Dear Roslyn Henry and Reviewers,

Thank you for your interest in our work and the opportunity to submit a revised draft of our manuscript. We appreciate the time and effort that you and the reviewers have dedicated to providing valuable feedback on our manuscript– We tried our best to address all comments, which hopefully improved the manuscript in a way that it is suitable for publication in Geoscientific Model Development. Here is a list of the main changes:

- Strengthening the focus as a model description and evaluation paper by expanding the model calibration and validation. In the revised version, we conducted a more thorough calibration and validation using the Biodiversity Exploratories dataset. A second dataset was added for the calibration and validation of intra-annual biomass dynamics. We added a local sensitivity analysis of all parameters after the calibration. A detailed discussion on the model evaluation was added. We removed the simple case study to maintain a clear focus on model evaluation.
- We elaborated on the reasons for building a new model by providing a qualitative comparison with existing grassland models. We hope that thereby the contribution to the field becomes clearer.
- Added a detailed discussion of potential applications, strengths and limitations of our model.
- Keeping the model description concise: We removed the description of simulation as a spatial grid, and we removed the inferior second option for light competition

In the following, we list our specific responses to each comment. Comments are in black, responses are coloured in blue. We hope these changes made the paper a more complete and valuable contribution.

Best regards,
Felix Nößler on behalf of all co-authors

Editor

Dear Felix,

Thank you for your email. While your paper is indeed a model description paper, it is important to demonstrate that the model functions effectively and represents a meaningful advancement in the field.

I tend to agree with the reviewers on this point. Although your model is well documented and supplemented by tutorials elsewhere, I strongly encourage you to expand the validation section. For instance, on line 556, you mention that your calibration/validation exercise highlights the challenges of achieving good results. This raises some concerns-does this suggests that results at other sites might be even less reliable? If so, could this undermine the model's overall applicability?

A well-crafted discussion addressing this concern might suffice, but I believe additional validation would significantly strengthen the paper. The guidelines for Geoscientific Model Development (GMD) explicitly state that 'sufficient verification and evaluation must be included to show that the model is fit for purpose and works as expected' (see bullet point #6: GMD manuscript types). In your response to referee #1, you mentioned a planned follow-up paper with a more extensive calibration and validation exercise. Could you incorporate some of those results here? Doing so would provide a stronger case that the model is robust and reliable.

Thank you for your response. We addressed the concerns by extending the calibration and validation section. We demonstrate that the model can reproduce the seasonal dynamics of one grass species across multiple sites in Europe. Moreover, we strengthened the calibration and validation section by including grassland sites from all three regions of the Biodiversity Exploratories dataset in our model calibration and validation. In addition, we added a detailed discussion of the model validation. In our initial submission, we included parts of the model evaluation in the supplementary material (for example, we presented model trajectories for all validation sites of one region of the Biodiversity Exploratories project on a website and in the supplementary material on Zenodo). To improve clarity and visibility, we have now incorporated (in a slightly reduced form) all calibration and validation into the main text or in the appendix.

I believe section 5 requires less attention. Your findings, showing that experimental results align with existing theories on responses to mowing and grazing, are reassuring. However, if necessary to keep the paper concise, you might consider moving this section to supplementary material. This would free up space, alongside referee #2's recommendations for conciseness, to expand on validation, which is arguably more critical.

Best wishes,
Roslyn

Thank you for making this suggestion. We followed this suggestion by removing the simple case study and strengthening the section on model calibration and validation. It ensures that the paper maintains a clear focus on model description and evaluation.

Reviewer 1

This manuscript presents GrasslandTraitSim.jl, a novel process-based model designed to simulate the effects of land use and climate on plant functional composition and ecosystem services in grasslands. Building on existing methods, the model links plant morphological traits with demographic and physiological processes, providing a comprehensive tool for researchers and land managers. As noted ("This article has a strong focus on model description"), the paper provides a detailed description of GrasslandTraitSim.jl, including its technical specifications and features (see Section 3 on "Technical details of the GrasslandTraitSim.jl model"), and seeks to position it as a tool for advancing ecological research. However, the manuscript vacillates between a technical and a scientific focus, leading to some ambiguity regarding its overall contribution. While GrasslandTraitSim.jl shows promise for advancing ecological research, particularly in understanding grassland dynamics, certain aspects of the paper may require further consideration by the journal editor to determine its suitability for publication.

Firstly, although the model is described in great detail, the paper contains only a limited number of case studies. While these examples are useful, they do not sufficiently demonstrate the application of the model to real-world scenarios or its potential for practical decision-making in grassland management. The authors could strengthen the manuscript by providing more robust examples of how the model works in different environments, or how it can inform specific management practices. These additions would be of particular benefit to readers looking for tangible applications.

Thank you for the suggestion to include more case studies. While we see the point of better demonstrating model capabilities, we followed the editor's recommendation to focus the manuscript more tightly on model description and evaluation. Given this, the editor suggested removing the simple case study from the main text to leave more space for a concise model description and model evaluation. As a result, we removed the simple case study previously included and expended the model evaluation section instead. We plan to present more applied case studies in future work.

Second, the paper presents several simplifications in the model that are not fully justified. Understanding the rationale behind these simplifications is crucial, as it will help readers to assess the reliability and limitations of the model. The authors should provide a clearer explanation of why these assumptions were made and how they might affect the results of the simulations. This would help to improve the overall credibility and robustness of the paper.

Thank you for the comment. In the newly formed discussion, we address this concern by discussing why chose to not include more complexity in our model (L. 630 - 633). Moreover, we described why we chose our specific set of functional traits (L. 660 - 669). Furthermore, we added a detailed description of the advantages and disadvantages of using morphological traits as species-specific inputs of our model (see also response to minor comment below, L. 108 - 112, L. 670 - 674). Moreover, we added Table 1 to illustrate the differences between our model and existing grassland models. This makes it clearer which parts we simplified or expanded.

Furthermore, although the paper mentions that the model can be used for grassland management, it does not elaborate on how it could be integrated into real-life management strategies or policies. A more detailed discussion of the practical applications of the model - particularly in terms of its potential impact on conservation or land-use practices - would make the work more relevant to practitioners in the field. Including such practical considerations would help to bridge the gap between model development and its application in real-world scenarios.

Thank you for the suggestion, we added a dedicated discussion section where we discuss the potential of using the models for analysing the influence of grassland management on the plant functional composition and on the plant functional diversity (L. 684 - 695).

Finally, the paper does not include a performance comparison with other similar models, nor does it assess the performance of the model in different settings. A comparative analysis with other existing models would provide useful context for readers and establish the position of GrasslandTraitSim.jl within the current body of ecological modelling tools. Such a comparison would also give readers a better understanding of the relative strengths and weaknesses of the model.

Thanks for the suggestion of a performance comparison with other models. We added the new Table 1 to qualitatively compare our model with similar models (see also extended versions in the appendix). This table highlights and illustrates the main strengths and weaknesses of our model compared to other models. We decided against doing a quantitative performance comparison with other models, as this would include calibrating the other grassland models to the datasets that we used. This is clearly beyond the scope of our current study, in our opinion.

