List of all relevant changes made to the manuscript

Please find below a list of all relevant changes made to the manuscript, in chronological order of appearance through the manuscript. Smaller changes, such as adopting consistent spelling for 'pan-Arctic' or 'parameterisation' are not included in this list. A point-by-point response to all reviewer comments can be found below, providing context for the listed changes. Line numbers refer to the track-changes file.

- L. 85-88: Rephrased shrub definition to be more general, following reviewer 2 comment #4.
- L. 145 – 167 and Table 2: Updated the eddy-covariance flux sites used for calibration and evaluation of simulated shrub CO_2 fluxes to include six instead of previously only three stations and updated description of selection and data analysis, in response to comments from all three reviewers (RC1 comment #1, RC2 comments #2, #6 and #10, RC3 comment #9).
- Figure 1: Updated to include the additional CO_2 flux evaluation locations and rephrased to clarify the parameterisation approach (calibration against pan-Arctic observations, not site-level calibration).
- L. 171-172 and appendix A1: Addition of a more detailed ORCHIDEE model description in response to comments from reviewers 1 (comment #2) and 3 (comment #1).
- L. 192-198: Clarification of the approach and justification of the reliance on mainly one model grid cell for parameter calibration and optimisation, in response to reviewer 1 comment #3 and reviewer 2 comment #2.
- L. 236 and 256-257: Addition of details to reconstruct ORCHIDEE parameter values and their physical interpretation to increase transferability between models, in response to reviewer 3 comment #7.
- L. 261-266: Clarification of the impact of nitrogen use efficiency on Vcmax, in response to reviewer 3 comment #2.
- L. 273-279: Clarifications regarding the optimisation approach and update of equation 4, following reviewer 2 suggestion in comment #7.
- L. 288-289: Clarification in response to reviewer 2 comment #8.
- L. 343-353: Clarification of selection of GPP evaluation datasets, in response to reviewer 1 comment #4.
- Table 4 caption: Clarification that parameter values not listed here are kept the same as in the underlying tree PFTs, in response to reviewer 3 comment #2.
- L. 378-420 and Figures 4, B2 and B3: Update of shrub CO_2 flux evaluation including the additional three eddy-covariance sites and small adjustments following reviewer 2 comment #10.
- L. 425-427: Clarification in response to reviewer 2 comment #13.
- L. 483: Slight rephrasing in response to reviewer 2 comment #14.
- Table 6: Reordered models by complexity of their shrub PFT implementation. Added information on shrub PFT height where available, following reviewer 1 comment #8.

- L. 487-509: Added an explicit discussion of the chosen approach to calibrate and optimise PFT parameters in a single model grid cell, in response to reviewer 1 comments #3 and #5, and reviewer 2 comment #2.
- L. 551-557: Rephrasing of discussion of snow-shrub interactions, in response to reviewer 2 comment #15.

Point-by-point response to Referee #1 comments

RC1: 'Comment on egusphere-2026-1071', Anonymous Referee #1, 18 Mar 2026

The manuscript "Introducing shrubs enhances the representation of high-latitude vegetation and carbon cycling in the ORCHIDEE land surface model" refined PFTs for shrublands and improved a LSM using updated PFT maps for the Arctic region. The study also investigated the influence of shrub PFT incorporation on estimates of biomass and carbon fluxes and used ground-truth data for model evaluation. In general, this manuscript is well written and follows a clear logical flow. The improved PFT methodology can benefit the broader modeling community and should be relevant to the general audience of Biogeosciences.

I have a few minor suggestions that may help the authors further improve the manuscript:

We sincerely thank the reviewer for their careful reading and constructive comments. The suggestions have been very helpful and have led to substantial improvements to the manuscript. All comments are addressed below.

1. L147. Three is a relatively small number, given that more sites could be categorized as shrublands according to the IGBP classification. Please provide more detail on why other sites were excluded based on specific criteria.

We thank Referee 1 (REF#1) for highlighting this limitation. The selection of sites was based on their vegetation descriptions, excluding, for instance, burn sites within the boreal forest, with fire succession vegetation, that consists mostly of young trees instead of shrubs (i.e. CA-NS6, CA-NS7, CA-SF3, CA-LU1, CA-LU2). Furthermore, we limited our selection to three sites for practical reasons and selected them based on their geographic location to account for maximum representativeness coverage within the pan-Arctic region.

However, to address the comments of Referee 1, 2, and 3 pointing to the same issue, we have revisited this choice and included three further sites into the analysis, including recently published data. Now the manuscript includes all sites suitable to represent high-latitude shrub vegetation according to our criteria from the FLUXNET and Ameriflux datasets at the time of writing. We have specified the site selection criteria in L.144 and following, as listed below:

> *"Data documenting the carbon fluxes between the atmosphere and shrub vegetation were extracted from the FLUXNET2015 (Pastorello et al., 2020) and AmeriFlux (Chu et al., 2023) datasets. We selected sites classified as open or closed shrubland (OSH/CSH), located above 50°N and with the full FLUXNET-standard post-processing available. Based on the available vegetation descriptions, further sites were excluded, including burn sites with fire succession vegetation, and one tussock-tundra site (US-ICt), resulting in six selected shrub sites (cf. Table 2).*

It is important to note that due to the nature of EC flux measurements, the recorded fluxes represent the aggregate of the entire ecosystem around the tower within a variable footprint, and the data does not specify the contributions of individual PFTs. Therefore, the sites were not assigned to specific shrub PFTs, but used instead as a benchmark for shrub tundra ecosystems to calibrate and evaluate all three ORCHIDEE shrub PFTs against. Notably, the available vegetation descriptions indicate that dwarf and low shrubs seem to be more prominent than tall shrubs at most sites (e.g. van der Molen et al. (2007); NEON (National Ecological Observatory Network); cf. Table 2). Across all six sites, a range of average daily gross primary productivity (GPP) of 0.73 (0.18-2.45) $gCm^{-2}d^{-1}$ (median (interquartile range (IQR)), and average annual GPP of 336 (209-413) $gCm^{-2}yr^{-1}$ was extracted. The GPP statistics were calculated using nighttime flux partitioning, after excluding negative GPP estimates and years without positive values."

Accordingly, we have updated Table 2, Figure 1, Figure 4 and Figures B2 and B3 to include the additional sites.

The additional sites show similar patterns as already described in the manuscript, both in the flux exchange magnitude and seasonality, and the degree of agreement with the ORCHIDEE flux estimations at the respective locations. Their patterns and fits with the simulated fluxes are added to the results in L. 346 and following, without changing the message. We have added a comparison between average GPP from all six sites and across the entire tundra region for all ORCHIDEE shrub PFTs.

