chains, 4000 total iterations per chain, and 2000 warmup iterations per chain. None of the models had divergent transitions, the minimum bulk effective sample size was larger than 400, and the largest rank-normalized $\hat{R}$ was 1.01, indicating that all chains converged (Vehtari et al., 2021). All models used the same priors for the same parameters and prior choices are listed and justified in supporting Tab. S1. Results of prior and posterior predictive checks are shown in supporting section S3.

We used power-scaling of the prior and likelihood distributions as implemented in the priorsense package (0.0.0.9000) (Kallioinen et al., 2024) to analyze the relative sensitivity of the posterior distribution to small perturbations of the prior and likelihood in HPM-leaching for HPM decomposition module parameters and peat properties. This is a computationally nonexpensive way to check whether the data provide information about a parameter and where prior and data may provide conflicting information (Kallioinen et al., 2024). Results of this analysis and further information on the data analysis are shown in supporting information S2.

2.3.3 Fit of model pedictions to estimated decomposition rates and observed remaining masses

To analyze how well the models fit remaining masses observed in the litterbag experiments, we plotted reported remaining masses versus remaining masses estimated by the litterbag decomposition model in HPM-peat, HPM-all, and HPM-leaching. HPM-standard is not linked to the litterbag decomposition model and therefore does not predict remaining masses.

To analyze how well all HPM decomposition module versions fit k_0 estimated by the respective litterag decomposition model, we created a similar plot for k_0. Here, we compared predictions of HPM-standard (equation (8)) against estimates of LDM-standard (equation (4)). We also computed the average difference of k_0 predicted by the HPM decomposition module and estimated from the litterbag data. We then computed the posterior probability that this average difference is different from zero. A large probability indicates a misfit of the model to available litterbag data. We also tested the same difference for specific species because graphical checks indicated that the decomposition rate prediction skill of the HPM decomposition module depends on species.

To test whether HPM-leaching has not only a better fit to available litterbag data, but also a better predictive accuracy for novel data than the model with standard parameter values (HPM-standard), we compared how well both can predict k_0 from litterbag experiments. HPM decomposition module parameters of HPM-standard are not estimated from data and therefore we could compute the root mean square error of prediction ($\text{RMSE}_{\text{test}}$) directly with k_0 predicted by HPM-standard and estimated with LDM-standard. HPM decomposition module parameters of HPM-leaching are estimated from the litterbag data and we therefore used cross-validation (CV) to estimate $\text{RMSE}_{\text{test}}$. Since decomposition rates from the same species and study usually are not independent, we defined blocks which were used as CV-folds. Each fold represents the data from one study, but only if there were still data for the same *Sphagnum* species left in the remaining data (we want to estimate the predictive accuracy not for new species). Species with data from one study only were always used for model training and not part of the testing folds. This procedure resulted in 5 folds. HPM-standard and HPM-leaching were tested against the same data. In the text, $\text{RMSE}_{\text{train}}$ is the RMSE computed with the data a model was estimated with (for HPM-standard, the data the litterbag decomposition model was estimated with), and $\text{RMSE}_{\text{test}}$ is the RMSE computed with independent test data.

2.3.4 Changes in k_0 and l_0 estimates of the litterbag decomposition models compared to LDM-standard

To analyze how parameter values of the litterbag decomposition model change when it is combined with different versions of the HPM decomposition module as prior, we estimated the average difference of k_0 and l_0 estimates of each model version to k_0 and l_0 estimates of LDM-standard. In particular, this allowed us to analyze whether there is any change in the relative magnitude of l_0 and k_0 because the litterbag decomposition module would adjust these parameter values to fit the respective HPM decomposition module prior and still fit the observed remaining masses.

2.3.5 Magnitudes of k_0 along the gradient from oxic to anoxic conditions

To analyze how k_0 changes along the gradient from oxic to anoxic conditions, we plotted k_0 estimated by LDM-standard versus the water table depth below the litterbags. To this plot, we added k_0 predicted by HPM-standard. To analyze how the relation of k_0 changes for the HPM decomposition module modifications compared to HPM-standard, we computed differences between k_0 estimated by HPM-peat, HPM-all, and HPM-leaching, respectively, and HPM-standard, and plotted these differences versus the water table depth below the litterbags.

2.3.6 Difference between values of $k_{0,i}$, c_1, W_{opt}, c_2, f_{min} estimated from litterbag data to the standard parameter values

For HPM-all and HPM-leaching, we computed the posterior probability that the HPM decomposition module parameter values estimated from litterbag data ($k_{0,i}$, c_1, W_{opt}, c_2, f_{min}) differ from the standard parameter values (Tab. 1). This way, we could identify discrepancies between standard parameter values and parameter values estimated from available litterbag data.

For HPM-leaching, we conducted in addition a sensitivity analysis, where we simulated decomposition of *S. fuscum* incubated at different depths in a peatland with water table depth of 40 cm below the surface, a porosity of 0.7 $\text{L}_\text{pores}\ \text{L}_\text{sample}^{-1}$, and a minimum water content at the surface of 0.05 $\text{g}_\text{water}\ \text{g}_\text{sample}^{-1}$. With these settings, we predicted five sets of average k_0: (1) with HPM-leaching ($k_{0,\text{modified}}$(HPM-leaching)). The remaining four sets were also predicted with HPM-leaching, but each time setting one of the HPM decomposition module parameters to their standard value ($k_{0,\text{standard}}$(HPM-leaching)): (2) c_1, (3) W_{opt}, (4) f_{min}, (5) c_2. We then computed the difference of k_0 from set (1) and (2) to analyze the effect of the new c_1 estimate, from set (1) and (3) to analyze the effect of the new W_{opt} estimate, and so on for sets (4) and (5). This gives the difference in decomposition rates of HPM-leaching if we would set individual HPM decomposition module parameters to their standard values. This way, we could analyze what HPM decomposition module parameters contribute to a change in k_0 predictions along the gradient from oxic to anoxic conditions.

3 Results

3.1 Fit and predictive accuracy of the different versions of the HPM decomposition module to available litterbag data

In each model, the litterbag decomposition model fitted the observed remaining masses similarly well (Fig. 2 (a) and supporting Fig. S2), no matter whether the HPM decomposition module was used as prior or not, and whether its parameters were estimated from data (HPM-all, HPM-leaching) or not (HPM-peat). Thus, remaining masses do not indicate large differences between the model versions.

For k_0, the picture is more nuanced: When the HPM decomposition module is not used as prior (HPM-standard), it fitted k_0 estimated by the litterbag decomposition model on average less well than when it was used as prior (all other model versions) (Fig. 2, Tab. 4). For example, HPM-standard had an average $\text{RMSE}_{\text{train}}$ of 0.11 yr^{-1}, whereas HPM-leaching had an average 385 $\text{RMSE}_{\text{train}}$ of 0.02 yr^{-1}. However, the cross-validation indicates that when applied to novel samples, both HPM-standard and HPM-leaching would perform similarly well if one considers the large uncertainty of the $\text{RMSE}_{\text{test}}$ estimates (Tab. 4). Interestingly, all model versions where the HPM decomposition module was used as prior had comparable fits ($\text{RMSE}_{\text{train}}$) (Tab. 4), even the version that still has standard parameter values for the HPM decomposition module (HPM-peat). This indicates that a change in HPM decomposition module standard parameter values is not required to make the HPM decomposition module fit k_0 values estimated from available litterbag data via the litterbag decomposition model, under the assumptions we made. Instead, the results indicate that parameter values of the litterbag decomposition model can be adjusted to fit predictions of this HPM decomposition module prior.

Table 4. Training and testing RMSE for decomposition rates as predicted by different versions of the decomposition module of the Holocene Peatland Model (see Tab. 3 for a description of the models) and number of misfits. $\text{RMSE}_{\text{train}}(k_0)$ is the root mean square error of model predictions for litterbag replicates used during model computation. $\text{RMSE}_{\text{test}}(k_0)$ is the RMSE for litterbag replicates used in blocked cross-validation. Where no $\text{RMSE}_{\text{test}}(k_0)$ is given, it was not computed for these models. Values are averages and lower and upper bounds of central 95% posterior intervals (yr^{-1}). Misfits counts the number of litterbag experiments for which k_0 predicted by the HPM decomposition module modification differed from k_0 as estimated from the litterbag decomposition model with a posterior probability of at least 99%. In total, k_0 was predicted with the HPM decomposition module modifications for 53 litterbag experiments ($\text{RMSE}_{\text{train}}(k_0)$) or 29 ($\text{RMSE}_{\text{test}}(k_0)$).

Model	$\text{RMSE}_{\text{train}}(k_0)$	$\text{RMSE}_{\text{test}}(k_0)$	Misfits
HPM-standard	0.105 (0.051, 0.191)	0.136 (0.06, 0.252)	13
HPM-peat	0.02 (0.013, 0.029)		0
HPM-all	0.014 (0.008, 0.021)		0
HPM-leaching	0.022 (0.012, 0.039)	0.088 (0.038, 0.179)	0
HPM-outlier	0.021 (0.013, 0.032)		0