In light of these points, the journal editor may wish to assess the relevance of this paper to the journal, given its strong technical focus on model development and limited exploration of its scientific and practical implications. Addressing these concerns could make the paper a more complete and valuable contribution to the field. The authors may also wish to incorporate the many suggestions made by this reviewer directly into the manuscript, as this may serve to improve its scientific depth and practical value.

Inline comments

L32: For losses of plant biodiversity associated with intensified practices, the authors may refer to meta-analyses like for instance Piseddu et al., 2021 (https://doi.org/10.1007/s13593-021-00722-y).

Thanks for the literature hint, we added the reference (L. 35).

L 37: A few more details on what exactly this shift is and under what conditions it might be desirable.

We added more details on the statement (L. 41 - 43):

"Furthermore, climate change is expected to alter the plant community composition of grasslands, particularly during periods of heat waves and droughts, for example by suppressing dominant species (Luo et al., 2025) and/or favouring plants with drought avoidance strategies (Griffin-Nolan et al., 2019; Schils et al., 2020)."

L 86: "exactly" can be deleted

We removed the word.

L 85: GrasslandTraitSim.jl still relies on morphological traits to represent species. Authors should discuss potential limitations of this approach, such as the role of genetic variation, environmental heterogeneity (e.g. soil conditions) or other non-morphological factors.

We acknowledge the reviewer's point and agree that relying solely on morphological traits in GrasslandTraitSim.jl presents certain limitations. Morphological traits, while useful proxies for ecological strategies, do not capture the full spectrum of factors influencing species dynamics. In particular, genetic variation within species can shape trait expression and adaptive potential, while environmental heterogeneity—such as variation in soil texture, nutrients, or moisture—can modulate trait performance and competitive outcomes. Additionally, physiological and biochemical traits, which are not always directly reflected in morphology, may play critical roles in species interactions and ecosystem functioning. We now briefly mention that this might come with some limitations, which are discussed later:

Introduction (L. 108 - 112):
"While this morphological trait-based approach enables broader species coverage and generality, it also comes with limitations. Morphological traits do not fully capture intra-specific genetic variation or phenotypic plasticity, both of which can be important for species' responses to environmental change. Additionally, environmental heterogeneity—such as soil texture, nutrient availability, and microclimate—may modulate the functional effects of traits in context-dependent ways." (starting from line 104)

Discussion on modelling intra-specific trait variability (L. 670 - 674):
"To some extent, our model can simulate intra-specific trait variability based on the functional representation rather than species identity. In our model, two simulated species can represent one species in the real world that exhibits different traits on different sites. However, this approach is not applicable to plant species whose traits change dynamically depending on variable environmental conditions. Furthermore, our model does not reflect changes in traits during the life stages of plant species."

L86 : The transition from the general overview of existing models to the specific justification for GrasslandTraitSim.jl could be made smoother. The authors may consider clarifying this connection by explicitly summarising how the challenges identified in earlier models set the stage for GrasslandTraitSim.jl.
For instance, what are the key limitations of simpler models (e.g. IBC-grass, GraS) and more complex ones (e.g. PaSim, LPJmL) that make an intermediate complexity model necessary? The authors could further substantiate their claims by briefly outlining why existing models are unsuitable for this purpose, for example by summarising the findings of Pulungan et al. (2019), Chalmandrier et al. (2021) and May et al. (2019). This should clarify how the design

of GrasslandTraitSim.jl allows it to be a more realistic representation of species richness than these alternatives.

Thank you for these helpful suggestions. We have added Table 1 (with extensions in Tables A1 and A2) summarising the overview of representative grassland models and their characteristics, including our GrasslandTraitSim.jl model. We also revised the text (starting from line 109) and gave more details on the difference of previously published models to our model. This revised section now more clearly motivates the need for morphological trait-based models like GrasslandTraitSim.jl with climate and management factors included. Chalmandrier et al. (2021, p. 6) already suggested that this transfer function approach could be applied to models that simulate more processes: “Beyond the scope of our particular community model, this new method paves the way to study more complex—and potentially better-performing models—of community assembly and dynamics in diverse ecosystems.”

L 90: GrasslandTraitSim.jl is presented as a model of "intermediate complexity". The authors should provide a more specific definition of this term and explain how GrasslandTraitSim.jl fits into this category, in particular how the model strikes a balance between the computational efficiency of simpler models and the detail of more complex models.

Thank you for the suggestion. We made it clearer what we mean by intermediate complexity (L. 127 - 130):
“Our model is of intermediate complexity compared to the above-mentioned models in terms of the number of equations, which is reflected in the number of simulated state variables and the number of parameters (species-specific and global, non-species-specific, parameters, see Tables 1, A1 and A2).”

L 92: GrasslandTraitSim.jl is compared to IBC-grass, but the similarities and differences are only partially explored. The authors should expand on this comparison and explicitly discuss how the use of continuous traits in GrasslandTraitSim.jl provides advantages over IBC-grass's reliance on discrete trait categories.

Thanks for pointing it out. We removed this misleading sentence. We showed the differences with IBC-grass better with the new Table 1.

L 94: The phrase "Therefore, we believe" may not be considered ideal in academic writing. Authors are encouraged to explore alternative vocabulary for transitions, e.g. "Consequently", to avoid making assumptions about the reader's understanding (which may not be the same as the authors') and to maintain a more objective tone. In addition, relying on phrases such as "we believe" can introduce an element of subjectivity. Instead, authors can strengthen their arguments by using more objective language, e.g. "we hypothesise".

We replaced “Therefore, we believe” with “Consequently” to improve objectivity and academic tone (L. 130).

Some subsections are thorough and grounded in established plant physiological and ecological principles. However, the model relies heavily on simplifying assumptions, such as linearity and species homogeneity in responses to stressors. Trait-based approaches are a strength of this model, but the limited set of traits (e.g. exudation rates) may constrain its applicability to diverse ecosystems.

Thank you for your comment. We responded to this comment in reply to one of the main comments above.

L 101: The introductory sentence of the model could be completed by mentioning the novelty compared to other models.

We revised the Introduction and added Table 1 to clearly highlight the novel aspects of GrasslandTraitSim.jl in comparison to existing grassland models.

L 103: The explanation of "spatially heterogeneous inputs" is vague. Examples of such inputs would be helpful.
L 100: Section 2 is overly technical, with excessive notation and long, dense sentences. Equations and processes are described without prioritising relevance or clarity for the reader. Important limitations (e.g. no patch interaction) are underemphasised.
L 119: The lack of patch interaction is a significant limitation and should be highlighted somewhere.
L 120: The statement is contradictory. If patch size is irrelevant, why does the model design limit it to 1 m^2 to 1 ha? This needs clarification.
The limits of 1 m^2 to 1 ha should be justified.
L 119: Authors should explain how this assumption affects the behaviour or accuracy of the model.