2. L155. Given the importance of the ORCHIDEE model for this study, consider adding a diagram and/or additional description of the model (perhaps in Appendix if space is limited) for a broader audience. This would help readers understand how this model compares with other LSMs and to what extent the findings can inform other modeling efforts.

We thank REF#1 for the good suggestion and admit that a conceptual diagram of ORCHIDEE in such context would indeed be useful. However, we also acknowledge that representing such a complex model in one single diagram is a complex task that may add additional generic ORCHIDEE details which can ultimately distract the reader from the main purpose of the paper. ORCHIDEE has been extensively reported in a large number of studies, and a dedicated publication describing the newest ORCHIDEE version (including updated diagrams and full description of the model) is currently in progress (Peylin et al. in preparation). To address REF#1 comment we chose to include a more detailed model description in the appendix:

"ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) is the land surface component of the French ESM IPSL-CM (Boucher et al., 2020) developed by the Institute Pierre Simon Laplace (IPSL). ORCHIDEE can be run in a coupled set-up interacting with the IPSL atmospheric, ocean and ice sheet models, or as a stand-alone model, offline, as done in this work. Inputs required for the simulations include atmospheric forcing (air temperature, humidity, wind

speed, surface pressure, and radiation), land surface characteristics (PFT distribution and dominant soil type), and initial conditions for energy, water, carbon, and nitrogen pools. These inputs are sourced from global datasets such as atmospheric reanalysis products, land cover maps, and soil texture databases. ORCHIDEE simulates carbon, nitrogen, energy and water cycling within the terrestrial biosphere, and the exchanges along the soil-vegetation-atmosphere interface (Krinner et al., 2005; Vuichard et al., 2019).

Different biomes are represented by fifteen generic PFTs, including forests, crops and grasslands. They are differentiated based on climate zone (tropical/temperate/boreal), leaf habit (evergreen/deciduous), leaf type (broadleaf/needleleaf) and photosynthetic pathway for grasses and crops (C3/C4) (Box, 1996) (see Table A1). Woody vegetation is represented using a dynamic, vertically discretised canopy scheme that enables explicit simulation of tree-scale demographic and growth processes, including light penetration allowing photosynthesis to be calculated at each canopy level, sapling establishment and density-dependent self-thinning (Naudts et al., 2015). Three diameter classes within each woody PFT allow simulation of different plant sizes within a stand. Mortality due to fire, wind, bark beetles and forest management are implemented in ORCHIDEE, but deactivated in this work (Chen et al., 2018; Marie et al., 2024; Yue et al., 2014; Yue et al., 2018).

The terrestrial carbon cycle, including processes of photosynthesis, respiration, soil carbon dynamics and plant carbon allocation and phenology, is simulated in the STOMATE module. Carbon assimilation and stomatal conductance are described following the concepts of Farquhar et al. (1980) and Ball et al. (1987), as implemented through the analytical solution proposed by Yin and Struik (2009). Photosynthetic processes are resolved at a half-hourly temporal resolution, whereas the allocation of carbon and nitrogen to vegetation compartments (namely leaves, fruits, roots, above-and belowground sapwood and heartwood, labile, and carbohydrate reserve pool), litter, and soil pools is simulated on a daily time step. The allocation of carbon to leaves, roots and wood follows allometric relationships, based on the pipe model theory (Shinozaki et al., 1964). Soil mineral properties are prescribed at the grid cell scale based on the 12 class USDA soil taxonomy (Forbes et al., 1987).

The hydrological and biophysical processes of water and energy exchange are described by the surface-vegetation-atmosphere transfer scheme SECHIBA (de Rosnay & Polcher, 1998; Ducoudré et al., 1993). Energy fluxes in the model are calculated uniformly at the grid cell scale, without explicit representation of sub-grid heterogeneity, with the exception of snow processes, for which a distinct energy balance is evaluated over the snow-covered fraction of each cell. In contrast, hydrological processes are simulated separately for different vegetation components within a grid cell, employing three distinct soil columns representing bare soil, tall vegetation (e.g. trees), and short vegetation (e.g. grasses). Vertical discretisation differs among model components: soil hydrology is represented using an 11-layer profile extending to a depth of 2 m, with layer thickness increasing with depth according to a geometric spacing. Thermal processes and soil carbon and nitrogen dynamics share the same grid, but extended downward by seven additional layers to a total depth of approximately 90 m."

3. L177. Provide more detail about the specific grid cell, including its resolution and representativeness, and explain why a single cell is sufficient for sensitivity tests and optimization.

We thank the referee for raising this concern and apologise that we were not clearer in the explanations around this point.
The site-level simulations were run at the same resolution as the pan-Arctic simulations, i.e. 2 x 2 degrees. We now clarify that in the manuscript in L.177.
Different locations were tested during the initial calibration phase and compared to ORCHIDEE outputs over the entire region. The grid cell around Toolik Lake was chosen because the average values of the relevant variables at this grid cell were close to the average values over the entire region, making it a representative site.
Generally, one location was deemed sufficient for calibration of variables like height and biomass, because in ORCHIDEE they do not show a strong dependence on site-specific conditions (cf. the relatively narrow ranges of height, diameter and biomass over the entire region given in Table 5). Parameterisation in one representative grid cell against target values capturing averages of the whole region is a reasonable setup to avoid site-specific biases, while still complying with model limitations, such as computational resources. In this context, optimisation with ORCHIDAS had to be limited to one site for computational reasons. However, test optimisations were also carried out at further sites. Please also note that the CO_2 related flux-variables, with a stronger variability based on site conditions, were evaluated at several locations to avoid site-specific biases.

To address the reviewers' comment, we have added these aspects to the discussion in L. 456 and following:

> *"Contrary to the site-level calibrations in many other shrub implementations (e.g. Wolf et al. 2008; Meyer et al. 2021; Sulman et al. 2021), the ORCHIDEE shrub PFT classification and parameterisation presented here are constrained by synthesised observations spanning the entire high-latitude region, increasing its transferability and limiting site-specific biases (Alton et al. 2011; Kuppel et al. 2014). This is in line with recent work by Murphy et al. (2025), highlighting the limitations of model parameterisations based on plot-level observations from a single site. Thus, the methodological choices, including the shrub PFT classification scheme, calibration targets and vegetation distribution, are applicable beyond the ORCHIDEE model.*
>
> *While parameter selection and optimisation were mainly carried out in a single grid cell due to computational limitations, the results indicate that parameterisations can be considered regionally representative, since simulated shrub characteristics match synthesised observations not only in the optimisation grid cell, but across the entire region. Parameterisation in a single grid cell was further deemed sufficient for the targeted variables like height and biomass, because in ORCHIDEE they are less sensitive to local environmental conditions*

than the observed variability indicates (cf. the relatively narrow ranges of height, diameter and biomass over the entire region given in Table 5, and the smaller variability in AGB in ORCHIDEE compared to observations in Figure 3). Notably, CO_2 flux-variables, with a stronger variability based on site conditions, were carefully evaluated at several locations to avoid site-specific biases. While this limited reproduction of spatial variability justifies the methodological choices in the present work, it calls for an enhanced representation of environmental dependencies in ORCHIDEE shrub parameter calculations in future work."