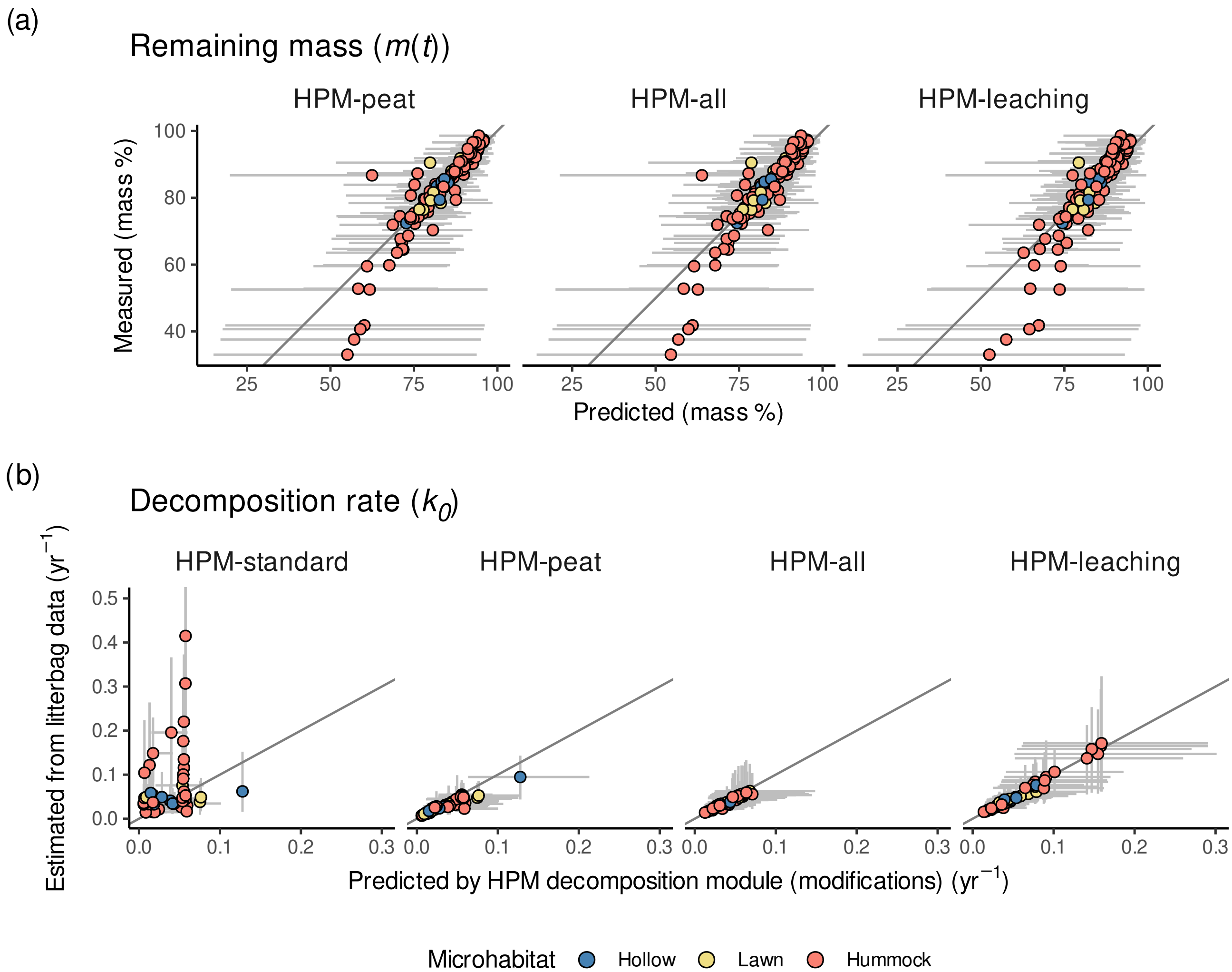

Figure 2. (a) Measured remaining masses versus remaining masses predicted by the litterbag decomposition model combined with each HPM decomposition module version. Values are shown for litterbag experiments with reported water table data. For HPM-standard no values are shown because it was not combined with a litterbag decomposition model. (b) k_0 estimated by the litterbag decomposition model versus k_0 predicted by different modifications of the HPM decomposition module (Tab. 3). For HPM-standard, y-axis values are k_0 estimates of LDM-standard. For all other model versions, y-axis values are k_0 estimates of the litterbag decomposition module with the respective HPM decomposition module version as prior. Points represent average estimates and error bars 95% posterior intervals. Points are colored according to the microhabitat classification of *Sphagnum species* (see the Methods section for details). In (b), error bars exceeding 0.5 yr^{-1} are clipped.

3.2 How are parameter values of the litterbag decomposition model adjusted when different versions of the HPM decomposition module are used as prior?

To understand how using the HPM decomposition module as prior changes k_0 and l_0 estimates of the litterbag decomposition model, we compared k_0 and l_0 estimates of the litterbag decomposition model of each model version with the k_0 and l_0 estimates of LDM-standard. We computed the average difference of k_0 estimates by the litterbag decomposition model for all models compared to the k_0 estimates of LDM-standard (using only litterbag experiments with reported WTD). Average differences compared to LDM-standard are in the order HPM-peat < HPM-all < HPM-leaching (average and 95% confidence interval: -0.04 (-0.06, -0.02) < -0.03 (-0.06, -0.01) < -0.01 (-0.04, 0.01) yr^{-1}). The magnitude (mean absolute difference) of adjustments of k_0 estimates is different for different species (species with at least 3 samples): The largest average absolute differences across all models were made for *S. angustifolium* (0.15 (0.06, 0.27) yr^{-1}) and the smallest for *Sphagnum* spec. (0.01 (0.01, 0.02) yr^{-1}). This indicates that for some species k_0 estimates of the litterbag decomposition model are forced to smaller values for HPM-peat and HPM-all, whereas differences are smaller for HPM-leaching.

With these changes in k_0 estimates, a similar fit to remaining masses as observed for all models (see the previous subsection) is only possible when l_0 estimates are changed in the opposite direction. To check this, we computed the average difference of l_0 estimates by the litterbag decomposition model for all model versions compared to the l_0 estimates of LDM-standard. Differences compared to LDM-standard are in the order HPM-leaching < HPM-all < HPM-peat (average and 95% confidence interval: 0.1 (-1.9, 2.2) < 2.8 (0.7, 4.8) < 3.3 (1.6, 5) mass %). Again, the magnitude (mean absolute difference) of adjustments of l_0 estimates is different for different species (species with at least 3 samples): The largest average absolute differences across all models were made for *S. angustifolium* (11.4 (7, 16.6) mass %) and the smallest for *Sphagnum* spec. (1.43 (0.86, 2.39) mass %). Thus, the smaller k_0 estimates are indeed compensated by larger l_0 estimates for HPM-peat and HPM-all, whereas the difference to LDM-standard is smaller for HPM-leaching.

Overall, this analysis indicates that errors in remaining masses observed in available litterbag experiments are large enough to support a range of k_0 and l_0 estimates. The equivalent fit of the different model versions is therefore caused by adjusting k_0 to the HPM proir, and adjusting l_0 as needed to fit observed masses.

3.3 How do HPM decomposition module parameter estimates differ to the standard values?

Two model versions estimated HPM decomposition module parameters (c_1, W_{opt}, f_{min}, c_2, $k_{0,i}$): HPM-all and HPM-leaching. These models indicate larger values for c_2 and W_{opt} than the standard values. Figure 3 shows marginal posterior densities of the HPM decomposition module parameters for HPM-all, with standard parameter values as defined in Frolking et al. (2010) indicated by vertical lines. For both HPM-all and HPM-leaching, there are large posterior probabilities that c_2 ($P_{\text{HPM-all}}(c_2 > 0.3\text{ m}) = 1$ and $P_{\text{HPM-leaching}}(c_2 > 0.3\text{ m}) = 1$) and W_{opt} ($P_{\text{HPM-all}}(W_{opt} > 0.45\text{ L}_{\text{water}}\text{ L}_{\text{pores}}^{-1}) = 1$ and $P_{\text{HPM-leaching}}(W_{opt} > 0.45\text{ L}_{\text{water}}\text{ L}_{\text{pores}}^{-1}) = 0.98$) have larger values than the standard parameter values, indicating a discrepancy between the HPM decomposition module and available litterbag data (Fig. 3 and supporting Fig. S11). In contrast, estimates for f_{min} do not differ much to the prior value and the power-scaling sensitivity analysis indicates a weak influence of the data (supporting information

S2) and therefore that currently available litterbag data provide only little information about minimum decomposition rates under anoxic conditions. HPM-all and HPM-leaching suggest a large variability of $k_{0,i}$ for individual species: Both models estimate a large posterior probability ($> 95\%$) that *S. russowii* and *S. rubellum* have a larger, and that *S. cuspidatum* has a smaller maximum possible decomposition rate ($k_{0,i}$) than the standard values for the respective PFT (Fig. 3 (b) and supporting Fig. S11). However, estimates for $k_{0,i}$ were very variable for the same species when different subsets of the litterbag data were used to estimate the model in the cross-validation. This indicates that samples of the same species from different studies have a large variability in $k_{0,i}$ values. In summary, when HPM decomposition module parameters are estimated from available litterbag data, estimates for W_{opt} and c_2 are larger than the standard values, differences to the c_1 and f_{min} standard value cannot be detected, and estimates for $k_{0,i}$ are variable and have large errors for different species.

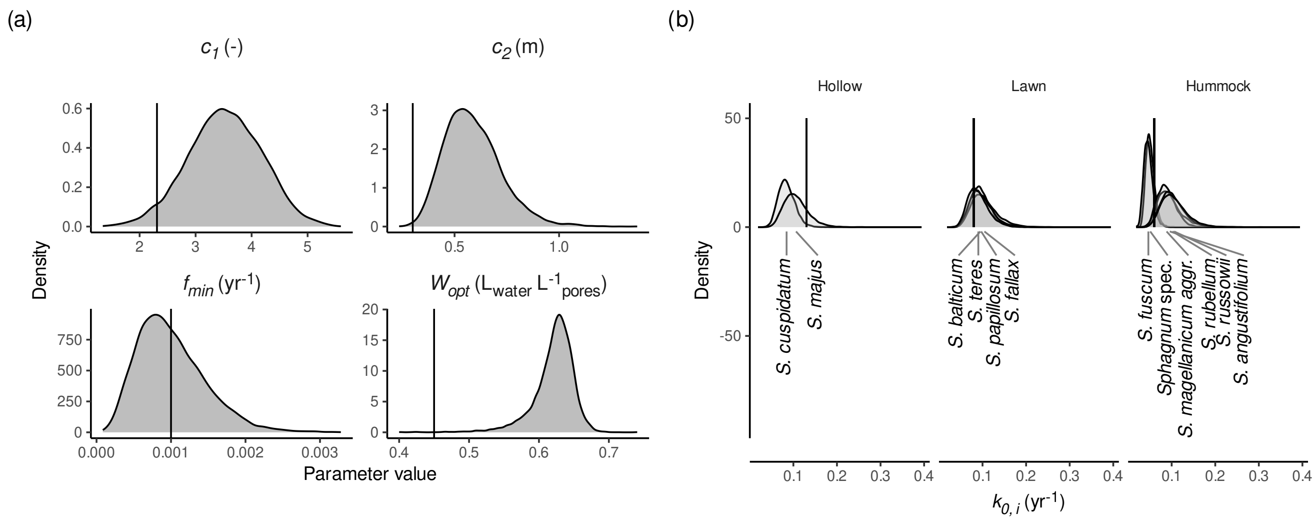

Figure 3. Marginal posterior distributions of HPM decomposition module parameters (see Tab. 1) as estimated by HPM-all. (a) Marginal posterior distributions for c_1, W_{opt}, f_{min}, and c_2. (b) Marginal posterior distributions for $k_{0,i}$ (maximum possible decomposition rate for species i). Species were assigned to HPM microhabitats as described in section 2.2.2. Vertical black lines are the standard parameter values from Frolking et al. (2010). *Sphagnum* spec. are samples that have been identified only to the genus level.