We removed the spatial dimension to be more precise about the model. As we have no patch interactions (we had this in a previous model version and removed it later), it is the same as running several model simulations. We changed the description to (L. 154 - 156): “The model is not spatially explicit and does not account for spatial heterogeneity. As the assumption of spatial homogeneity is only met approximately for smaller spatial dimensions, we suggest using the model for areas between 1 m^2 and 1 ha”

L 103-104: Redundant after subsequent introduction of species-specific variables. Authors may consider merging with later descriptions.

We removed the following sentence, for the explanation see comment above about the spatial dimension:
“Within each patch, many plant species (denoted by the subscript s) can grow.”

L 105: To avoid overwhelming the reader with notation too early, variables could be introduced gradually.

We appreciate the suggestion but opted to retain the current structure for clarity and consistency within the technical section, where full notation is introduced upfront for reference.

L 107: "a set of coupled differential equations" is too abstract. A brief explanation of what these equations capture would add clarity.

We have added a brief explanation that the difference from one day to the next day is calculated to clarify what the equations represent, as follows (L. 142 - 143):
“Changes in the state variables from one day to the next are described by a set of difference equations (for details see Table A5).”

L 108: The term "morphological functional traits" is introduced without examples or prior context.

We have clarified the term by adding an example of morphological functional traits (L. 143 - 144):
“The morphological functional traits of all plant species are fixed (time-invariant inputs, for example the maximum plant height) and linked by model parameters to the species’ demographic processes (Fig. 2).”

L 110: This sentence lacks specificity. The authors could define "performance" and give an example of emergent dynamics.

We have revised the sentence to define “performance” (L. 144 - 146):
“As a result of the differences in the demographic rates of all species, the performance of individual plant species differs (biomass increase or decrease under particular conditions), leading to the emergence of plant community dynamics.”

L 111: This is distracting in the middle of a technical explanation. It can be moved to the conclusion.
Thank you for the suggestion. We think that it is important to refer to the online documentation and made a small change to the sentence (L. 146 - 148):
“While reading the model description, we encourage the reader to have a look at the online documentation, which contains many interactive graphics and flowcharts that make the model description more accessible (see data accessibility statement).”

L 113: The sentence may need some elaboration on examples of "functional traits" and "daily management data". This reviewer argues that the plant functional traits are variables. They should be considered as species-specific parameters.

Thanks for the comment. We added an example of management data and restructured the sentence. The revised sentence is as follows (L. 149 - 150):
“The required model inputs are the plant functional traits of each species, soil properties, daily climatic data and daily management data (e.g., timing and intensity of grazing, Table A3).”

L 114: "parameters"
This could be vague and unhelpful if there is no elaboration on the meaning or examples of these parameters.

Thanks for the comments, we want to refer in this sentence to the table which gives a quick overview about all parameters in the model. We refer to the Table A4 for the specific meaning and examples of the model parameters, as follows (L. 151 - 152):

“The model has in total 54 global parameters (for details see Table A4) that are neither site, time nor species dependent .”

L 115-117: Overly dense sentence: authors may consider splitting the sentence into two, with one focusing on state variables and the other on diversity metrics.

Thanks for the comment. We split the sentences (L. 151 - 154):
“Outputs include the state variables, grazed and mown biomass, community-weighted mean and variance of each trait. Additionally, taxonomic diversity indices (e.g., Simpson diversity) and plant functional diversity indices (e.g., Functional dispersion and Functional evenness) are provided. Both state variables and diversity metrics can be calculated for each day.”

Defaults (e.g. patch size, initial biomass) are often arbitrary and unexplained.
L 124: The choice of default values (e.g. 5000 kg ha-1) is not justified.
There should also be a justification for dividing the biomass equally between above and below ground.
L 126: This arbitrary assumption could influence early growth dynamics.

We decided to shorten this section and write that users can choose initial values (what is possible with our model). We will justify the initial values that were used for calibration/validation of our model in the corresponding sections (L. 157 - 159):
“During the initialisation, the state variables (height, above-ground and below-ground biomass of species, and soil water content) are set to user supplied initial values.”

L 126: Arbitrary threshold (10^{-30}) is unexplained.

We added an explanation, why we added a relative low threshold (L. 159 - 162):
“Very low or negative values (< 10−30) of the height and biomass state variables are set to zero to avoid numerical problems. We have deliberately kept the threshold at a low level because the plant species should be able to recover even from a very low biomass level.”

L 126-127: Authors should be careful not to oversimplify a critical process. Details of loop mechanics or computational requirements would be valuable.

Computational requirements are written in section 3: “Technical details of the GrasslandTraitSim.jl model”. As we removed the possibility to run the model over several patches to keep the model and the manuscript concise to the purpose, the loop mechanism simplified. For example, the order of patches is not important any more. Now, the loop is only run over each day (L. 159 - 160):
“During the simulation, a loop is run over each day.”

L 129: The long list of equation ranges (e.g. 5-34) overwhelms the reader without summarising their content.

The idea of referencing the equations for all the processes was to guide the reader in which equation is which process described. We added the corresponding Sections in the parentheses to make this guidance more useful (L. 162 - 163):

“After that, the main part of the model is executed with the calculation of growth (Sections 2.1-2.1.7, Eqs. 5-33), senescence (Section 2.1.8, Eqs. 34-35), biomass removal by management (Section 2.1.9, Eqs. 36-42), height dynamics (Section 2.2, Eq. 43), and soil water dynamics (Section 2.3, Eqs. 44-52).”

L 131: It needs further justification as it seems to contradict standard practice in dynamic modelling, where order is often important.

We completely removed our justification, that we had in the manuscript, why the order is not important to be more concise and avoid misunderstanding:
“However, the order of the execution has no influence on the results, because the change of the state variables is calculated based on forcing variables (input variables) of the day and state variables of the previous day.”

L 133: Overuse of subscripts makes equations visually complex and intimidating. Authors may wish to simplify or group terms.

We simplified the subscripts by removing the spatial dimension. Since there is no interaction between patches, this is equivalent to running several simulations. Therefore, we decided to remove it from the model description. Consequently, only two subscripts remain: one for time and one for species. These are helpful for understanding the equations.

L 157-158: The sentence structure must be checked.

Thanks, we changed it to (L. 191 - 192):
This concept was widely adopted in grassland modelling studies (Schapendonk et al. 1998; Jouven et al. 2006; Moulin et al. 2021; for a review see Pei et al. 2022).

L 195: "different" can be deleted.

We deleted it.

Given the two methods described for light competition, authors can indicate when users should choose one over the other.

As we advocate for using the more complex method, we completely removed the simple method, also to keep the manuscript concise.

L 242: Mentioning these traits is appropriate, but the rationale for choosing these specific traits over others could be briefly elaborated. For example, why are other potential traits such as root exudate production or root depth not considered?