4. L318-322. The writing in this paragraph is somewhat unclear. Please clarify how the three datasets were selected and what distinct aspects they represent, such as ground-truth data versus data-fusion products. In the later discussion section, the comparison with these datasets could also mention a bit more on uncertainty from the observations/ products.

We thank the reviewer for this constructive feedback, which highlights an important aspect of our benchmarking strategy. We agree that clarifying the distinct nature of these datasets and discussing their inherent uncertainties strengthens the manuscript.

1. Clarification on Dataset Selection: Our rationale for selecting these three specific datasets was deliberate: we aimed to evaluate the model against an ensemble of independent, methodologically distinct products. Rather than relying on a single estimate of regional GPP, we selected datasets that represent fundamentally different ways of quantifying the same metric. Specifically, we chose FLUXCOM to represent machine-learning upscaling of eddy-covariance data, Virkkala et al. (2021) for statistical upscaling of synthesised CO_2 flux data, and CARDAMOM to represent a model-data-fusion approach. This approach ensures that our model evaluation is not biased by the specific assumptions, algorithms, or spatial scaling limitations inherent to any single observational methodology. We will rewrite the paragraph at L.318-322 to explicitly categorise each dataset and outline the distinct perspective it provides as follows:

"To ensure a robust evaluation of regional GPP, we benchmarked our simulations against an ensemble of three independent data-based products, all remapped to the ORCHIDEE grid. These were deliberately chosen to represent distinct methodological approaches: (1) the FLUXCOM product (Jung et al., 2019), representing a machine-learning approach that upscales FLUXNET eddy-covariance observations using remote-sensing and meteorological predictors; (2) the Virkkala et al. (2021) dataset, representing a statistical modelling approach to upscale synthesised CO_2 flux data; and (3) CARDAMOM (Hugelius et al., 2024; López-Blanco et al., 2019), representing a model-data-fusion approach that assimilates remote sensing and gridded datasets into a structurally simple ecosystem carbon model, allowing for observation-constrained parameter retrieval. By utilising this structurally diverse ensemble, we mitigate the risk of evaluating our model against the biases or scaling assumptions inherent to any single upscaling method."

2. Addressing Observational Uncertainty in the Discussion: We entirely agree with the reviewer's point regarding the uncertainty of the benchmark data itself. Observational and data-fusion products are often treated as absolute benchmarks, yet they carry significant uncertainties stemming from sensor limitations, gap-filling algorithms, spatial heterogeneity, and upscaling errors. In fact, as noted in Section 4.4, the spread among these upscaled products implies that absolute GPP bias cannot be uniquely diagnosed from any single regional benchmark. The divergence among the three independent datasets provides a practical boundary of this observational uncertainty. We will expand the discussion to emphasise that model-data mismatch must be viewed through the lense of these observational constraints.

5. L336 and L345: Elaborate more on why the simulated range is much smaller than the observed range (Figure 3) and discuss the implications.

That is an interesting point, we thank the reviewer for raising this. The simulated ranges in ORCHIDEE of aboveground biomass (as well as height, diameter, belowground biomass) show a narrower range than the observations. This suggests that the relevant processes in ORCHIDEE are somewhat less sensitive to environmental conditions than the observed variability indicates. The limitations in reproducing local heterogeneity are a known weakness in large-scale modelling in general. ORCHIDEE is working on addressing them by introducing further dependencies on environmental conditions (e.g. climate or soil type) in parameter calculations, such as making the calculation of tree height dependent on precipitation (as mentioned in the discussion, L.471). The limited variability in shrub height and biomass indicates that the model's shrub representation would benefit from dynamic parameter calculations to capture the full range of observed variability. We hope to address this in future model developments and have now explicitly addressed it in the revised discussion from L.456 on, as shown already above:

> *"Parameterisation in a single grid cell was further deemed sufficient for the targeted variables like height and biomass, because in ORCHIDEE they are less sensitive to local environmental conditions than the observed variability indicates (cf. the relatively narrow ranges of height, diameter and biomass over the entire region given in Table 5, and the smaller variability in AGB in ORCHIDEE compared to observations in Figure 3). Notably, CO_2 flux-variables, with a stronger variability based on site conditions, were carefully evaluated at several locations to avoid site-specific biases. While this limited reproduction of spatial variability justifies the methodological choices in the present work, it calls for an enhanced representation of environmental dependencies in ORCHIDEE shrub parameter calculations in future work."*

6. Figure 4: It appears that the uncertainty range is truncated for GPP and NEE in some cases.

We thank REF#1 for pointing this out. We adapted the figure margins accordingly, so that the full range of standard deviation is visible now.

7. L405-406: Explain why the modeled trends show a clear increase over time for biomass (Figure 6) and GPP (Figure 7), whereas this pattern is much less evident in the datasets used for comparison.

We thank the reviewer for raising that interesting point. The increase in aboveground biomass over the shown period can be almost exclusively attributed to the boreal tree PFTs dynamics. Since the PFT maps prescribe no noteworthy increase in tree cover over the analysed time period (1992-2020), the trend indicates a growth response to warming that is stronger than observed in the data products. This may partly be explained by the absence of forest disturbance mechanisms in the simulations, which should provide a balance to the warming growth response. In ORCHIDEE, forest disturbances through fire, wind, bark beetles and anthropogenic management are implemented (Chen et al., 2018; Marie et al., 2024; Yue et al., 2014; Yue et al., 2018), but were not activated in the present setup, since the focus was on the implementation of new shrub PFTs.

8. Table 6. Consider adding a column that specifies the thresholds or definitions associated with these shrub classes, since these definitions can differ substantially across ESMs.

This is a great suggestion. We will add a column with information on the shrub PFTs to the table accordingly, as far as it is available for the listed models.

Point-by-point response to Referee #2 comments

RC2: 'Comment on egusphere-2026-1071', Wu Sun, 21 Apr 2026

The manuscript by Kirchner et al. implemented a parsimonious, fine-grained classification of shrub plant functional types (PFT) for the global Arctic–boreal region in the land surface model ORCHIDEE, meticulously identified and calibrated relevant parameters, and evaluated model simulations of biomass density and carbon fluxes against benchmark data products. The shrub PFT implementation is nicely done and well documented. In my opinion, two highlights of the work include: a parsimonious classification of shrub PFTs to capture the main variability of functional diversity among dominant Arctic–boreal shrub species while not overcomplicating it, and the systematic approach toward parameter selection and parameter optimization using ORCHIDAS, which puts the work on a rigorous footing. The manuscript is overall well written, although some clarifications would be helpful. My comments are mostly minor, listed as follows:

We thank reviewer Wu Sun (hereafter referred to as REF#2) for their constructive and thorough feedback on the manuscript. All comments are addressed below and implemented in the manuscript, strengthening it further.