3.4 Magnitude and change of decomposition rates along the gradient from oxic to anoxic conditions

A comparison of k_0 estimates of LDM-standard and k_0 estimates of HPM-standard shows that the HPM decomposition module with standard parameter values implies a steeper decrease of decomposition rates from oxic to anoxic conditions than LDM-standard and, for some species, smaller anaerobic decomposition rates. Figure 4 (a) shows k_0 estimated by LDM-standard and k_0 predicted by HPM-standard versus water table depths below the litterbags reported in the studies for species with at least three litterbag experiments. Regression lines were fitted to both sets of k_0 values and they indicate an on average steeper slope for HPM-standard than for LDM-standard for many species (with large uncertainties). Moreover, under anoxic conditions (negative water table depth), k_0 estimates by LDM-standard are larger on average for many of the litterbag experiments than what HPM-standard predicts (Fig. 4 (b)).

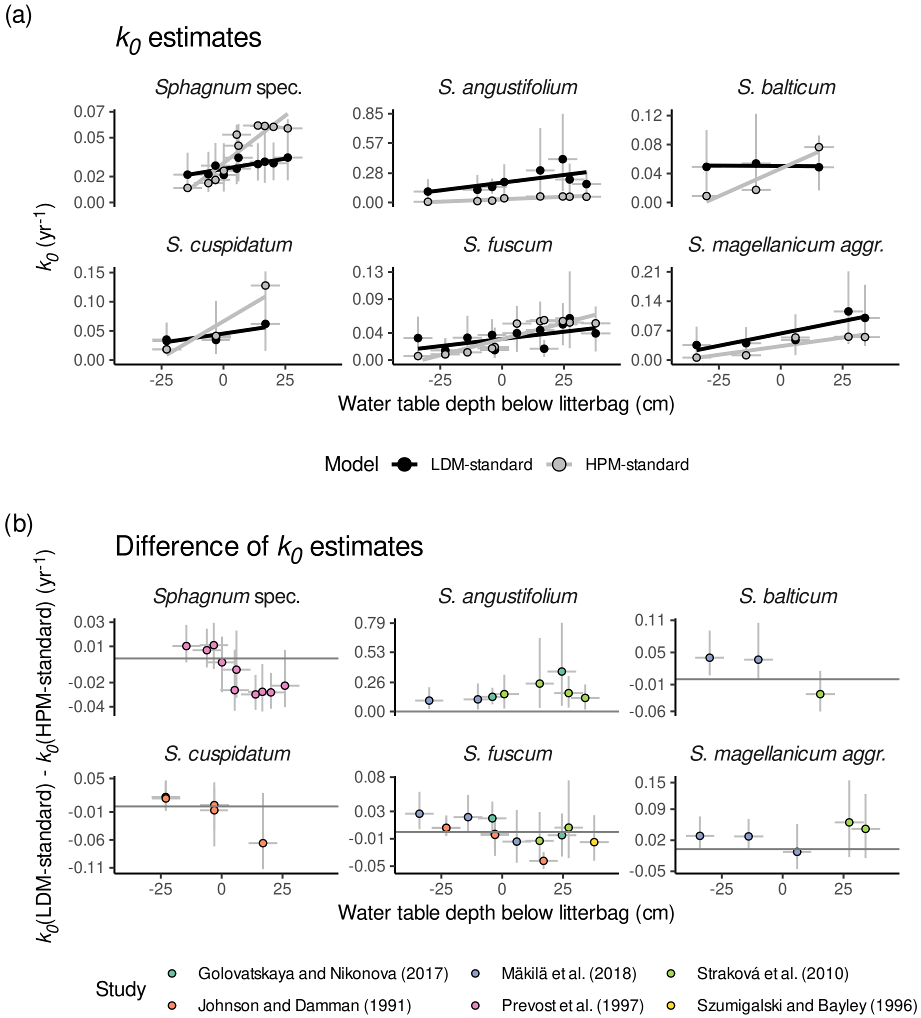

Figure 4. Comparison of k_0 estimates of HPM-standard and LDM-standard (Tab. 3) for species with at least three litterbag experiments. (a) k_0 estimates of HPM-standard (grey) and k_0 estimates of LDM-standard (black) versus reported average water table depths below the litterbags (negative values represent litterbags placed below the water table, positive values represent litterbags placed above the water table in the unsaturated zone). (b) k_0 estimates of LDM-standard minus k_0 estimates of HPM-standard versus reported average water table depths below the litterbags (i.e., the difference of the values shown in (a)). Grey horizontal lines indicate a difference in k_0 of 0 yr^{-1}. Points represent average estimates and error bars 95% posterior intervals. Lines are predictions of linear models fitted to the average estimates. *Sphagnum* spec. are samples that have been identified only to the genus level.

A comparison of k_0 estimates of HPM-standard and the other modifications of the HPM decomposition module suggests that when HPM decomposition module parameters are estimated, larger anaerobic decomposition rates and a less steep decrease of decomposition rates from oxic to anoxic conditions are predicted, similar to LDM-standard. We computed the difference of k_0 predicted by HPM-standard and the other HPM decomposition module versions (Fig. 5). When the HPM decomposition module with standard parameter values is used as prior for the litterbag decomposition module (HPM-peat), it predicts k_0 nearly identical to HPM-standard. In contrast, both model versions where HPM decomposition module parameters were estimated predict larger anaerobic decomposition rates and less of an increase under oxic conditions relative to anoxic conditions than HPM-standard. Thus, the HPM decomposition module with standard parameter values predicts a steeper decrease of decomposition rates from oxic to anoxic conditions and overall smaller anaerobic decomposition rates than LDM-standard and the models that estimate HPM decomposition module parameters from available litterbag data.

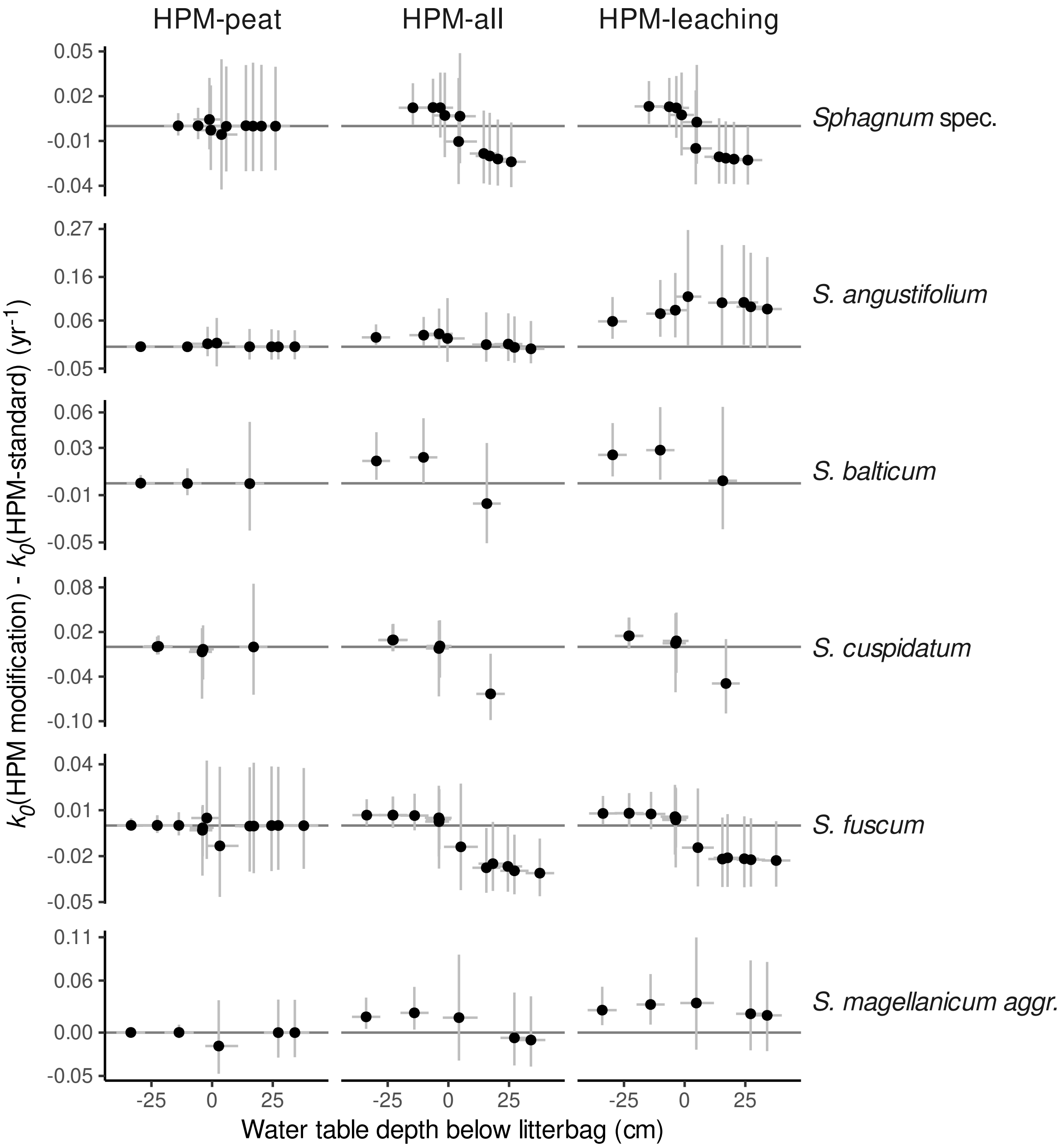

Figure 5. k_0 predicted by HPM decomposition module modifications (either HPM-peat, HPM-all, or HPM-leaching) minus k_0 predicted by the HPM decomposition module with standard parameter values (HPM-standard) versus estimated average water table depths below the litterbags (negative values represent litterbags placed below the water table, positive values represent litterbags placed above the water table in the unsaturated zone). Points represent average estimates and error bars 95% posterior intervals. *Sphagnum* spec. are samples which that been identified only to the genus level. Only data for species with at least three replicates are shown.