We added a section in the discussion to clarify the choice of the traits (L. 660 - 669):
“We chose the morphological functional traits that represent main trade-offs in plant physiology. Rather than reflecting one process in detail with many traits (e.g., more traits dealing with water stress, such as stomatal conductance and rooting depth), we aimed to represent the following main trade-offs of plants: (1) The slow-fast continuum of the leaf

economic spectrum states that plants with thinner leaves have a higher light use efficiency per unit of biomass, but also a higher senescence rates (as reflected by specific leaf area; Reich et al. 1992; Wright et al. 2004). (2) Taller plant species can overtop other plant species and are therefore less affected by shading. However, they are more susceptible to mowing and grazing (as reflected by maximum plant height; Díaz et al. 2007; Klimešová et al. 2008). (3) Investing in roots and mycorrhizae enhances nutrient and water uptake, but this comes at the cost of maintaining fine roots and the collaboration with mycorrhiza (as reflected by above-ground biomass per plant biomass, root surface per below-ground biomass, arbuscular mycorrhizal colonisation rate; Reich 2014; Prieto et al. 2015; Bergmann et al. 2020)."

L 242: This opening statement would benefit from a brief mention of what constitutes "poorly adapted" plants in this context, to set the stage for the discussion.

We clarified the sentence by writing (L. 265):
"Plant growth may be reduced when soil nutrient availability is low and plants are inefficient at taking up nutrients."

L 246: This is an important clarification, but stress can also result from nutrient imbalances (e.g. excess nitrogen relative to phosphorus) that are not addressed.

Thanks for the comment. In our model, we only consider one nutrient (nitrogen) to be a limiting factor for plant growth. Taking into account more nutrients (e.g. phosphorus) would make our model more complex, which we tried to avoid. We briefly mention this general point in the new discussion (L. 630 - 632):
"In our model, we tried to keep a balance between a model that can reproduce the basic patterns in biomass production and functional community composition, but does not have too many global parameters, so that it is possible to calibrate all parameters with datasets that are readily available."

L 248: Although equations are clearly presented, the logic of using the maximum of two stresses could be better explained. The rationale given - that plants prefer one strategy over the other - is valid, but assumes no synergistic effects between the two strategies.
L 251: This assumes that plants adopt only one strategy, but empirical evidence often shows a mix of strategies depending on environmental conditions. The binary distinction can be misleading.

Thanks for the comment. To make the simplification clearer, we changed the sentence as follows (L. 272 - 273):
"The maximum of the two nutrient stress factors is used because, for simplicity, we assume that plants can invest either in a high root surface area per total biomass or in a high rate of arbuscular mycorrhizal colonisation."

L 255-257: The explanation of the parameters is detailed, but their empirical basis is not clear. Are these parameters derived from experiments, literature or assumptions? If from the literature, specific studies that validate these assumptions should be cited.

Thank you for your comment. We have made it clearer in Table A4 how the parameter values are derived by adding statements in the column “reference”. The parameters in this section are theoretical in nature; therefore, no literature references could be used to determine their values. Some parameters were set manually by comparing simulated trajectories with measured data, while all other parameters were calibrated in the Bayesian calibration using the Biodiversity Exploratories dataset.

L 261: This statement is correct and supported by the literature, but it generalises the response to all plant species. Non-crop species may behave differently and the text would benefit from a caveat recognising this variability.

We added two references to support the point in a more general way (L. 285 - 286):
“It has been shown that the root-to-shoot ratio increases in many plants under nitrogen-poor conditions (Jiang et al., 2016; Meurer et al., 2019; Lopez et al., 2023).”

L 302: Similar to 2.1.4, this opening is clear but lacks specificity. It could briefly outline what "poorly adapted" means in terms of traits or behaviours related to water stress.

We made it more explicit what we mean by “poorly adapted” (L. 332 - 333):
“Plant growth may be reduced if soil water is low and the plants are poor at taking up water.”

L 304: While this is valid, it ignores other characteristics such as stomatal conductance, root depth or hydraulic conductivity that are critical in determining water uptake efficiency. This omission may be a limitation.

Thanks for the suggestion. We added the following paragraph in the discussion to justify the choice of the traits (L. 660 - 669):
“We chose the morphological functional traits that represent main trade-offs in plant physiology. Rather than reflecting one process in detail with many traits (e.g., more traits dealing with water stress, such as stomatal conductance and rooting depth), we aimed to represent the following main trade-offs of plants: (1) The slow-fast continuum of the leaf economic spectrum states that plants with thinner leaves have a higher light use efficiency per unit of biomass, but also a higher senescence rates (as reflected by specific leaf area; Reich et al. 1992; Wright et al. 2004). (2) Taller plant species can overtop other plant species and are therefore less affected by shading. However, they are more susceptible to mowing and grazing (as reflected by maximum plant height; Díaz et al. 2007; Klimešová et al. 2008). (3) Investing in roots and mycorrhizae enhances nutrient and water uptake, but this comes at the cost of maintaining fine roots and the collaboration with mycorrhiza (as reflected by above-ground biomass per plant biomass, root surface per below-ground biomass, arbuscular mycorrhizal colonisation rate; Reich 2014; Prieto et al. 2015; Bergmann et al. 2020).”

L 306: This exclusion is understandable, but could lead to oversights in ecosystems where waterlogging or flooding is common. A short justification would strengthen this choice.

Thanks for the comment. We added that these systems are not the target system that we want to model (L. 334 - 335):
“Here, we only consider too little water leading to water stress conditions, not too much

water, as our primary goal of our model is not to model systems with regular flooding or waterlogging."

L 307: This simplification assumes that water and nutrient stress act in the same way, which may not always be the case. The biological processes underlying these stressors are very different and the model may benefit from differentiating between them.

We know that water and nutrient stress act differently. Here we have made the design choice to use the same form of equations. This type of equation (logistic functions) is still flexible depending on the parameter values.

L 311: This scaling is standard, but ignores the potential effects of temporal dynamics, such as rapid drying or rainfall events, which can cause transient stresses not captured by the model.

We acknowledge this limitation and have added a sentence to the manuscript noting that the current scaling approach does not account for short-term fluctuations and may cause transient stress responses not captured by the model, highlighting the direction for future development (L. 344 - 346):
"This formulation of plant available water does not take into account some short-term temporal dynamics. For example, after a rainfall event, plants are often not water stressed at all, even if the soil water content is not replenished to the water holding capacity."

L 327: These parameters are critical, but are not biologically intuitive. The model would benefit from examples or ranges from empirical data to support their use.
L 315: This is a good start, but the discussion could benefit from quantifying costs or providing comparative examples from the literature.

Thank you for the comment. These parameters are calibrated with the Biodiversity Exploratories dataset, we made it clearer in Table A4 that documents all parameters. Both parameters have a strong influence on the total growth dynamics (biomass) and on which species has a competitive advantage by experiencing less growth reduction. Therefore, functional composition is influenced by these parameters.

L 319: This statement is clear but simplistic. It might be useful to mention how these costs scale with environmental conditions or plant size.

We made this clearer that cost change depending on the root-to-shoot ratio and added the following sentence (L. 359 - 360):
"Therefore, the cost of maintaining fine and roots and mycorrhizae does change with time depending on the ratio between above-ground and below-ground biomass."

L 320: This equation is simple, but it lacks biological nuance. Are there interactions between the two? For example, does a large root surface area reduce dependence on mycorrhizae?

We made the assumption that the cost is still there and the mycorrhizal colonisation rate for one plant species is constant over time. Changes in the community reflect a change in strategies.