1. A limitation of this work is that the partition from the total shrub coverage into the three shrub PFTs (tall deciduous, short deciduous, and short evergreen) is simplistic, with latitude as the only factor determining shrub PFT fractions. Empirical knowledge tells us that there is a strong west–east gradient in the northern limit of treelines over North America: the same latitude in southern Alaska where boreal forest dominates would be tundra around the Hudson Bay. There is also an east–west gradient in shrub coverage going from Finland to Norway. Potential misclassification caused by the latitude-only shrub distribution (Eqs. 5 and 7) may need to be discussed in comparison with the original CAVM and Macander et al. (2017) maps. In addition, the empirical representation of shrub PFT fractions seems ill-suited for CMIP-style future projections where PFT fractions are forced by evolving climate.

We agree with REF#2 that the partitioning of shrub coverage into PFTs is suboptimal. This is a direct consequence of the limited availability of shrub cover mapping products, and we did the best we could with the limited data available. Notably latitude is not the only determinant of shrub cover partitioning: within the CAVM region deciduous and evergreen shrub cover are derived from the different shrub tundra types mapped (P1, P2, S1 to evergreen, S2 to deciduous). Since CAVM does not explicitly distinguish between evergreen and deciduous shrubs, this assumption was necessary. The latitude-dependent relationship described in Equation 5 is only applied for gap-filling. South of the CAVM region we decided to partition evergreen and deciduous shrub following a spatially uniform fraction derived from a detailed mapping product over Alaska (Macander et al., 2022). We consciously decided to keep this partition simple, since we cannot extrapolate shrub distribution patterns from Alaska to the rest of the panarctic region.

We wonder, how is “a strong west–east gradient in the northern limit of tree lines over North America” related to the partitioning of different shrub types? Note that the total shrub cover, as well as the tree line is prescribed directly from Harper et al. (2023) in our maps.

We further agree that the prescription of vegetation cover presents a limitation for future projections. Future projections would benefit from dynamic vegetation cover, evolving in response to climate, as calculated by dynamic vegetation modules, as included in some land surface models (e.g. LPJ-GUESS). The absence of a dynamic vegetation module in the current ORCHIDEE version is a generic limitation of the model, which is being addressed with ongoing developments in a parallel modelling effort. Our shrub implementation operates under the restrictions of the ORCHIDEE model and thus requires empirical vegetation cover maps. Note that the present manuscript consciously focuses on the historical period (where prescribed vegetation cover changes yearly in accordance with the Harper et al. (2023) maps), and does not include future projections. We are currently preparing a manuscript on future simulations with the new shrub PFTs in ORCHIDEE, focusing exactly on those limitations of prescribed vegetation cover.

2. This study claims to have obtained “regionally representative shrub traits.” If this means in a narrow sense that simulated aboveground biomass and GPP align well with pan-Arctic/boreal data products, then yes. But this claim is tenuous if it refers to parameter optimization, given that “sensitivity tests and optimisation were performed … primarily using the grid cell around Toolik Lake, Alaska” (L176–177) and that the GPP targets are based on data from only three eddy covariance sites (Table 2). Further clarifications on regional representativeness are needed. Besides, it’s unclear which target variables from Toolik Lake are used for the optimization.

We thank the reviewer for this relevant yet constructive critique. We would like to clarify that even though parameterisation and optimisation took place mainly in a single grid cell, this is not to be confused with site-level parameterisations as in many other model studies, since they were based on synthesised observations from across the pan-Arctic region, not on site-level observations. We tested different sites during manual sensitivity tests and optimisation, and found little sensitivity to the choice of grid cell, also due to the limited dependency on local environmental conditions in the targeted ORCHIDEE parameterisations. The fact that simulated shrub characteristics match synthesised observations not only at the optimisation grid cell, but across the pan-Arctic region, further strengthens our claim at having obtained regionally representative parameterisations.

To address the reviewer’s comment, we have:

a.) Expanded the analysis to include three more EC sites for GPP calibration targets and evaluation, now including all suitable sites from the FLUXNET and Ameriflux databases (for more details see answer to point 6 below).

b.) Clarified the approach in line (L.) 176 and following:

"Sensitivity tests and optimisation were performed separately for each shrub PFT at grid cell level, primarily using the 2° x 2° grid cell around Toolik Lake, Alaska (67.6°-69.6°N, 148.6-150.6°W). This area was found suitable both because it includes a well-studied diverse shrub tundra environment, and because initial ORCHIDEE simulations across the high-latitude region showed that it represents average growing conditions for the region. Different grid cells were tested during the development process, and parameterisation proved to be only weakly sensitive to grid cell-choice. While parameterisation and optimisation were limited to a single grid cell for computational reasons, they were based on regionally representative targets derived from synthesised panarctic observations."

c.) Explicitly addressed the question of parameterisation and optimisation in a single grid cell in the discussion, L. 456:

"Contrary to the site-level calibrations in many other shrub implementations (e.g. Wolf et al. 2008; Meyer et al. 2021; Sulman et al. 2021), the ORCHIDEE shrub PFT classification and parameterisation presented here are constrained by synthesised observations spanning the entire high-latitude region, increasing its transferability and limiting site-specific biases (Alton et al. 2011; Kuppel et al. 2014). This is in line with recent work by Murphy et al. 2025, highlighting the limitations of model parameterisations based on plot-level observations from a single site. Thus, the methodological choices, including the shrub PFT classification scheme, calibration targets and vegetation distribution, are applicable beyond the ORCHIDEE model.

While parameter selection and optimisation were mainly carried out in a single grid cell due to computational limitations, the results indicate that parameterisations can be considered regionally representative, since simulated shrub characteristics match synthesised observations not only in the optimisation grid cell, but across the entire region. Parameterisation in a single grid cell was further deemed sufficient for the targeted variables like height and biomass, because in ORCHIDEE they are less sensitive to local environmental conditions than the observed variability indicates (cf. the relatively narrow ranges of height, diameter and biomass over the entire region given in Table 5, and the smaller variability in AGB in ORCHIDEE compared to observations in Figure 3). Notably, CO_2 flux-variables, with a stronger variability based on site conditions, were carefully evaluated at several locations to avoid site-specific biases. While this limited reproduction of spatial variability justifies the methodological choices in the present work, it calls for an enhanced representation of environmental dependencies in ORCHIDEE shrub parameter calculations in future work."

d.) The question about target variables in the optimisation was addressed by explicitly adding the optimisation target variables in L. 252, see answer to point 7 below.