3.5 HPM decomposition module parameters that are responsible for the less steep gradient in decomposition rates from oxic to anoxic conditions

To analyze which of the HPM decomposition module parameters ($c_1, W_{opt}, f_{min}, c_2$) cause the less steep gradient in decomposition rates from oxic to anoxic conditions, we conducted a sensitivity analysis, where we made predictions with HPM-leaching for the same species and the same gradient from oxic to anoxic conditions, each time setting one of the four parameters to their standard values (four sets of predictions in total). We then computed the difference of the predicted k_0 values to predictions of HPM-leaching (with no parameter value set to its standard value). This difference is plotted versus the depth of the water table below the litter, as shown in Fig. 6. This analysis suggests that W_{opt} and c_2 cause the less steep gradient in decomposition rates from oxic to anoxic conditions, whereas the other two parameters have no qualitative influence.

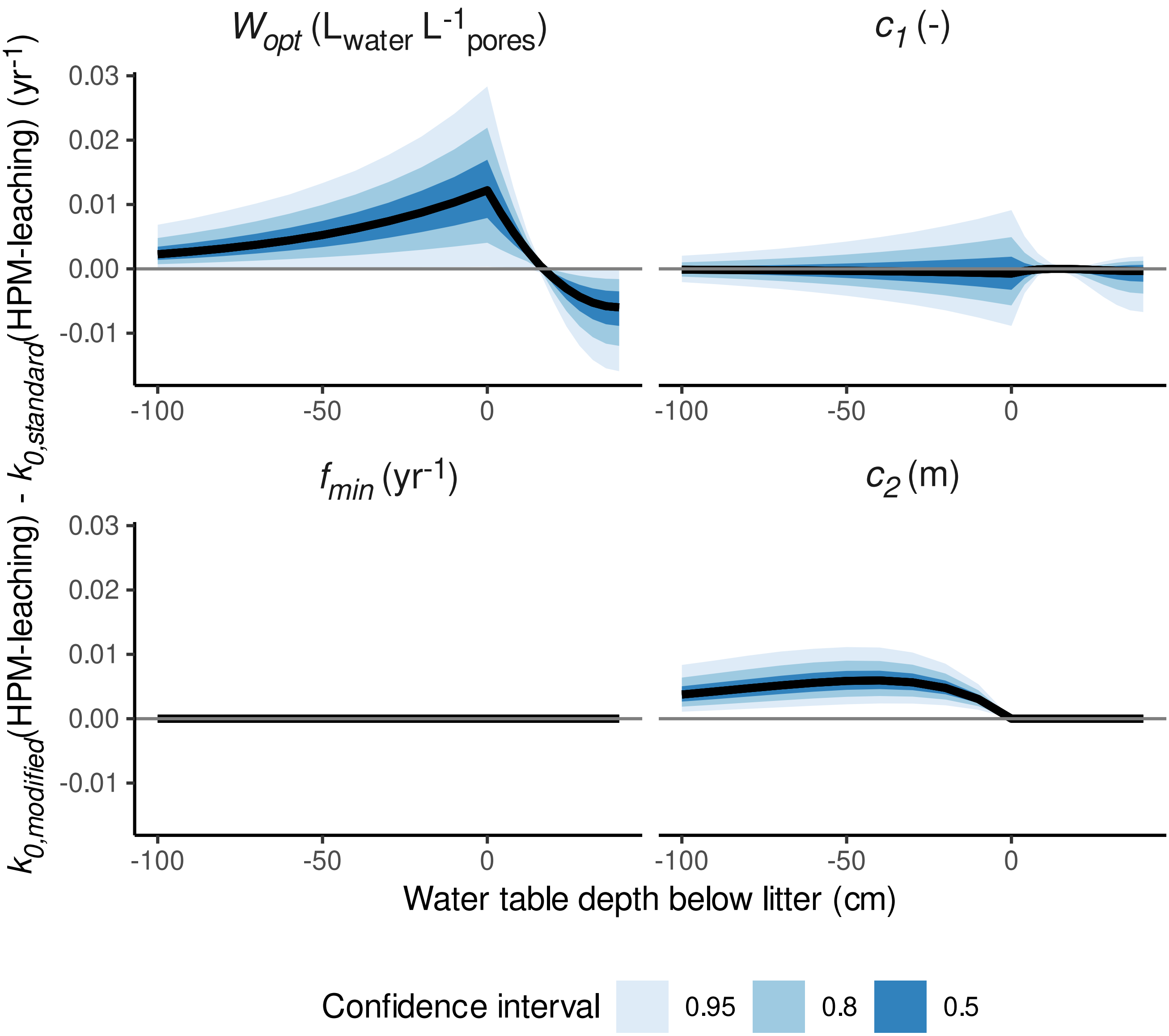

Figure 6. Difference between decomposition rates for *S. fuscum* predicted with parameter values estmated by HPM-leaching ($k_{0,\text{modified}}$(HPM-leaching)), and when setting the HPM decomposition module parameter in the panel title to its standard value ($k_{0,\text{standard}}$(HPM-leaching)), versus the water table depth below the litter (negative values represent litter placed below the water table, positive values represent litter placed above the water table in the unsaturated zone). Panels show results when different parameters are set to their standard values. Positive $k_{0,\text{modified}}(\text{HPM-leaching}) - k_{0,\text{standard}}(\text{HPM-leaching})$ means that decomposition rates are larger when using the estimated parameter value compared to using the standard parameter value. Shaded areas are central confidence intervals with probabilities given in the figure legend.

3.6 Relation of l_0 to the degree of saturation

In model HPM-leaching, we included a logistic regression model that estimates the relation between l_0 and the degree of saturation. The parameter estimates suggest that both positive and negative relations of l_0 to the degree of saturation are compatible with available litterbag data (95% confidence intervals for the slope (logit scale): (-0.28, 0.15)). Thus, available litterbag data do not allow to conclude whether l_0 are positively related to the degree of saturation or not.

4 Discussion

Our aims were to test whether the HPM decomposition module fits decomposition rates estimated from available litterbag experiments, to estimate HPM decomposition module parameters from available litterbag experiments, to understand what factors could cause differences in parameter estimates to the standard values, and to check whether the estimates from litterbag data could imply significant differences in peat accumulation predicted by the HPM compared to the standard parameter values.

The parameter estimates derived from available litterbag data suggest differences in the control of decomposition rates compared to the standard parameter values: the HPM decomposition module with standard parameter values predicts a steeper decrease of decomposition rates from oxic to anoxic conditions and smaller anaerobic decomposition rates for several species than estimated from LDM-standard and the models that estimate HPM decomposition module parameters from available litterbag data. These differences imply larger estimates for W_{opt}, the degree of saturation where decomposition rates are maximal, and c_2, the anoxia scale length (the parameter that controls how strong decomposition rates decrease below the water table depth). We will show here, by comparing parameter estimates to results from sensitivity analyses of the HPM, that the new parameter estimates can cause large differences in long-term peat accumulation predicted by the HPM.

Our analysis suggests that the HPM decomposition module with standard parameter values fits available litterbag data, but our modifications, where W_{opt}, c_2, and (for some species) $k_{0,i}$ estimates significantly differ from the standard values, have equivalent fit. This can be explained by two mechanisms: first, the litterbag decomposition model explains mass loss by initial leaching and decomposition. Thus, remaining masses reported in a litterbag experiment can be fitted either by assuming a larger l_0 and smaller k_0, or by assuming a smaller l_0 and larger k_0. By this first mechanism, the litterbag decomposition model can first estimate k_0 to agree with the HPM decomposition module and then adjust l_0 to fit the remaining masses of the litterbag experiments. The second mechanism is the impact of the design of available *Sphagnum* litterbag experiments on the accuracy of l_0 and k_0 estimates: initial leaching losses can explain mass losses only at the start of the experiment (equation (2)), but decomposition explains a continuous mass loss. It is therefore possible to estimate l_0 and k_0 accurately when remaining masses shortly after the start of the experiment are recorded, but the majority of litterbag experiments collects the first litterbags only after half a year or later (Teickner et al., 2024b). This causes large errors in l_0 and k_0 estimates and therefore allows the model to adjust l_0 and k_0 by the first mechanism, such that all model versions have equivalent fit to remaining masses while also fitting decomposition rates suggested by different HPM decomposition module priors. Improved litterbag experiments are needed for more accurate tests of any peatland decomposition module and for obtaining parameter estimates accurate enough to allow even only approximate predictions of long-term peat accumulation. Applications of the HPM should consider this variability in parameter estimates compatible with available litterbag experiments.

In the next subsections, we first evaluate the reliability of our test. We discuss whether the identified parameter value differences could be an artifact of using heterogeneous litterbag data, and we discuss how compatible the new HPM decomposition module parameter estimates are with other studies that analyzed how decomposition rates differ in dependency of water availability or that estimated c_2 from peat core data. Second, we address the remaining aims: we discuss what factors could cause the larger anaerobic decomposition rates and, in some cases, smaller aerobic decomposition rates estimated by the litterbag decomposition model, and we discuss what implications the differences between estimated and standard parameter values have for peat accumulation predicted by the HPM. Finally, we give recommendations for improving tests of peat decomposition modules.

4.1 Reliability of the identified discrepancies

Before analyzing potential causes of the discrepancies found for c_2 and W_{opt} we first ask if combining different litterbag experiments is reliable evidence for the less steep gradient in decomposition rates from oxic to anoxic conditions.

If we take a look at the misfits of the standard HPM decomposition module (HPM-standard) shown in Fig. 4, many, but not all underestimations of aerobic decomposition rates could have been caused by other factors: for example for *S. balticum* the difference may have been caused by differences in the two litterbag experiments from which we collected the data because the replicate with positive water table depth is from Straková et al. (2010), whereas the two others are from Mäkilä et al. (2018) (Fig. 4). The less pronounced gradient in measured decomposition rates above the water table depth is, however, also visible for *S. fuscum* replicates within the same study and in addition similar across these (independent) studies (Fig. 4, supporting information S6) (Johnson and Damman, 1991; Golovatskaya and Nikonova, 2017; Mäkilä et al., 2018), indicating that this pattern cannot be explained in all cases by differences between studies. In addition, during the cross-validation, we removed data from individual studies from the model and the remaining subsets still resulted in similar estimates for c_2 and W_{opt} (supporting Fig. S12). Finally, numerous previous studies suggest that water table depth is an important control of decomposition rates (e.g., Blodau et al., 2004) and one may therefore expect that also between different studies decomposition rate differences should be controlled to a large degree by differences in water table depths. Thus, even with the heterogeneous litterbag data currently available, a less steep gradient of decomposition rates from oxic to anoxic conditions appears to be replicable between studies and species. To fully rule out that this pattern may be biased by heterogeneous litterbag data and biases of the litterbag decomposition model, controlled litterbag experiments that systematically estimate decomposition rates along the gradient from oxic to anoxic conditions are needed.