L 330: This reuse of parameters assumes that the same trait values are equally affected by multiple stressors, which may oversimplify trait-environment interactions.

We are aware that the reuse of the parameters for different processes could oversimplify trait-environment interactions. Still, we have to make some simplifications in our model. We added a section in the discussion to describe why we did not want to increase the number of global parameters (L. 630 - 632):
"In our model, we tried to keep a balance between a model that can reproduce the basic patterns in biomass production and functional community composition, but does not have too many global parameters, so that it is possible to calibrate all parameters with datasets that are readily available."

L 338: This is a valuable point, but it would benefit from a brief mention of the thresholds at which photoinhibition becomes significant.
L 352: Although logical, this linearity assumption may not capture non-linear responses observed in nature, such as thermal acclimation.

Thanks for these comments. We made the simplification that the stress due to excess radiation is not temperature dependent. We decided to keep this part of the model simple, to have more thorough formulations in other parts of the model. Therefore, we used the same formulation as in Schapendonk et al. (1998).

L 354: This is an important addition, but appears to be based on empirical step functions rather than mechanistic understanding. A brief justification for this choice would improve transparency.

We made it clear, that we used empirical step functions by changing the following sentence (L. 381 - 382):
"The temperature adjustment factor is based on the empirical step functions by Schapendonk et al. (1998) that were adjusted by Jouven et al. (2006):"

L 366: Authors may wish to briefly explain why these particular factors were chosen.

We are not sure what is meant by that comment. The scaling of the senescence rate is explained later (L. 404 - 406):
"To facilitate interpretation, we have chosen to use the basic senescence rate per month αSEN. Consequently, αSEN has been converted to a senescence rate per day, assuming a monthly duration of 30.44 days."
Moreover, the use of the specific leaf area to scale the senescence rate is explained in detail later.

L 377: This is correct, but would benefit from a citation beyond Wright et al. (2004) to show wider applicability.

Thanks for the hint, we added two more references (L. 412 - 414):
"This means that they tend to be highly photosynthetically efficient, modelled here with400

a higher leaf area index per biomass, but have a short leaf lifespan and therefore a high senescence rate (Reich et al., 1992; Wright et al., 2004; Onoda et al., 2017)"

L 380: This assumes that seasonality affects all species similarly, which may not reflect interspecific variability. Authors may wish to consider addressing this limitation.

This was our deliberate choice to not include species-specific seasonality. In the text, this should be clear because we included it in the section "Community environmental and seasonal factors". We made this choice because we focused mostly on the slow-fast continuum in plant strategies.

L 484: The description of the hardware requirements should be more specific to make the claims of efficiency verifiable.

Thanks for the comment. We clarified that we used a standard personal computer and not a cluster (L. 739 - 741):
"The model can be run on computers with low hardware requirements. For example, a 10-year simulation involving 70 species typically takes less than half a second to run on a standard personal computer."

The online documentation is mentioned, but no details are given on its depth, examples or unique features (e.g. visualisations, tutorials).

We already wrote quite a lot about the online tutorials and are not sure that more details are needed here.

The GUI and flowcharts are important usability features, but more detail on their functionality would help readers to understand their full potential.

Thanks for the comment. We think it is enough to provide the link to the online documentation, where both are explained in detail.

The section does not explicitly mention whether the installation of additional dependencies (besides Julia) is required.

We added the following sentence (L. 751):
"All dependencies will be installed with this command."

While contributions are encouraged, practical steps or links for contributors (e.g. guidelines, issue tracker) are not mentioned.

We added that contributions can be made on Github (L. 752):
"Contributions are welcome and can be made via GitHub."

However, the level of technical detail provided in the current section 3 is not appropriate for the main text of a scientific article. Instead, these details should be relocated to a user guide or supplementary material, with the main article focusing on the scientific innovation, methodology and applications of the model. Including a link to the user guide ensures that interested users can still access the full implementation details.

We moved this section to Appendix B.

L 511: The decision to include only one long grazing season, although justified by incomplete data, may oversimplify grazing dynamics and potentially reduce the ability of the model to simulate finer temporal dynamics.

Thanks for the comment. As we had the information on the grazing type, we implemented a technique to create several short grazing periods for some grazing types. We changed the description as follows (L. 758 - 761):
"The exact dates of grazing were not available, only the type of grazing, the number of days and the start and end month of a grazing period. We assumed different numbers of consecutive grazing days (2 for rotational grazing type I - "Portionsweide", rotational grazing, 5 for rotational grazing type II - "Umtriebsweide" and all days for permanent grazing) and distributed them equally over the whole grazing period.

L 528: Although the collection of trait data is thorough, there are some inconsistencies in the availability of traits for all species. For example, setting a fixed value (e.g. 80% leaf biomass per plant biomass) for missing data could introduce bias or reduce the accuracy of the trait representation.

Thanks for this comment. We are aware that it would be more accurate to have species-specific trait values for leaf mass per above-ground biomass. Still, we think that the model is usable with one trait without species-specific values.

L 550: Authors may wish to provide more detailed justification for fixed parameter values and their potential influence on simulation results.

We added a local sensitivity analysis of all parameters after the calibration with the FAO and the Biodiversity Exploratories dataset (Tables A12 and A13). Moreover, we expanded the description on how we fixed the parameter values in the Table A4.

L 556: change "good results" to "satisfactory results"

We changed that in the new description.

L 558: While the authors acknowledge that site-specific characteristics may not be adequately reflected, there is no discussion of how this limitation might be addressed in future iterations of the model.

Thanks for the comment. We discussed in more detail that site-specific characteristics may not be adequately reflected for the FAO data set. In the future, this could be addressed by allowing that parameter values are dependent on the location of the site (L. 651 - 658):
"For the independent validation site with the highest error of the FAO dataset (FAO45 in Spain, see Figure 4), our model predicts too high above-ground biomass in spring. Thereby, we see that the model is not flexible enough to simulate production in a very wide range of regions. Our step function for seasonal growth adjustment assumes that the growth

increases in spring after 200 ° C have been accumulated (see Eq. 32). This might be a reasonable assumption for Lolium perenne in the Netherlands, but not for sites in Spain. The strong growth starts too early for the site in Spain. For the calibration of the LINGRA model with the same dataset, it was assumed that species-specific parameters are different for the northern and southern sites (Bouman et al., 1996). We did not calibrate the model here for spatial subsets of the sites, as we wanted to analyse whether our model is in general applicable to a variety of sites."

L 566: The scenarios (mowing only and grazing only) are valuable for illustrative purposes, but may oversimplify real grassland management, which often involves a mixture of both practices. This limits the applicability of the results to complex systems.
L 571: The authors could provide a more detailed discussion of how functional dispersion and evenness respond to land use intensity, and link these metrics to ecosystem functioning or resilience.
L 572: The decision to exclude functional richness to avoid setting an arbitrary extinction threshold limits the scope of functional diversity analysis. The authors could explore potential alternative methods or thresholds that could have been used.
L 576: This statement seems to contradict the rest of the passage, which gives a fairly detailed interpretation of the results. The authors might consider emphasising their cautious approach, acknowledging that the results are primarily illustrative and that extensive speculation should be avoided.
L 578: The text could be more specific: instead of saying "stronger land uses", it could specify the types of land use (e.g. intensive grazing, frequent mowing).
L 589: The results are presented as illustrative, but lack a broader context of their potential applications in grassland management. For example, it is not clear how these trends might inform decisions about optimal mowing frequency or grazing intensity.
Figure 5: While the figures show the variability at site level, the text should explore how or why certain sites deviate from the mean. This could be valuable in understanding the role of site-specific factors.