3. L57–62: It's worth discussing in which aspects this shrub PFT implementation differs from Druel et al. (2017) and why the implementation presented in this study makes better sense.

As explained in L.57–62, the shrub implementation by Druel et al. (2017) was based on an old, fundamentally different version of ORCHIDEE. It is not possible to integrate their work into the current version of ORCHIDEE, necessitating an entirely new implementation of shrub PFTs, as done in this manuscript. Since they are based on essentially different models, no direct comparison of the implementations is possible. Regarding the choice of shrub PFTs, Druel et al. (2017) included only one type (broadleaf deciduous shrubs). The arguments for a more detailed shrub classification are discussed in the manuscript already (e.g. L.100 and following, L.445 and following).

4. L85: "adapted to harsh high-latitude conditions" - This statement only refers to Arctic/boreal shrubs because shrublands are common in mediterranean, xeric, and alpine ecosystems.

We thank REF#2 for pointing out this inconsistency. We adjusted the sentence to include shrublands in general and made the focus on high-latitude shrubs more explicit:

L.85: *"Shrubs are low-stature woody plants - both deciduous and evergreen - adapted to limited growing conditions, such as in arid or cold regions (CAVM Team, 2024). In high-latitude regions, shrubs typically have lower height and stem diameter and therefore lower biomass, relative to boreal trees, [...]"*

5. L99–104: Are there evergreen tall shrubs in the Arctic?

To the best knowledge of the authors, evergreen tall shrubs are not generally treated as a distinct pan-Arctic shrub class and were therefore not included in the classification. This is supported by the classifications in other models and data products, frequently including tall deciduous shrubs, but not tall evergreen shrubs, such as Macander et al. (2017), or the four-tier hierarchical PFT classification scheme suggested by Salmon et al. (2025).

6. Table 2: Is there a scientific reason to use different partitioning methods for different sites? How does the discrepancy between daytime and nighttime partitioning affect the calibration?
   Also, I have the same question as Reviewer #1 - Why did the authors select only three sites? There are more AmeriFlux and ICOS shrub sites in the high latitude (for example, US-EML).

We thank the reviewer for this important comment and address the second part of it first.

To address the second part of the comment, as well as comments by the other two referees, we have updated the selection to include in total six sites, clarified the selection criteria in the manuscript, and updated the relevant figures, as detailed below:

> L. 144: *"Data documenting the carbon fluxes between the atmosphere and shrub vegetation were extracted from the FLUXNET2015 (Pastorello et al., 2020) and AmeriFlux (Chu et al., 2023) datasets. We selected sites classified as open or closed shrubland (OSH/CSH), located above 50°N and with the full FLUXNET-standard post-processing available. Based on the available vegetation descriptions, further sites were excluded, including burn sites with fire succession vegetation, and one tussock-tundra site (US-ICt), resulting in six selected shrub sites (cf. Table 2)."*

To address the first part of the comment: Both nighttime (Reichstein et al., 2005) and daytime (Lasslop et al., 2010) flux partitioning are standard eddy-covariance post-processing approaches. We originally chose the daytime method, where NEE is fitted using a light-response formulation combined with temperature-dependent respiration (see López-Blanco et al. (2017) for further details), since this approach is particularly useful at high-latitude sites, where true nighttime observations can be sparse or absent during summer. However, the northernmost site (RU-Cok, at 70.8°N) showed strong differences in GPP estimates between the two methods (see Table 1), due to positive GPP values throughout the winter season in the daytime method, while the nighttime partitioning yielded more realistic estimates, in line with the estimates of the other sites.

We note that the choice of partitioning method can influence the inferred GPP/Reco split. This is consistent with previous work, including López-Blanco et al. (2017), which showed that different partitioning approaches can lead to different estimates and associated uncertainties in Arctic CO_2 fluxes.

For the sake of consistency, and because Reviewer #2 raised the issue of using different partitioning methods across sites, we have revised the analysis to apply a single partitioning method consistently across all sites. This was done to improve comparability among sites, while acknowledging that uncertainty related to the partitioning approach remains and that an exhaustive assessment of partitioning uncertainty is beyond the scope of the present study.

After comparing both partitioning methods, we applied the nighttime partitioning to all six sites. As shown in Table 1, the nighttime method provided a more realistic estimate for the RU-Cok site, while the estimated annual GPP was similar at the other five sites for both methods, albeit slightly lower using the nighttime method.

Table 1: Median (interquartile range) annual GPP in $gCm^{-2}yr^{-1}$ for nighttime and daytime flux partitioning methods

	Five EC sites (excl. RU-Cok)	RU-Cok
Nighttime	335.73 (208.90-413.05)	360.72 (303.02-519.78)

Daytime	353.60 (234.23-494.48)	725.71 (466.71-866.64)

7. L252: What are the target variables in Y? Please spell them out. Also, better use an upright capital "T" to indicate matrix transpose in Eq. (4).

We thank REF#2 for this clarifying question. The description was adjusted to include *"The observational targets (Y) included height, peak summer AGB, minimum fraction of belowground biomass, mean GPP, peak summer leaf area index, and age (ref. Table A4)."* in line 254, moving this information from the supplementary material into the main manuscript.
The equation was adjusted following the reviewers' suggestions.

8. L264: "boreal tree PFTs (80%) and boreal grass (20%)" - This feels ambiguous to me. Does this mean that a high-latitude grid cell would have 80% fractional coverage of boreal tree PFTs, and 20% fractional coverage of boreal grass? Or does it mean that among all high-latitude grid cells, 80% of them are assigned boreal tree PFTs?

Thank you for raising this ambiguity. Here, we refer to a methodological choice made by the ORCHIDEE core team and implemented in the standard version of ORCHIDEE without shrub PFTs. In the absence of explicit shrub PFTs in ORCHIDEE, shrub fractional cover from Harper et al. (2023) maps is redistributed to existing ORCHIDEE PFTs.
In practice, across the high-latitude region, this means that in every grid cell where Harper et al. (2023) identifies shrub cover, that shrub fraction is reassigned to the existing boreal PFTs. Of the shrub cover fraction, 80% is allocated to boreal tree PFTs (ORCHIDEE PFT 7, 8, or 9, depending on shrub type), and the remaining 20% is allocated to boreal grass (PFT 15).
To make this clearer, we have rephrased line 264 as follows:

> L 264: *"Prior to the implementation of shrub PFTs in ORCHIDEE, the high-latitude shrub cover fraction in Harper et al. (2023) was redistributed at grid cell level to boreal tree PFTs (80%) and boreal grass (20%)."*

9. Fig. 3: Is there a similar comparison for belowground biomass, or any related measurable belowground traits?

Great question, but since belowground biomass is more difficult to measure, fewer data products on belowground biomass are available. As described in L. 139-143, we used two synthesis datasets to derive average fractions of belowground biomass for the shrub PFTs. In our manuscript we evaluate belowground biomass implicitly by evaluating the fraction of biomass allocated belowground (in Table 5) and aboveground biomass (in Figure 3).