The W_{opt} estimate suggested by HPM-leaching is near the average optimum of heterotrophic respiration estimated across a range of mineral soils (Moyano et al., 2013). The estimate is also in line with a study where the largest decomposition rates of the same litter type were observed at or just above the average water table level in hummocks (Belyea, 1996), and with maximum CO_2 production rates around 13 cm above the water table level in a mesocosm study (Blodau et al., 2004). According to the ModGberg model the degree of saturation at this depth is near the W_{opt} estimate suggested by HPM-all and HPM-leaching. For example, for our simulation analysis used to produce Fig. 6, the average W_{opt} estimated by model HPM-leaching (0.57 L_{water} L_{pores}^{-1}) is reached around 16 cm above the water table level, as shown in Fig. 7. At shallower depths, the degree of saturation decreases below the W_{opt} estimate and this would decrease decomposition rates as observed in Belyea (1996). In contrast, according to the the ModGberg model, a degree of saturation corresponding to the standard W_{opt} value (0.45 L_{water} L_{pores}^{-1}) is reached at shallower depths and in the same simulation with this standard W_{opt} value, no pronounced subsurface peak in decomposition rates is observed (supporting Fig. S15). In hollows, the optimum degree of saturation suggested by HPM-leaching is reached near the surface for either W_{opt} value (supporting Fig. S15). Thus, a larger value for W_{opt} would be compatible with results from several previous studies.

Larger and smaller c_2 than the standard value have been estimated for several peatland cores with the HPM and a modified version with monthly time step (Quillet et al., 2015; Treat et al., 2021, 2022). Smaller values have been estimated for tropical peatlands (Kurnianto et al., 2015). To our knowledge, no litterbag experiment directly estimated c_2. A difficulty is that available litterbag experiments cover only a comparatively small depth range below the water table level (at most around 30 cm, Fig. 4) and therefore gradients in anaerobic decomposition rates across larger depths below the water table currently cannot be estimated with available litterbag data.

The estimates for the maximum possible decomposition rate ($k_{0,i}$) have large errors and removal of data during the cross-validation caused larger relative differences in $k_{0,i}$ estimates compared to W_{opt} and c_2 (supporting Fig. S12). On the one hand, this variability indicates that available litterbag data are not sufficient to estimate $k_{0,i}$ accurately and that our assignment of *Sphagnum* species to HPM PFT may not be optimal, but on the other hand, this variability may also indicate that categorizing *Sphagnum* species into three PFT may not accurately describe the variability of maximum possible decomposition rates. Several studies suggest that diverse aspects of litter chemistry may increase $k_{0,i}$ (Turetsky et al., 2008; Bengtsson et al., 2018). However, we are not aware of studies that systematically analyze what factors control $k_{0,i}$ within the same species.

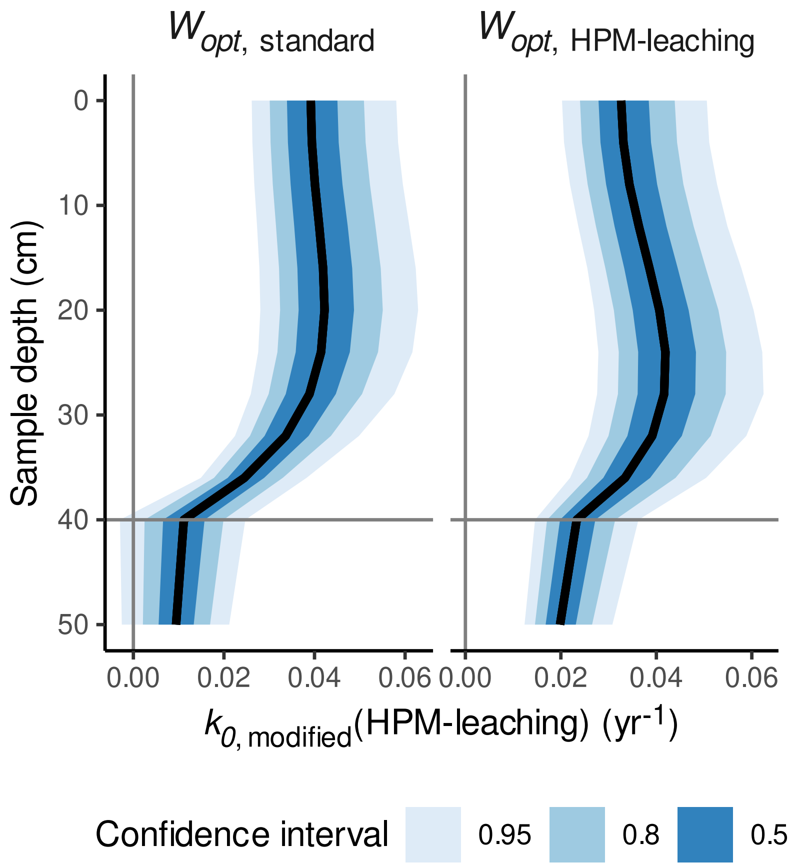

Figure 7. Decomposition rates predicted with HPM-leaching ($k_{0,\mathrm{modified}}$(HPM-leaching)) for *S. fuscum* (hummocks), using either the standard value for W_{opt} or the W_{opt} value estimated by HPM-leaching versus depth of the litter below the peat surface. The horizontal line is the average water table depth. Shaded areas are central confidence intervals with probabilities given in the figure legend.

4.2 Water table fluctuations may explain the discrepancies in c_2 and W_{opt} and larger anaerobic and smaller aerobic decomposition rates.

The HPM decomposition module predicts decomposition rates based on average annual water table depths (Frolking et al., 2010) and ignores water table fluctuations. Our evaluation of the HPM decomposition module also assumed an average water table depth during the litterbag experiments and the HPM decomposition module translated this into a clear pronounced transition between anaerobic and aerobic decomposition rates (Fig. 4). In reality, water table levels fluctuate and this causes transient and nonlinear changes in decomposition rates due to variations in the availability of oxygen and other electron acceptors, flushing of end products of anaerobic decomposition, and possibly other factors (Siegel et al., 1995; Blodau and Moore, 2003; Blodau et al., 2004; Beer and Blodau, 2007; Knorr and Blodau, 2009; Walpen et al., 2018; Campeau et al., 2021; Kim et al., 2021; Treat et al., 2022; Obradović et al., 2023). A possible explanation why the gradient in decomposition rates from oxic to anoxic decomposition is less steep across litterbag experiments, on average, than suggested by the standard HPM decomposition module could therefore be that an averaging effect of fluctuating water table levels on both aerobic and anaerobic decomposition rates is neglected by the HPM decomposition module. An additional factor may be that litterbags cover a depth range and therefore the decompsition rate estimate is an average over the depth covered by the litterbag. If moisture conditions vary over this depth, the decomposition rate estimate also averages over moisture conditions, with similar effects as the temporal average caused by water table fluctuations.

If this is the case, c_2 would have to be re-interpreted as transition parameter that accounts for both limitation of anaerobic decomposition under anoxic conditions and the effects of periodically oxic conditions. Similarly, W_{opt} would have to be re-interpreted as the optimum average degree of saturation for decomposition under water table level variations and its value would be necessarily different from the optimum degree of saturation for depolymerization under static degree of saturation.

Adjusting the HPM decomposition module parameters as implied by our modified models may be an easy way to account for the effect of sub-annual variation in water table levels on decomposition rates, if the discrepancies are caused by fluctuating water tables and if the model is representative for different effects variations in water table level may have on decomposition rates (e.g., short-term fluctuations compared to seasonal water table variations compared to prolonged droughts). What we have not considered due to limited data is that c_2 can be expected to depend on long-term changes in groundwater flow (e.g., Siegel et al., 1995) or site-specific differences in hydrology and other factors (e.g., Frolking et al., 2010; Treat et al., 2021, 2022). Therefore, c_2 can be expected to differ between litterbag studies and our data only indicate that c_2 is larger on average, whereas more research is necessary to estimate and understand site-specific controls of c_2 and how a change in hydrology controls c_2. Similarly, W_{opt} may differ between sites and over time. It would be interesting to know whether litterbag experiments can quantify these controls and whether c_2 estimated from litterbag experiments is generally larger in peatlands with larger water table fluctuations.

It is also worth mentioning that a modification of the HPM, HPM-Arctic (Treat et al., 2021), has a seasonally dynamic WTD and this modification may account for at least a part of the discrepancies we observed here. Unfortunately, most available litterbag data do not report WTD at sufficient temporal resolution to test whether standard HPM parameter values are more compatible with litterbag data when such seasonal variations in WTD are considered.

4.3 Implications of the discrepancies in W_{opt}, c_2, and $k_{0,i}$ for long-term C accumulation

A larger c_2 implies larger anaerobic decomposition and may thus indicate that the HPM decomposition module underestimates anaerobic decomposition rates. Previous global and local sensitivity analyses, where HPM parameter values were varied in broad ranges and environmental conditions were varied, identified c_2 as influential for C accumulation in the HPM (Quillet et al., 2013a, b).

If c_2 is varied within the range from the standard value (0.3 m) to the average posterior estimate from HPM-leaching (0.64 m), this would cause differences in predicted C accumulation of a maximum of ca. 20% in the sensitivity experiment of Quillet et al. (2013a) (depending on precipitation, Fig. 1 c in Quillet et al. (2013a)). If values are changed across the complete posterior range compatible with litterbag data and if other HPM parameters would also be varied, the effect would be even larger (Fig. 2 c in Quillet et al. (2013a)).

Due to parameter interactions and feedbacks, an increase in anaerobic decomposition rates can result in smaller or larger C accumulation of the HPM, depending on environmental conditions (Quillet et al., 2013a). Small anaerobic decomposition may cause too rapid C accumulation resulting in a low water table level, a thick aerobic zone, and thus smaller overall C accumulation after a longer time. Larger anaerobic decomposition may result in higher water table levels and this can increase C accumulation in the long-term. Too large anaerobic decomposition decreases C accumulation (Quillet et al., 2013a).