We decided to completely remove the section with the simple case study (see above the response to the editor).

L 600: While the authors acknowledge the low number of coexisting species, they do not provide much detail on how this limitation affects the ecological realism of the model. This is an area that could benefit from further investigation, e.g. how these few species dominate the biomass, or whether this is a result of specific model assumptions or a more general limitation of current trait-based approaches to modelling plant communities.

Thanks for the comment. We discuss this point now more in detail in the discussion (L. 675 - 683):
"The number of coexisting species (e.g., with biomass > 2 %) is rather low, with three to five species accounting for most of the biomass in most scenario analyses. This is a common challenge in grassland models. For example, in a model comparison study with the GRASSMIND and LPJmL models, it was noted that in a two-species simulation always one species always accounted for most of the biomass (Wirth et al., 2021). We noticed that by including a density-dependent senescence rate (not shown in the model equations above), the simulated functional diversity is increased, and the distance between modelled

and observed community trait distributions can be lowered. A density-dependent senescence rate can be explained, for example, by negative plant-soil feedbacks (Bonanomi et al., 2005; Liu et al.,2022; Goossens et al., 2023). This shows the potential to explore in future studies how the incorporation of coexistence mechanisms can lead to more realistic predictions of functional community composition."

L 603: While the authors acknowledge the inclusion of negative density dependence in nutrient competition and explore species replacement patterns, a critical aspect missing from their analysis is an evaluation of how simplifying assumptions, such as proportional biomass removal by grazing and mowing, affect the ecological realism of the model.

Thanks for the comment. The mentioned biomass removal by mowing and grazing is already quite complex in our model. It depends on the above-ground biomass and the height of the plant species, and for grazing also on the leaf nitrogen content. We refer again to the following part of the discussion that we did not aim to add too many details in our model (L. 630 - 632):
"In our model, we tried to keep a balance between a model that can reproduce the basic patterns in biomass production and functional community composition, but does not have too many global parameters, so that it is possible to calibrate all parameters with datasets that are readily available."

L 611: This is critical as the paper focuses on describing the model and the practical applications of the model are not fully explored in the conclusion. For example, while the authors acknowledge that the model simulates land use effects, they could discuss how their results can inform grassland management practices or what stakeholders might find most useful in terms of ecosystem service provision.

Thanks for the comment. We added a detailed description of potential model applications in the discussion (L. 684 - 695):
"We argue that our model is well suited for analysing the effects of management (grazing, mowing and fertilization), of edaphic factors (soil nitrogen, permanent wilting point and water holding capacity), and of climatic factors (temperature, radiation, potential evapotranspiration and precipitation) on the productivity and the functional composition of diverse plant communities of temperate semi-natural grasslands. We envisage the model as a useful tool for conducting scenario analyses (e.g., what would happen if the input X were to change, and why?), rather than as a model with superior predictive performance compared to conventional statistical models. For example, the influence of management type and intensity on achieving a balance between creating highly productive grasslands and maintaining plant diversity could be analysed. Furthermore, the influence of the initial species composition on the productivity under fluctuating climate conditions (e.g., years with drought) could be studied by answering the question whether a more diverse community can buffer extreme climatic events. Moreover, we consider the potential application of including or excluding certain processes (e.g., a specific transfer function, which links traits to demographic rates) and analyse whether the agreement between simulations and measured data improves."

L 612: The authors mention the problem of not calibrating below-ground biomass, height and soil water content, but there is no mention of how they plan to overcome these challenges in the future. The authors may be planning to address these challenges in the future, e.g. by collecting new data or considering model-based approaches to estimate these values. There could also be more discussion of why these particular aspects (e.g. below-ground biomass) are crucial for improving the model, especially as they affect key ecosystem processes such as nutrient cycling and root competition.

Thank you for your comment. We have restructured the discussion section so that we no longer discuss the calibration of state variables that were not included in the calibration or validation process. Including all state variables in the validation could increase the reliability of the model outcomes. However, for the variables we are interested in — mainly the community composition of above-ground biomass due to the morphological functional traits of the species — the exact values of some state variables are not important. For example, the exact value of soil water content is unimportant, since it is rescaled, and the parameters of the growth reducers are then calibrated to these values. The growth reducer formulation is flexible enough to produce similar results by changing the parameter values, even if the rescaled soil water content is lower or higher. Therefore, we view the other state variables as tools for reliably simulating above-ground biomass and are not directly interested in interpreting their values.

Appendix A: A technically rigorous derivation of species-specific water and nutrient growth inhibitors is presented in Appendix A. While the use of logistic growth curves is a well-established and effective method in ecological modelling, the complexity of the equations presented and the limited biological context may be challenging for some readers. To improve the accessibility of this section, the authors could consider:
- Using a more concise and intuitive notation system to improve readability and comprehension;
- Providing clearer and more detailed explanations of the biological basis of the equations to strengthen the link between the mathematical formulations and the ecological processes they represent;
- Including practical examples and case studies to help readers visualise the application of the growth reducers in real-world scenarios, making the concepts more concrete and easier to grasp.

Thanks for the suggestions. We answered about case studies in the main response. These equations are not easy to convey, we chose them so that the parameters are easier to interpret. The biological basis of these equations are already described in the main text.

Appendix B: The model's detailed inputs and parameters reflect a thorough approach to understanding plant growth dynamics under different environmental and management conditions. While its complexity and reliance on calibrated parameters are strengths, further clarity on calibration methods and model sensitivity could improve its accessibility and reliability in broader contexts.

Thanks for the comments, we improved the calibration and validation section and included a local sensitivity analysis of the parameters after the model calibration with the FAO and the Biodiversity Exploratories dataset in the Tables A12 and A13.

Table B2: Unit MJ-ha-1, spelling

Thanks, we corrected the spelling mistakes. The correct unit for the parameter $\gamma_{RAD,1}$ is [$MJ^{-1} \cdot ha$] because the resulting growth reducer is dimensionless.

Reviewer 2

The manuscript presents a new model for grasslands, that allows simulations of the response of grasslands to land use and environmental drivers. Overall, I agree with the authors that grassland models need to be improved, particularly in the context of large-scale/global models that often focus on forests. I appreciate the trait-based approach that allows analyses of trait diversity and how it responds to grazing and other drivers. Thereby, a detailed representation of individual species in not required. Understanding grassland dynamics is important and timely due to increasing pressures by land use and climate change.

The model description is detailed and comprehensive, and it is published together with examples and tutorials. This makes the model accessible to the modelling community, and results are reproducible. Nonetheless, some aspects of the model could be better motivated and justified. Therefore, I suggest being more specific about these points, particularly by adding supporting references where possible (see also minor comments below for examples).