10. Fig. 4: This is a central figure that requires some clarifications.
    First, do we expect these sites to host only one shrub PFT? It is unclear what this comparison aims to achieve. It is true that if we restrict the grid cell to only one PFT, then PFT 18 is the obvious winner. But is reality a mixture of multiple shrub PFTs? It seems a fair comparison can only be made between observations and PFT-aggregated grid-cell-scale carbon flux estimates.
    Second, why does PFT 18 (evergreen dwarf shrubs) show a more gradual spring onset than do PFTs 16 (tall deciduous shrubs) and 17 (low deciduous shrubs)?
    In addition, consider changing "FLUXNET" to "observations" in figure annotations to avoid potential confusion (FLUXNET usually indicates the global network of flux towers).

We thank the reviewer for raising this very valid point, which has actually been thoroughly discussed during the preparation of this manuscript.
This part of the results (section 3.1) aims at evaluating the three shrub PFTs individually. The second part of the evaluation (sections 3.3 and 3.4) then focuses on the overall model outputs taking all PFT contributions together. It is important to evaluate the three shrub PFTs individually, to exclude potential biases from other PFTs. For instance, the boreal grass PFT CO_2 fluxes include biases in ORCHIDEE at the moment, but those biases should not interfere with evaluation of the shrub PFTs.

The reviewer is correct in raising that the eddy-covariance CO_2 flux measurements capture a real-world mixture of PFTs and a comparison with individual PFTs is difficult. While individual studies have conducted a footprint analysis to separate contributions from different PFTs (e.g. Ludwig et al. (2024), as pointed out by referee #3), this information is not available for the selected FLUXNET and Ameriflux sites, and thus the best we could do was carefully select sites dominated by shrub tundra and use those shrub tundra ecosystem fluxes. We do believe that there is a value in evaluating the simulated PFT fluxes against this data, but have adjusted the manuscript to make the reader more aware of the limitations of this comparison:

- Adapted L. 148 and following to clarify that the data represent shrub tundra ecosystems:

"It is important to note that due to the nature of EC flux measurements, the recorded fluxes represent the aggregate of the entire ecosystems around the tower within a variable footprint, and the data does not specify the contributions of individual PFTs. Therefore, the sites were not assigned to specific shrub PFTs, but used instead as a benchmark for shrub tundra ecosystems to calibrate and evaluate all three ORCHIDEE shrub PFTs against. Notably, the available vegetation descriptions indicate that dwarf and low shrubs seem to be more prominent than tall shrubs at most sites (e.g. van der Molen et al. (2007); NEON (National Ecological Observatory Network); cf. Table 2). Across all six sites, a range of average daily gross primary productivity (GPP) of 0.73 (0.18-2.45) $gCm^{-2}d^{-1}$ (median (interquartile range (IQR)), and average annual GPP of 336 (209-413) $gCm^{-2}yr^{-1}$ was extracted. The GPP statistics were calculated using nighttime flux partitioning, after excluding negative GPP estimates and years without positive values."

- Adapted Figure 4, as shown above, to make it clearer that the comparison happens between ORCHIDEE shrub PFT fluxes at the grid cell level and EC fluxes recording an entire tundra ecosystem at site level. We include in total six FLUXNET/Ameriflux sites, four of which are located in the same two ORCHIDEE grid cells. By reframing the figure as comparing individual PFT fluxes in a 2°x2° ORCHIDEE grid cell against composite flux estimates from individual EC tower sites, the reader should be more aware of the limitations of this comparison.

To address the second sub-question: The reason for the more abrupt spring onset in the deciduous shrub PFTs (PFT 16 and 17) compared to PFT 18 is explained in L. 360-363: *"In particular, the deciduous shrub PFTs experience a rapid increase in photosynthetic activity at the onset of the growing season - a behaviour that can be traced to model functionality where leaf onset for deciduous PFTs occurs within one single day."*.

We thank the reviewer for highlighting the potential confusion by using the FLUXNET label in Figure 4. We have changed it to the actual station names.

11. L329: What does the recruitment beta mean?

'RECRUITMENT_BETA' is an ORCHIDEE parameter controlling recruitment of new saplings based on light availability and stand density, as listed in Table 4. Through introducing young saplings into the shrub stands, it controls average plant size, as indicated in L. 329.

12. L362–363: Does ORCHIDEE ramp up Vcmax gradually after leaf onset or use a fixed Vcmax value throughout the growing season?

Vcmax is calculated dynamically throughout the growing season, as a function of leaf nitrogen content and nitrogen use efficiency. Therefore, it shows a dynamic increase and decrease throughout the course of the growing season.

13. L389: What does "forest-prescribed tundra" mean in this context?

We agree that this wording is confusing and thank the reviewer for pointing it out. We were referring to the prescription of boreal forest in place of shrubs in ORCHIDEE PFT maps without shrub PFTs, as explained above in point 8. We have changed the formulation in the manuscript (L. 389) as follows to be less ambiguous :

- From: *"This redistribution reduced the extent of forest-prescribed tundra and established the spatial baseline for subsequent changes in simulated carbon stocks and fluxes."*
- To: *"Introducing explicit shrub PFTs prevented the reassignment of shrub fractional cover to boreal tree and grass PFTs in ORCHIDEE PFT maps, thereby removing erroneously prescribed forest cover in tundra regions, and established*

the spatial baseline for subsequent changes in simulated carbon stocks and fluxes."

14. L445: "more exhaustive and complete" - This seems pretty subjective. I would say that the classification is more fine-grained, for sure.

We thank the reviewer and have changed the wording from *"more exhaustive and complete"* to *"more detailed"* to be more objective.

15. L504–506: It is also worth noting that the shrub–snow albedo feedback goes the other way: more shrub coverage prompts earlier snow melt and amplifies warming, because protruding shrub branches are darker than snow. See, for example: Domine et al. (2025) JGR-Biogeosciences, https://doi.org/10.1029/2024JG008593; Belke-Brea et al. (2020) J Climate, https://doi.org/10.1175/JCLI-D-19-0318.1.