A larger W_{opt} implies that the largest aerobic decomposition rates are reached under more saturated conditions. W_{opt} has not been identified as influential in a sensitivity analysis of the HPM (Quillet et al., 2013a), but as shown above, it contributes to the less steep decrease of decomposition rates from oxic to anoxic conditions. Importantly, since the HPM does not have a seasonally resolved water table depth, the two sensitivity analyses did not consider how seasonal variations of the water table depth may control long-term C accumulation, and consequently the re-interpreted W_{opt} may be more important to long-term C accumulation than previously assumed. In addition, HPM-leaching suggests an average W_{opt} value of 0.57 $\mathrm{L_{water}\ L_{pores}^{-1}}$, which is outside the range of values tested in Quillet et al. (2013a) (0.3 to 0.5 $\mathrm{L_{water}\ L_{pores}^{-1}}$). This implies that the sensitivity of long-term C accumulation to W_{opt} has been evaluated over a too small range.

A larger $k_{0,i}$ increases decomposition rates for a species and *Sphagnum* $k_{0,i}$ are particularly relevant for many peatlands because the bulk of the peat is *Sphagnum* peat. In the sensitivity analysis in Quillet et al. (2013b), $k_{0,\mathrm{hummock}}$ had large interaction effects with other parameters of the HPM and therefore could either cause larger or smaller peat accumulation, depending on environmental conditions, other parameters, and what vegetation shifts occur in a specific case. Similar to W_{opt}, our $k_{0,i}$ estimates have errors that are larger than the range of values tested in Quillet et al. (2013b). For example, for hummock *Sphagna*, $k_{0,i}$ was varied from 0.04 to 0.06 $\mathrm{yr^{-1}}$, whereas average estimates for $k_{0,i}$ of HPM-leaching for species assigned to the hummock PFT range from 0.04 to 0.19 $\mathrm{yr^{-1}}$. As mentioned above, this range of $k_{0,i}$ estimates may be biased because of the difficulty to assign *Sphagnum* species to HPM PFT, but from a different perspective, this is an additional error source for $k_{0,i}$ estimates that should be considered in sensitivity analyses unless more evidence becomes available to define PFT and their maximum possible decomposition rates.

A further aspect that needs to be considered is that HPM-all and HPM-leaching estimate parameter distributions based on available data, whereas existing studies defined fixed parameter values or ranges of parameter values based on expert knowledge. Based on Quillet et al. (2013a), the uncertainties would cause non-negligible differences in predicted long-term C accumulation. For example, values within the uncertainty range of c_2 estimated by HPM-leaching ((0.4, 0.97), 95% confidence interval), would imply differences up to 100 kg $\mathrm{m^{-2}}$ of accumulated C over 5000 years in some simulations (Fig. 1 (c) in Quillet et al. (2013a), with a maximum total accumulation of ca. 430 $\mathrm{kg_C\ m^{-2}}$). Simulations of remaining masses for different *Sphagnum* species under different conditions also indicate large uncertainties in predicted remaining masses (supporting info S9). This implies that more work is required to estimate parameters accurately enough to detect even relative large differences among peatland models and between model predictions and peat cores.

Summarized, based on existing sensitivity analyses of the HPM the parameter discrepancies suggested by HPM-all and HPM-leaching can translate into non-negligible differences in long-term C accumulation rates. They also imply gaps in previous sensitivity analyses of the HPM, namely that W_{opt} and possivly $k_{0,i}$ (for some species) have been analyzed over a too restricted value range and may play a more important role if water table fluctuations are taken into account.

4.4 How can we improve tests of peatland decomposition modules?

We suggest the following steps to estimate peatland decomposition module parameters more accurately and therefore also to improve the accuracy of tests of peatland decomposition modules:

1. High temporal resolution measurements of WTD: For many available litterbag studies, it is not clear whether reported WTD estimates are unbiased estimates of average WTD (i.e., are derived from high-resolution measurements during the incubation) or biased (due to a too small temporal resolution or coverage). This limitation could be reduced by reporting high temporal resolution WTD measurements along litterbag experiments. Such data are also necessary to investigate whether HPM decomposition module parameters are controlled by WTD fluctuations.

2. Eliminate the need of auxiliary models to estimate the degree of saturation: There is a lack of data on the degree of saturation (or porosity and volumetric water content, from which the degree of saturation could be computed) for available litterbag experiments. For this reason, we used the modified Granberg model to estimate the degree of saturation based on reported WTD and an assumed peat porosity. The modified Granberg model, reported WTD, and our assumed peat porosity are error sources for our test. This limitation could be reduced by measurements of peat porosity and high temporal resolution measurements of volumetric water content during litterbag experiments.

3. Implementing a standard for how to assign *Sphagnum* species to model PFT: The HPM does not specify how to assign *Sphagnum* species to PFT (Frolking et al., 2010), which makes it difficult to compare litterbag experiments to parameters for HPM PFT. Ideally, peatland models should provide lists of species they assign to certain PFT to facilitate tests. Moreover, available niche data used here to assign species to PFT may be biased by short term measurements during summer that are not in line with average niches defined in peatland models, similar to how transfer model for testate amoebae are suggested to be biased (Swindles et al., 2015).

4. Decreasing errors in k_0 and l_0 estimates from litterbag experiments: Our analysis suggests that a comparatively large range of c_2, W_{opt}, and $k_{0,i}$ estimates in the HPM decomposition module are compatible with available litterbag data because errors in remaining masses are large enough to support a range of k_0 and l_0 estimates and because of deficiencies in the design of the litterbag experiments. As a consequence, k_0 estimates of the litterbag decomposition model can be adjusted to fit predictions of the HPM decomposition module for a range of HPM decomposition module parameter values. We also assume that because of these large errors and a large variability of initial leaching losses due to differences in litter handling (Teickner et al., 2024b), we could not detect an expected positive relation of l_0 to the degree of saturation

(Lind et al., 2022). Future litterbag experiments that aim to improve peatland models should reduce errors of k_0 and l_0 estimates (e.g., Teickner et al., 2024b).

5. Systematic litterbag experiments along the gradient from oxic to anoxic conditions: There are few litterbag experiments available that systematically analyze how decomposition rates differ along the gradient from oxic to anoxic conditions. Problems are that many studies test only few conditions and do not cover depth ranges large enough to estimate the minimum decomposition rate (f_{min}) and c_2. An ideal study would use litter material of the same species and origin (thus making sure $k_{0,i}$ would be the same for all replicates) and systematically record remaining masses under different degrees of saturation in the same peat material to accurately estimate W_{opt} and c_1. Another ideal study would systematically record remaining masses at many depth levels, and deeper than 30 cm below the average annual WTD to allow accurate estimation of c_2. Similar experiments could be used to estimate how WTD fluctuations affect decomposition rates along the gradient from oxic to anoxic conditions and how this would change estimates for W_{opt} and c_2.

6. Understanding the controls of $k_{0,i}$: Values of $k_{0,i}$ can be assumed to be controlled, among other factors, by litter chemistry. Even though there are studies that analyze how litter chemistry controls decomposition rates (e.g., Turetsky et al., 2008), there are few that do this systematically (e.g., Bengtsson et al., 2018) and these do not consider initial leaching losses and thus may confound initial leaching and decomposition, both of which may depend on initial litter chemistry. Studies that systematically change litter chemistry within species would be required to estimate $k_{0,i}$. These estimates would also be useful to define PFT for decomposition modules.

7. Understanding how c_2 and W_{opt} vary between sites and in dependency of peat characteristics: Too few litterbag experiments with too few replicates are available to estimate c_2 and W_{opt} separately for individual sites (or how they may vary over time). Systematic litterbag experiments are needed to estimate how environmental conditions control the magnitude of these parameters, for example due to temporal variations in water and oxygen availability or differences in availability of alternative electron acceptors under anoxic conditions.

Systematic and high-quality litterbag experiments that are designed specifically to test peatland decomposition modules are required to achieve these improvements. To support the design of such experiments, we created an R package (hpmdpredict, supporting information S10) that allows to make predictions with HPM-leaching for hypothetical litterbag experiments and that also allows to change parameter values (Teickner and Knorr, 2024a). This could for example be useful to estimate the sample sizes that are required to detect specific differences in remaining masses, to test to what extent litterbag experiments are compatible with HPM-leaching, or to analyze the effect of changing HPM decomposition module parameter values from the standard values or our estimates.

5 Conclusions

Based on the litterbag data, the degree of saturation where decomposition is largest (W_{opt}) and the anoxia scale length (c_2, controls how fast decomposition rates decrease below the average annual WTD) are significantly larger than the standard parameter values. Moreover, maximum possible decomposition rates ($k_{0,i}$) for individual species are overall more variable than implied by the standard HPM decomposition module parameter values. According to previous sensitivity analyses, these parameter estimates imply differences in predicted C accumulation rates of up to 100 $\mathrm{kg_C\ m^{-2}}$ over 5000 years (with a maximum total C accumulation of ca. 430 $\mathrm{kg_C\ m^{-2}}$) when compared to the standard parameter values. The differences in HPM parameter estimates imply larger anaerobic decomposition rates for several species and a less steep gradient of decomposition rates from oxic to anoxic conditions. This pattern may be caused by water table fluctuations, differences in groundwater flow, or spatial averaging in litterbag experiments; factors that are currently not explicitly considered both in the HPM decomposition module and available litterbag experiments.

Our analysis suggests that the HPM decomposition module with standard parameter values fits available *Sphagnum* litterbag data, but model versions where HPM decomposition module parameters were estimated from available litterbag data have an equivalent fit. This is caused by two mechanisms: First, remaining masses in litterbag experiments can be explained by initial leaching losses and decomposition. If remaining masses are reported only some time after the initial leaching loss has happened, they can be explained either by small initial leaching losses and a large decomposition rate or by large initial leaching losses and a smaller decomposition rate. Second, the majority of available *Sphagnum* litterbag experiments reports remaining masses only a long time after the initial leaching loss happened. Taken together, this means that available litterbag data are compatible with a broad range of decomposition rates suggested by HPM decomposition module versions with large differences in parameter values. Improved litterbag experiments are needed for more accurate tests of any peatland decomposition module and for obtaining parameter estimates accurate enough to allow even only approximate predictions of long-term peat accumulation. Applications of the HPM and any other peatland model that relies on litterbag data to parameterize its decomposition process should consider that a broad range of decomposition module parameter values is compatible with available litterbag experiments.