We appreciate the reviewer's acknowledgment of the relevance of our model and its trait-based approach, as well as the importance of accessibility and reproducibility. We added. We added the new Table 1 to distinguish our model from similar grassland models and clearly show why there is a need for building a new grassland mode. Moreover, we added a discussion section to discuss strengths, weaknesses and potential applications of our model. For example, we better motivated the choice of the plant functional traits in the discussion (L. 660 - 669).

I understand the argument that the manuscript is about presenting the model and that including full case studies would make the manuscript extremely long. A caveat of this approach is that it is more challenging to convince readers that the model is a step forward and to highlight the benefits compared to other models. In addition, a discussion of model uncertainties and deviations between the presented model result and data remains open. I recommend adding a short discussion on these points, and therefore shorten elsewhere. I think there is potential to shorten by being more concise.

Thanks for these suggestions. We shortened the model description by removing the description that the model can be simulated in a spatial grid, and we removed the description of the inferior second option for light competition. We added a more complete calibration and validation of our model by adding a second data set and using more spatial replicates for the data set of the Biodiversity Exploratories. Moreover, we expanded the discussion on the validation by adding a completely new discussion section (L. 619 - 658).

Minor comments:

l 27: "soils that are neiter too dry nor too wet" this is a bit vague. Another factor in this context is fire, even though it's not so relevant in the grasslands considered in the study.

We restricted the sentence to the semi-natural grasslands in the temperate zone and added

an explanation that regular disturbances can lead to maintenance of grassland ecosystems. The changes sentences are now (L. 28 - 31):
“The key factor in maintaining the semi-natural grasslands in the temperate zone is management, as well as regular natural disturbances, such as low-intensity fires or avalanches, without which grasslands would become woodlands. This is because the abiotic conditions on most grassland sites favour tree growth, by having the sufficient temperature, precipitation, soil moisture and nutrients (Petermann and Buzhdygan, 2021).”

l 49: "focussing on different parts of the model" I would rather argue that different model focus on different components and aspects of the ecosystem such as grazing, diversity, fire impacts, etc.

Thanks, that is true, we reformulated it (L. 56 - 58):
“Historically, different research questions on grasslands, ranging from ecology to biogeochemistry, have led to the development of different grassland models by focusing on some parts of the grassland system while simplifying others.”

l 50, 51 remove "for example" in one of the sentences?

Thanks, we removed the second “for example”.

l 71: most DGVMs such as LPJmL include grass PFTs or herbaceous PFTs for C3 and C4 photosynthetic pathways, but their representation is typically simplified compared to trees.

Thanks for the hint, we included a sentence that two plant functional types are simulated in LPJmL (L. 81 - 83):
“However, the representation of plant functional diversity in these models is limited. For example, in LPJmL only two plant functional types (C3 and C4 grasses) are simulated in natural and managed grasslands (Rolinski et al., 2018). “

l 118: "possible to run simulation with just one patch" I suggest rewording, it should also be possible to do this in a model with interaction, even though interactions don't play a role in this case.

For clarity, we removed this part of using spatially heterogeneous inputs to simulate a spatial grid, because, without interactions, this is the same as running several simulations in parallel.

l 124: 5000 kg/ha, is this a common biomass in the study system and a reference?

We decided to shorten this section and write that users can choose initial values (what is possible with our model). We justify the initial values that were used for calibration/validation of our model in the corresponding sections (L. 157 – 158):
“During the initialisation, the state variables (height, above-ground and below-ground biomass of species, and soil water content) are set to user supplied initial values.”

l 127: can negative values occur in the model? Does that indicate caveats in the model design?

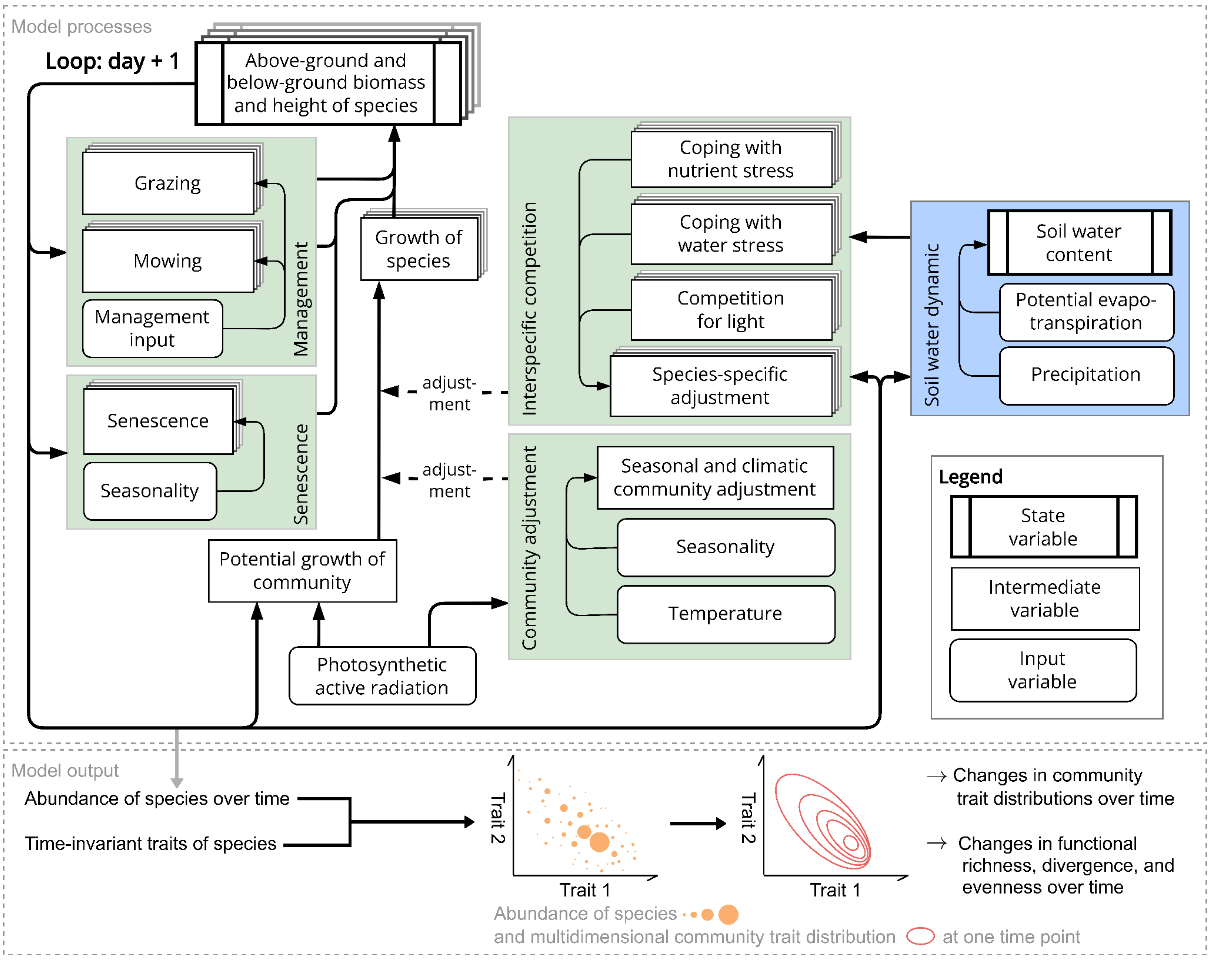
Negative values can occur, especially because we use an integration step of one day. With a smaller time step, this wouldn't be possible. After each time step, we set low or negative values to zero to avoid problems in the simulation.

l 129: based on the list of processes, is there also plant mortality, or is this implicitly considered in senescence?