We thank the reviewer for highlighting the complicated and contradicting processes involved in snow-shrub feedbacks. We have adjusted the text (L. 498-506) to reflect the shrub–snow albedo feedback better as follows:

"Central feedback processes to consider when modelling shrubs include their interactions with snow, impacts on soil thermal regimes and permafrost dynamics, as well as albedo changes (Wullschleger et al., 2014; Mekonnen et al., 2021). For instance, snow cover influences shrub growth, by insulating and protecting shrub buds and tissues from frost, and shaping shrub height and growth form (Myers-Smith et al., 2011; Wheeler et al., 2014). Conversely, shrub presence increases snow accumulation and reduces its compaction, with impacts on soil temperatures and permafrost through the thermal insulation of the snowpack (Myers-Smith et al., 2011; Domine et al., 2016; Heijmans et al., 2022). Depending on shrub height and stature, shrub branches protruding above the snowpack can reduce winter albedo and accelerate snow melt (Belke-Brea et al., 2020; Shuman et al., 2025). These dynamics may potentially represent feedbacks on shrub growth and climate warming, and will thus be important to capture in modelling efforts of future high-latitude ecosystem-climate interactions in ORCHIDEE"

Point-by-point response to Referee #3 comments

RC3: 'Comment on egusphere-2026-1071', Anonymous Referee #3, 21 Apr 2026

The manuscript by Kirchner et al. substantially improves the representation of tundra vegetation in a LSM by introducing three plant functional types describing Arctic-Boreal shrubs. They define shrub PFT parameters using literature estimates that are then constrained by calibrating ORCHIDEE model outputs of vegetation structure against pan-Arctic observational datasets. This is timely and important work given the current need to understand feedbacks between climate and vegetation change at high-latitudes, in particular the consequences of shrub expansion. The paper is well written and results are generally presented clearly. However, elaborating on several of the methodological choices could help readers better understand how transferable this PFT parameterization is beyond ORCHIDEE and how scalable this approach is as Arctic-Boreal datasets continue to improve in spatial resolution and coverage. Some comments are provided below to this end.

We thank the reviewer for their thorough evaluation and insightful comments. The comments were very helpful and led to substantial improvements overall. We addressed all comments in detail below.

1. L70. Including a brief framework overview of the carbon cycle processes relevant to shrubs in ORCHIDEE (e.g. expanding on the description in L155 (perhaps nearer the beginning of the methods, such as here) could help orient the reader to methodological choices further on.

We thank the reviewer for this suggestion and agree that further understanding of the ORCHIDEE model will be helpful for readers. Considering the current length of the manuscript, we suggest to include additional details on the ORCHIDEE model (including but not limited to carbon cycle processes) in the supplementary material, as suggested also in the reply to reviewer #1.

2. L170: A more detailed discussion on how targeted parameters (Table 3 and Table 4) were chosen, and in particular which were kept the same across trees-to-shrubs and across new shrub PFT's, could be helpful. Some relevant traits seemed missing: for example, Vcmax varies across evergreen and deciduous species, but seems to be the same (across trees and shrubs) for all PFT's? Leaf nitrogen and leaf carbon construction costs would also vary, which could be taken into account by the "NUE_OPT" parameter, but without more elaboration on what is going on in the model, it is somewhat unclear how these parameter changes would translate (particularly beyond the ORCHIDEE-specific framework).

We thank the reviewer for raising this question about parameter selection. As explained in lines 186-190, we focused on selected key differences between shrubs and trees (height, diameter, biomass) and tested parameters controlling relevant processes

(Carbon uptake, release, allocation, recruitment, mortality, …) for their impacts on the targeted output variables. The parameters detailed in Table 3 and 4 were selected for their specific control on our targeted outputs. It is important to emphasise that we do not claim this parameter set to be the only pathway to transform trees PFTs into shrubs PTFs, but rather one approach that proved successful in achieving the intended shrub characteristics.
Please note that parameters not listed were kept at the same values as in the underlying tree parameterisations (PFT 5 and PFT 8). Due to the high number of parameters in the ORCHIDEE model (~375 PFT-specific parameters alone), it was not deemed fruitful to list them all in the manuscript.
Since deciduous and evergreen shrubs were derived from deciduous and evergreen trees respectively, the traits mentioned by the reviewer vary indeed between deciduous and evergreen species, but not between shrubs and trees. For instance, Vcmax is calculated in ORCHIDEE as a function of nitrogen use efficiency and leaf nitrogen content, all three of which are lower for evergreen compared to deciduous species.
Maintaining most parameters at the same values for tree and shrub PFTs is based on assuming that the underlying processes (e.g. photosynthesis) follow the same principle for trees as for shrubs. We agree with the reviewer that implementing further tree-shrub differences will be important future work, as has already been highlighted in the discussion (from L. 496).

To address the reviewers' comments, we have:

a) Clarified the impact of nitrogen use efficiency on Vcmax in L.244:

"For evergreen dwarf shrubs, originally based on a temperate tree PFT, we adapted key physiological parameters to reflect boreal conditions. Specifically, the reference temperature used to calculate critical leaf age (leaf_age_crit_tref) was reduced and nitrogen use efficiency was increased via the parameter nue_opt~~,~~ to stabilise growth in nitrogen-limited high-latitude locations, in line with estimated values reported by Kattge et al. (2009). Nue_opt represents the photosynthetic nitrogen use efficiency, defined as the carboxylation capacity at 25°C (V_{max}^{25}) per leaf nitrogen content, based on Kattge et al. (2009), and is used for calculation of V_{max}^{25} after accounting for sugar loading and leaf efficiency."

b) Explicitly stated in Table 4 caption that all parameters not listed kept the same values of the respective tree PFTs 8 (for deciduous) and 5 (for evergreen shrubs):

"Table 4. Selected PFT-specific model parameters, their function in ORCHIDEE, original values in the underlying forest PFTs and recalibrated values for the new shrub PFTs. The given values are the final values after parameter optimisation, except the parameter values marked with an asterisk, which were set manually. '-' indicates that a parameter value was not changed from the baseline. PFT 8, 16 and 17 are all based on metaclass (MTC) 8, while PFT 5 and 18 are based on MTC 5. <u>All PFT-specific parameters not listed here were kept at the same values for shrub PFTs as in the underlying forest PFTs (5 and 8, for deciduous and evergreen respectively).</u>"

3. L175: Even after reading the appendix, I remained unclear about this process.
    - How sensitive were the PFT model parameter calibrations to the choice of grid cell for optimization?
    - Was a certain distribution of each shrub PFT assigned to this cell during the calibrations based on vegetation maps of the region? Or was the whole cell separately assigned to each PFT to run the calibrations?
    - While shrubs are prevalent in the Toolik area, a large majority of the area is dominated by mixed or non-shrub vegetation cover. Comparing how well observational targets identified in Section 2.1.2 match the substantial amount of vegetation biomass (e.g. shrub biomass maps, such as Greaves et al. 2018, https://doi.org/10.1016/j.rse.2016.07.026) seems perhaps relevant.
    - Perhaps more relevant to the method, how much did the optimization procedure change the manual estimates of parameters and in which directions?