The modeling approach used here can be combined with different data sources and peatland decomposition modules and therefore may serve as blueprint for future tests and to obtain more accurate parameter estimates once improved litterbag experiments are available. In light of the large differences in long-term peat accumulation suggested by the parameter estimates, we conclude that it is worth to conduct such litterbag experiments, not only to improve the decomposition module of the HPM, but to improve peatland models in general.

. Data and code to reproduce this manuscript are available from Teickner et al. (2024a). The data used in this study are derived from Teickner and Knorr (2024b). An R package to make predictions for litterbag experiments with model HPM-leaching is available from Teickner and Knorr (2024a).

. HT: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing - original draft, visualization, project administration. EP: supervision, funding acquisition, writing - review & editing. KHK: supervision, funding acquisition, writing - review & editing.

. The authors declare no competing interests.

. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant no. KN 929/23-1 to Klaus-Holger Knorr and grant no. PE 1632/18-1 to Edzer Pebesma.

References

Asada, T. and Warner, B. G.: Surface Peat Mass and Carbon Balance in a Hypermaritime Peatland, Soil Science Society of America Journal, 69, 549–562, https://doi.org/10.2136/sssaj2005.0549, 2005.

Bartsch, I. and Moore, T. R.: A Preliminary Investigation of Primary Production and Decomposition in Four Peatlands near Schefferville, Québec, Canadian Journal of Botany, 63, 1241–1248, https://doi.org/10.1139/b85-171, 1985.

Bauer, I. E.: Modelling Effects of Litter Quality and Environment on Peat Accumulation over Different Time-Scales: Peat Accumulation over Different Time-Scales, Journal of Ecology, 92, 661–674, https://doi.org/10.1111/j.0022-0477.2004.00905.x, 2004.

Beer, J. and Blodau, C.: Transport and Thermodynamics Constrain Belowground Carbon Turnover in a Northern Peatland, Geochimica et Cosmochimica Acta, 71, 2989–3002, https://doi.org/10.1016/j.gca.2007.03.010, 2007.

Belyea, L. R.: Separating the Effects of Litter Quality and Microenvironment on Decomposition Rates in a Patterned Peatland, Oikos, 77, 529–539, https://doi.org/10.2307/3545942, 1996.

Bengtsson, F., Rydin, H., and Hájek, T.: Biochemical Determinants of Litter Quality in 15 Species of *Sphagnum*, Plant and Soil, 425, 161–176, https://doi.org/10.1007/s11104-018-3579-8, 2018.

Blodau, C. and Moore, T. R.: Experimental Response of Peatland Carbon Dynamics to a Water Table Fluctuation, Aquatic Sciences - Research Across Boundaries, 65, 47–62, https://doi.org/10.1007/s000270300004, 2003.

Blodau, C., Basiliko, N., and Moore, T. R.: Carbon Turnover in Peatland Mesocosms Exposed to Different Water Table Levels, Biogeochemistry, 67, 331–351, https://doi.org/10.1023/B:BIOG.0000015788.30164.e2, 2004.

Bona, K. A., Hilger, A., Burgess, M., Wozney, N., and Shaw, C.: A Peatland Productivity and Decomposition Parameter Database, Ecology, 99, 2406–2406, https://doi.org/10.1002/ecy.2462, 2018.

Bona, K. A., Shaw, C., Thompson, D. K., Hararuk, O., Webster, K., Zhang, G., Voicu, M., and Kurz, W. A.: The Canadian Model for Peatlands (CaMP): A Peatland Carbon Model for National Greenhouse Gas Reporting, Ecological Modelling, 431, 109 164, https://doi.org/10.1016/j.ecolmodel.2020.109164, 2020.

Breeuwer, A., Heijmans, M., Robroek, B. J. M., Limpens, J., and Berendse, F.: The Effect of Increased Temperature and Nitrogen Deposition on Decomposition in Bogs, Oikos, 117, 1258–1268, https://doi.org/10.1111/j.0030-1299.2008.16518.x, 2008.

Campeau, A., Vachon, D., Bishop, K., Nilsson, M. B., and Wallin, M. B.: Autumn Destabilization of Deep Porewater CO_2 Store in a Northern Peatland Driven by Turbulent Diffusion, Nature Communications, 12, 6857, https://doi.org/10.1038/s41467-021-27059-0, 2021.

Chaudhary, N., Miller, P. A., and Smith, B.: Biotic and Abiotic Drivers of Peatland Growth and Microtopography: A Model Demonstration, Ecosystems, 21, 1196–1214, https://doi.org/10.1007/s10021-017-0213-1, 2018.

Clymo, R. S., Turunen, J., and Tolonen, K.: Carbon Accumulation in Peatland, Oikos, 81, 368–388, https://doi.org/10.2307/3547057, 1998.

Frolking, S., Roulet, N. T., Moore, T. R., Richard, P. J. H., Lavoie, M., and Muller, S. D.: Modeling Northern Peatland Decomposition and Peat Accumulation, Ecosystems, 4, 479–498, https://doi.org/10.1007/s10021-001-0105-1, 2001.

Frolking, S., Roulet, N. T., Tuittila, E., Bubier, J. L., Quillet, A., Talbot, J., and Richard, P. J. H.: A New Model of Holocene Peatland Net Primary Production, Decomposition, Water Balance, and Peat Accumulation, Earth System Dynamics, 1, 1–21, https://doi.org/10.5194/esd-1-1-2010, 2010.

Frolking, S., Talbot, J., Jones, M. C., Treat, C. C., Kauffman, J. B., Tuittila, E.-S., and Roulet, N.: Peatlands in the Earth's 21st Century Climate System, Environmental Reviews, 19, 371–396, https://doi.org/10.1139/a11-014, 2011.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. B.: Bayesian Data Analysis, Chapman & Hall/CRC Texts in Statistical Science, CRC Press, Boca Raton, third edition edn., ISBN 978-1-4398-4095-5, 2014.

Golovatskaya, E. A. and Nikonova, L. G.: The Influence of the Bog Water Level on the Transformation of Sphagnum Mosses in Peat Soils of Oligotrophic Bogs, Eurasian Soil Science, 50, 580–588, https://doi.org/10.1134/S1064229317030036, 2017.

Granberg, G., Grip, H., Löfvenius, M. O., Sundh, I., Svensson, B. H., and Nilsson, M.: A Simple Model for Simulation of Water Content, Soil Frost, and Soil Temperatures in Boreal Mixed Mires, Water Resources Research, 35, 3771–3782, https://doi.org/10.1029/1999WR900216, 1999.

Hagemann, U. and Moroni, M. T.: Moss and Lichen Decomposition in Old-Growth and Harvested High-Boreal Forests Estimated Using the Litterbag and Minicontainer Methods, Soil Biology and Biochemistry, 87, 10–24, https://doi.org/10.1016/j.soilbio.2015.04.002, 2015.

Hassel, K., Kyrkjeeide, M. O., Yousefi, N., Prestø, T., Stenøien, H. K., Shaw, J. A., and Flatberg, K. I.: *Sphagnum Divinum* (*Sp. Nov.*) and *S. Medium* Limpr. and Their Relationship to *S. Magellanicum* Brid., Journal of Bryology, 40, 197–222, https://doi.org/10.1080/03736687.2018.1474424, 2018.

Hayward, P. M. and Clymo, R. S.: Profiles of Water Content and Pore Size in *Sphagnum* and Peat, and Their Relation to Peat Bog Ecology, Proceedings of the Royal Society of London. Series B, Biological Sciences, 215, 299–325, 1982.

Heijmans, M. M., Mauquoy, D., Van Geel, B., and Berendse, F.: Long-term Effects of Climate Change on Vegetation and Carbon Dynamics in Peat Bogs, Journal of Vegetation Science, 19, 307–320, https://doi.org/10.3170/2008-8-18368, 2008.

Heinemeyer, A., Croft, S., Garnett, M. H., Gloor, E., Holden, J., Lomas, M. R., and Ineson, P.: The MILLENNIA Peat Cohort Model: Predicting Past, Present and Future Soil Carbon Budgets and Fluxes under Changing Climates in Peatlands, Climate Research, 45, 207–226, https://doi.org/10.3354/cr00928, 2010.

Hoffman, M. D. and Gelman, A.: The No-U-turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, Journal of Machine Learning Research, 15, 1593–1623, 2014.

Johnson, L. C. and Damman, A. W. H.: Species-Controlled *Sphagnum* Decay on a South Swedish Raised Bog, Oikos, 61, 234, https://doi.org/10.2307/3545341, 1991.

Johnson, M. G., Granath, G., Tahvanainen, T., Pouliot, R., Stenøien, H. K., Rochefort, L., Rydin, H., and Shaw, A. J.: Evolution of Niche Preference in *Sphagnum* Peat Mosses, Evolution, 69, 90–103, https://doi.org/10.1111/evo.12547, 2015.

Kallioinen, N., Paananen, T., Bürkner, P.-C., and Vehtari, A.: Detecting and Diagnosing Prior and Likelihood Sensitivity with Power-Scaling, Statistics and Computing, 34, 57, https://doi.org/10.1007/s11222-023-10366-5, 2024.

Kettridge, N. and Baird, A.: In Situ Measurements of the Thermal Properties of a Northern Peatland: Implications for Peatland Temperature Models, Journal of Geophysical Research: Earth Surface, 112, https://doi.org/10.1029/2006JF000655, 2007.

Kim, J., Rochefort, L., Hogue-Hugron, S., Alqulaiti, Z., Dunn, C., Pouliot, R., Jones, T. G., Freeman, C., and Kang, H.: Water Table Fluctuation in Peatlands Facilitates Fungal Proliferation, Impedes *Sphagnum* Growth and Accelerates Decomposition, Frontiers in Earth Science, 8, 9, 2021.

Knorr, K.-H. and Blodau, C.: Impact of Experimental Drought and Rewetting on Redox Transformations and Methanogenesis in Mesocosms of a Northern Fen Soil, Soil Biology and Biochemistry, 41, 1187–1198, https://doi.org/10.1016/j.soilbio.2009.02.030, 2009.

Kurnianto, S., Warren, M., Talbot, J., Kauffman, B., Murdiyarso, D., and Frolking, S.: Carbon Accumulation of Tropical Peatlands over Millennia: A Modeling Approach, Global Change Biology, 21, 431–444, https://doi.org/10.1111/gcb.12672, 2015.