Yes, the plant mortality is implicitly considered in senescence.

Fig 1: This scheme is very helpful, but I was uncertain where to start and why the last step is at the top. So maybe add some letters or numbers to guide readers through the graph (and refer to the letters in the figure caption) or restructure that the first step is at the top.

We rearranged the scheme, the analysis part is below the main simulation. Thanks for pointing that out. Here is the updated figure:



l 175: "easier to think about growth reducer" this is vague, and I recommend clarification.

Thanks for the comment, we changed the respective sentence to (L. 207 - 208):
“The 0.2 m has been arbitrarily set, and the parameter $\alpha_{RUE,cwmH}$ is inversely calibrated.”

l 183: in my experience, it's better to calculate total LAI by caculating the total leaf area in a patch divided by patch size, instead of summing LAI. But maybe this is the same in your model.

First, we calculate the leaf area index of each species and, second, the total leaf area index is the sum of the leaf area of all species. We don’t see a problem as the unit does not change, and leaf area indices can be summed up - there is no non-linearity in calculating leaf area indices.

l 184: if abp is abovground per total biomas - shouldn't it be multiplied in the equation?

For clarity, we change the input trait lbp from leaf mass / total biomass to leaf mass / aboveground mass and, therefore, don’t use abp in the equation for calculating the leaf area index.

In the equation, we wanted to calculate the proportion of leaf biomass on the aboveground biomass. We used the leaf biomass per total biomass (lbp) and the aboveground biomass per total biomass (abp). The total biomass was cancelled out in the equation:

lbp / abp =
(leaf biomass / total biomass) / (aboveground biomass / total biomass) =
leaf biomass / total biomass * total biomass / aboveground biomass =
leaf biomass / aboveground biomass

l 190: "we want to distribute..." I recommend rewording, you don't distribute the growth but this describes the rule how it's done in the model.

Thanks for the comment, we updated the introduction to the section (L. 221 - 222):
“The proportion of the potential growth of each plant species to the potential growth of the community is based on the leaf area index and height of the species.”

l 245: nutrient deficit instead of "too litle nutrients"

Thanks, changed (L. 268).

l 397: "distributes the grazed biomass" do you mean that the amount of biomass removal by grazers is species-specific?

Thanks for the comment, we changed the sentence to (L. 432 - 434)::
“The grazing function is divided into two parts: the first part defines the total grazed biomass and the second part the proportion between the grazed biomass of each species and the total grazed biomass.”

l 413-414: why does fodder supply not fully compensate the requirements of animals? Is there are reference for this assumption?

We made it clearer, that if no reachable biomass is left, the farmers will fully compensate the requirement (L. 444 - 447):
“Furthermore, we assume that if the overall reachable above-ground biomass is low, the farmers will gradually increase the supply of additional fodder resulting in less grazed biomass. If no reachable above-ground biomass is left, the farmers will fully compensate the requirements of the livestock animals.”

l. 437: "die along the stem" is unclear to me, please clarify.

We simplified the sentence to (L. 474 - 475):
“Since leaves can die without reducing height, we assume that senescence has no effect on plant height”

l. 475: use SI units instead of bar (ie Pa)

We changed the unit from bar to kPa.

l. 500-507: you could add a table in the supplement to provide more details for the sites

Thanks we added the Tables A7, A9, A10, A11 to give readers a quick overview of all sites used in the calibration and validation of our model.

l. 510: does this simplification influence the model results? Time between grazing periods allows vegetation to recover, such that the grazing impacts may differ compared to a long grazing period.

Thanks for the comment. As we had the information on the grazing type, we implemented a technique to create several short grazing periods for some grazing types. We changed the description as follows:
“The exact dates of grazing were not available, only the type of grazing, the number of days and the start and end month of a grazing period. We assumed different numbers of consecutive grazing days (2 for rotational grazing type I - "Portionsweide", rotational grazing, 5 for rotational grazing type II - "Umtriebsweide" and all days for permanent grazing) and distributed them equally over the whole grazing period.

l. 535: "more trait data" than? Not clear to me what more refers to.

Thanks for pointing it out. The sentence did not make sense. The description of the data preparation for the Biodiversity dataset changed and can be found in Appendix C.

l. 537: "callibration data from 2010 to 2021" why was the end of the entire period used for callibration? Often, the first period is used for calibration and the second for forecasting/model evaluation.

Thank you for your comment. We understand that it often makes sense to use the second part of a time series for validation purposes. However, as we had sufficient spatial replications for this purpose, it was unnecessary to split the time series.

l. 538-539: are the sites in the north and the south similar enough such that a model calibrated sith northen sites can be applied to the southern sites? Why did you not use a randomized design for calibration and evaluation sites?

We changed the calibration with the Biodiversity Exploratories and included from each of the three regions sites in the calibration of the model.

l. 551: replace "make sense" by "are plausible", can the ranges also be backed up with references?

Now, we included tables to summarize all used priors for the calibration with both datasets. The priors for the calibration are backed up with references (see Table A6) and the other priors are chosen so that simulated trajectories are in the range of the data (see Table A8.)

l. 576: "we do not interpret the results" I suggest removing this statement or - even better - interpreting the resuls, even if this is the aim of a follow-up study. As it is now, I think this is unsatisfying for the readers to see model results but no interpretation.
l. 587: are consumed more instead of "fed more"?
l. 591: "we think ... model can be used..." This does not sound very convincing, I recommend a more convincing statement highlighting the novelty and potential of the model.

We decided to completely remove the section with the simple case study (see above the response to the editor).

l. 595: climate change was not studied, is this an aim of a subsequent study?

Thanks for the comment. One of the potential applications with the model is to study the influence of climate variables on the plant functional composition and the productivity of temperate semi-natural grasslands. We included the following sentences in the discussion (L. 684 - 689): “We argue that our model is well suited for analysing the effects of management (grazing, mowing and fertilization), of edaphic factors (soil nitrogen, permanent wilting point and water holding capacity), and of climatic factors (temperature, radiation, potential evapotranspiration and precipitation) on the productivity and the functional composition of diverse plant communities of temperate semi-natural grasslands. We envisage the model as a useful tool for conducting scenario analyses (e.g., what would happen if the input X were to change, and why?), rather than as a model with superior predictive performance compared to conventional statistical models. ”

l. 613: maybe provide examples of what data sets you have in mind.

Thank you for the comment. We directly included the FAO dataset in our study to calibrate and validate intra-annual dynamics of above-ground biomass.