Lind, L., Harbicht, A., Bergman, E., Edwartz, J., and Eckstein, R. L.: Effects of Initial Leaching for Estimates of Mass Loss and Microbial Decomposition—Call for an Increased Nuance, Ecology and Evolution, 12, https://doi.org/10.1002/ece3.9118, 2022.

Liu, H. and Lennartz, B.: Hydraulic Properties of Peat Soils along a Bulk Density Gradient-A Meta Study, Hydrological Processes, 33, 101–114, https://doi.org/10.1002/hyp.13314, 2019.

Loisel, J., van Bellen, S., Pelletier, L., Talbot, J., Hugelius, G., Karran, D., Yu, Z., Nichols, J., and Holmquist, J.: Insights and Issues with Estimating Northern Peatland Carbon Stocks and Fluxes since the Last Glacial Maximum, Earth-Science Reviews, 165, 59–80, https://doi.org/10.1016/j.earscirev.2016.12.001, 2017.

Mäkilä, M., Säävuori, H., Grundström, A., and Suomi, T.: *Sphagnum* Decay Patterns and Bog Microtopography in South-Eastern Finland, Mires and Peat, pp. 1–12, https://doi.org/10.19189/MaP.2017.OMB.283, 2018.

Morris, P. J., Baird, A. J., and Belyea, L. R.: The DigiBog Peatland Development Model 2: Ecohydrological Simulations in 2D, Ecohydrology, 5, 256–268, https://doi.org/10.1002/eco.229, 2012.

Moyano, F. E., Manzoni, S., and Chenu, C.: Responses of Soil Heterotrophic Respiration to Moisture Availability: An Exploration of Processes and Models, Soil Biology and Biochemistry, 59, 72–85, https://doi.org/10.1016/j.soilbio.2013.01.002, 2013.

Nichols, J. E. and Peteet, D. M.: Rapid Expansion of Northern Peatlands and Doubled Estimate of Carbon Storage, Nature Geoscience, 12, 917–921, https://doi.org/10.1038/s41561-019-0454-z, 2019.

Obradović, N., Joshi, P., Arn, S., Aeppli, M., Schroth, M. H., and Sander, M.: Reoxidation of Reduced Peat Organic Matter by Dissolved Oxygen: Combined Laboratory Column-breakthrough Experiments and In-field Push-pull Tests, Journal of Geophysical Research: Biogeosciences, 128, e2023JG007 640, https://doi.org/10.1029/2023JG007640, 2023.

Prevost, M., Belleau, P., and Plamondon, A. P.: Substrate Conditions in a Treed Peatland: Responses to Drainage, Écoscience, 4, 543–554, https://doi.org/10.1080/11956860.1997.11682434, 1997.

Quillet, A., Frolking, S., Garneau, M., Talbot, J., and Peng, C.: Assessing the Role of Parameter Interactions in the Sensitivity Analysis of a Model of Peatland Dynamics, Ecological Modelling, 248, 30–40, https://doi.org/10.1016/j.ecolmodel.2012.08.023, 2013a.

Quillet, A., Garneau, M., and Frolking, S.: Sobol' Sensitivity Analysis of the Holocene Peat Model: What Drives Carbon Accumulation in Peatlands?, Journal of Geophysical Research: Biogeosciences, 118, 203–214, https://doi.org/10.1029/2012JG002092, 2013b.

Quillet, A., Garneau, M., Van Bellen, S., Frolking, S., and Tuittila, E.-S.: Integration of Palaeohydrological Proxies into a Peatland Model: A New Tool for Palaeoecological Studies, Ecohydrology, 8, 214–229, https://doi.org/10.1002/eco.1501, 2015.

R Core Team: R: A Language and Environment for Statistical Computing, Vienna, Austria, 2022.

Rydin, H., Jeglum, J. K., and Bennett, K. D.: The Biology of Peatlands, Biology of Habitats, Oxford University Press, Oxford, 2nd ed edn., ISBN 978-0-19-960299-5 978-0-19-960300-8, 2013.

Scheffer, R. A., Van Logtestijn, R. S. P., and Verhoeven, J. T. A.: Decomposition of *Carex* and *Sphagnum* Litter in Two Mesotrophic Fens Differing in Dominant Plant Species, Oikos, 92, 44–54, https://doi.org/10.1034/j.1600-0706.2001.920106.x, 2001.

Siegel, D. I., Reeve, A. S., Glaser, P. H., and Romanowicz, E. A.: Climate-Driven Flushing of Pore Water in Peatlands, Nature, 374, 531–533, https://doi.org/10.1038/374531a0, 1995.

Stan Development Team: Stan Modeling Language Users Guide and Reference Manual, 2021a.

Stan Development Team: RStan: The R Interface to Stan, 2021b.

Straková, P., Anttila, J., Spetz, P., Kitunen, V., Tapanila, T., and Laiho, R.: Litter Quality and Its Response to Water Level Drawdown in Boreal Peatlands at Plant Species and Community Level, Plant and Soil, 335, 501–520, https://doi.org/10.1007/s11104-010-0447-6, 2010.

Swindles, G. T., Holden, J., Raby, C. L., Turner, T. E., Blundell, A., Charman, D. J., Menberu, M. W., and Kløve, B.: Testing Peatland Water-Table Depth Transfer Functions Using High-Resolution Hydrological Monitoring Data, Quaternary Science Reviews, 120, 107–117, https://doi.org/10.1016/j.quascirev.2015.04.019, 2015.

Szumigalski, A. R. and Bayley, S. E.: Decomposition along a Bog to Rich Fen Gradient in Central Alberta, Canada, Canadian Journal of Botany, 74, 573–581, https://doi.org/10.1139/b96-073, 1996.

Teickner, H. and Knorr, K.-H.: hpmdpredict: Predictions with Model HPMe-LE-peat-l0 from Teickner et al. (2024), Zenodo, https://doi.org/10.5281/ZENODO.11339733, 2024a.

Teickner, H. and Knorr, K.-H.: Peatland Decomposition Database (1.0.0), https://doi.org/10.5281/ZENODO.11276065, 2024b.

Teickner, H., Pebesma, E., and Knorr, K.-H.: Compendium of R Code and Data for "A Synthesis of *Sphagnum* Litterbag Experiments: Initial Leaching Losses Bias Decomposition Rate Estimates" and "Underestimation of Anaerobic Decomposition Rates in *Sphagnum* Litterbag Experiments by the Holocene Peatland Model Depends on Initial Leaching Losses", Zenodo, https://doi.org/10.5281/ZENODO.11472955, 2024a.

Teickner, H., Pebesma, E., and Knorr, K.-H.: A Synthesis of *Sphagnum* Litterbag Experiments: Initial Leaching Losses Bias Decomposition Rate Estimates, https://doi.org/10.5194/egusphere-2024-1686, 2024b.

Thormann, M. N., Bayley, S. E., and Currah, R. S.: Comparison of Decomposition of Belowground and Aboveground Plant Litters in Peatlands of Boreal Alberta, Canada, Canadian Journal of Botany, 79, 9–22, https://doi.org/10.1139/b00-138, 2001.

Treat, C. C., Jones, M. C., Alder, J., Sannel, A. B. K., Camill, P., and Frolking, S.: Predicted Vulnerability of Carbon in Permafrost Peatlands with Future Climate Change and Permafrost Thaw in Western Canada, Journal of Geophysical Research: Biogeosciences, 126, e2020JG005 872, https://doi.org/10.1029/2020JG005872, 2021.

Treat, C. C., Jones, M. C., Alder, J., and Frolking, S.: Hydrologic Controls on Peat Permafrost and Carbon Processes: New Insights from Past and Future Modeling, Frontiers in Environmental Science, 10, 892 925, https://doi.org/10.3389/fenvs.2022.892925, 2022.

Trinder, C. J., Johnson, D., and Artz, R. R.: Interactions among Fungal Community Structure, Litter Decomposition and Depth of Water Table in a Cutover Peatland, FEMS Microbiology Ecology, 64, 433–448, https://doi.org/10.1111/j.1574-6941.2008.00487.x, 2008.

Tuittila, E.-S., Juutinen, S., Frolking, S., Väliranta, M., Laine, A. M., Miettinen, A., Seväkivi, M.-L., Quillet, A., and Merilä, P.: Wetland Chronosequence as a Model of Peatland Development: Vegetation Succession, Peat and Carbon Accumulation, The Holocene, 23, 25–35, https://doi.org/10.1177/0959683612450197, 2013.

Turetsky, M. R., Crow, S. E., Evans, R. J., Vitt, D. H., and Wieder, R. K.: Trade-Offs in Resource Allocation among Moss Species Control Decomposition in Boreal Peatlands, Journal of Ecology, 96, 1297–1305, https://doi.org/10.1111/j.1365-2745.2008.01438.x, 2008.

van Breemen, N.: How *Sphagnum* Bogs down Other Plants, p. 6, 1995.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C.: Rank-Normalization, Folding, and Localization: An Improved R^ for Assessing Convergence of MCMC (with Discussion), Bayesian Analysis, 16, https://doi.org/10.1214/20-BA1221, 2021.

Vitt, D. H.: Growth and Production Dynamics of Boreal Mosses over Climatic, Chemical and Topographic Gradients, Botanical Journal of the Linnean Society, 104, 35–59, https://doi.org/10.1111/j.1095-8339.1990.tb02210.x, 1990.

Walpen, N., Lau, M. P., Fiskal, A., Getzinger, G. J., Meyer, S. A., Nelson, T. F., Lever, M. A., Schroth, M. H., and Sander, M.: Oxidation of Reduced Peat Particulate Organic Matter by Dissolved Oxygen: Quantification of Apparent Rate Constants in the Field, Environmental Science & Technology, 52, 11 151–11 160, https://doi.org/10.1021/acs.est.8b03419, 2018.

Yu, Z., Turetsky, M., Campbell, I., and Vitt, D.: Modelling Long-Term Peatland Dynamics. II. Processes and Rates as Inferred from Litter and Peat-Core Data, Ecological Modelling, 145, 159–173, https://doi.org/10.1016/S0304-3800(01)00387-8, 2001.

Yu, Z. C.: Northern Peatland Carbon Stocks and Dynamics: A Review, Biogeosciences, 9, 4071–4085, https://doi.org/10.5194/bg-9-4071-2012, 2012.

Zhao, B., Zhuang, Q., Treat, C., and Frolking, S.: A Model Intercomparison Analysis for Controls on C Accumulation in North American Peatlands, Journal of Geophysical Research: Biogeosciences, https://doi.org/10.1029/2021JG006762, 2022.