We thank the reviewer for their interest in the optimisation procedure and provide further information below:

- Regarding the choice of grid cell, we did not explicitly quantify the sensitivity of the optimisation to grid-cell selection in the original manuscript. However, previous experience from manual parameter tuning indicated little dependence on the choice of grid cell used, and additional optimisation tests for selected parameters at grid cells located at different latitudes gave consistent results. This is also consistent with the fact that the variation in simulated shrub characteristics across the test environments was smaller than the variability in the observational targets (e.g. Figure 3), suggesting a relatively limited influence of local environmental conditions on the parameterised shrub characteristics.
- For the calibration and optimisation, a fixed equal fraction of all high-latitude PFTs was prescribed within the selected grid cell, but only the outputs of the shrub PFT under consideration were evaluated. In other words, the calibration was performed separately for each shrub PFT, while retaining a common background vegetation configuration in the grid cell.
- We also agree that the use of a grid cell around Toolik Lake could have given the impression of a site-level calibration, although this was not the intention. The parameterisation and optimisation were explicitly designed to derive regionally representative parameter values, using target values based on synthesised observations across the broader pan-Arctic region. Toolik Lake was used only as the location of the 2° × 2° grid cell for computational reasons, not as a site-specific calibration target. For this reason, a direct evaluation against Toolik-specific vegetation or shrub biomass products, such as Greaves et al. (2018), would not be fully aligned with the purpose of this analysis. To make this clearer, we have revised the text in L. 176 ff. as follows:

    "Sensitivity tests and optimisation were performed separately for each shrub PFT at grid cell level, primarily using the 2° x 2° grid cell around Toolik Lake, Alaska (67.6°-69.6°N, 148.6-150.6°W). This area was found suitable both because it

includes a well-studied diverse shrub tundra environment, and because initial ORCHIDEE simulations across the high-latitude region showed that it represents average growing conditions for the region. Different grid cells were tested during the development process, and parameterisation proved to be only weakly sensitive to grid cell-choice. While parameterisation and optimisation were limited to a single grid cell for computational reasons, they were based on regionally representative targets derived from synthesised panarctic observations."

- We agree that clarifying the role of the optimisation step and its influence on the final parameter values is important. However, in this study, the optimisation was intentionally designed as a fine-tuning step within a narrowly constrained parameter space, rather than as a broad explanatory search. As a result, most parameters changed only modestly, and reporting individual parameter shifts would provide limited additional insights beyond the final parameter values and their prescribed ranges. This was expected because the optimisation was restricted to relatively narrow parameter ranges defined from prior manual sensitivity analyses. These manual tests had already identified a balance between reproducing the intended shrub characteristics and maintaining realistic, stable long-term growth. This was particularly important because the shrub PFTs were sensitive to parameter changes, and larger shifts often resulted in unstable or unrealistic long-term behaviour. The optimisation should therefore be understood primarily as a fine-tuning step, rather than as a broad unconstrained search across parameter space.

  We also note that summarising parameter changes from the manual estimates to the final values is not entirely straightforward, because the parameterisation process was highly iterative. It involved several rounds of manual adjustment and optimisation, and was carried out across successive ORCHIDEE developments to remain consistent with ongoing model updates. In addition to the prior values themselves, the optimisation outcome was also sensitive to the prescribed parameter ranges that defined the search space, and these ranges were likewise informed by manual testing. Overall, the final parameter values should therefore be understood as the result of an iterative process in which manual sensitivity analysis provided the main constraint, and ORCHIDAS was used primarily to fine-tune and validate those manually derived values. In our experience, finding a stable balance while adapting ORCHIDEE's woody tree-based functionality to represent small shrubs was better achieved through iterative manual testing followed by targeted optimisation than through a broad Monte Carlo search alone.

4. L205: What is the relationship between PFT height and stand height? Was stand height (or some combination of pipe_tune2) used to constrain the height targets in Table 1?

PFT height is equal to the average stand height. In ORCHIDEE, tree (and now also shrub) demographic processes are represented through three diameter classes, so the 'stand'

of each woody PFT in a grid cell includes three different sizes. This information was included in the additional ORCHIDEE information added to the supplementary material as included above.

5. L244: Is NUE the only aspect in which nitrogen dynamics interface with shrub PFT traits? In general, how ecosystem impacts of shrubs beyond their contribution to GPP is not particularly clear -- are these ecosystem-feedback processes implemented in the model and carried over from grass or tree PFT's or not yet implemented? For example, are shrubs altering albedo like trees (PFT5), modifying litter inputs, altering soil temperature and moisture, or nitrogen availability? Parameterizing PFT's based on mainly GPP and standing biomass alone is a worthy task within itself, but it would still be useful to clarify which shrub-feedback process are and are not included.

We thank the reviewer for this comment. Since the shrub implementation relies on the functionality of trees in ORCHIDEE, every parameter and function not explicitly mentioned in the manuscript is carried over from the tree PFTs 8 (for deciduous shrubs) and 5 (for evergreen shrubs). Therefore, shrubs impact albedo, litter inputs, soil carbon, soil temperature and moisture, and nitrogen dynamics based on the same processes as tree PFTs. As pointed by the referee, implementing key shrub-tree differences, such as morphological differences, nitrogen use efficiency, carbon uptake, and belowground carbon allocation, already alters certain environmental interactions and ecosystem feedbacks for shrubs compared to trees. This includes, for instance, roughness length, which depends on plant height. We agree with the reviewer that additional shrub-specific feedback processes could further improve the representation and should be included in future work, as is already noted in the discussion (from L. 496). All the above considerations are already included in the discussion in lines 485 – 498.

6. Table 3: Should there be a column for each shrub PFT? Combining with some higher-level info from Table 4 seems nice here.

We thank the reviewer for their suggestions for Table 3. We have chosen not to include a separate column for each shrub PFT, because the direction of change of the parameters is identical for each PFT, so we decided to keep it simple for ease of reading. Table 4 then provides the detailed values for each shrub PFT. We consciously kept Table 3 simple, using it to introduce the targeted variables and their direction of change, but not including parameter values, since those are listed in Table 4 to avoid overwhelming the reader with duplicate information.

7. Table 4: A column w/more general names (or translations to related metrics) could help increase reach of results. For example, if possible to translate outputs from the pipe-allometry sub-model back into the power-law allometric parameters typically used for Arctic shrubs (e.g. as in Berner et al. 2015), this would be of general interest. While parameters like "alpha_self_thinning" are specific to ORCHIDEE model infrastructure, model outputs such as maximum stand density seem more relevant (and transferable) to report even if not the exact parameters tuned.

We thank the reviewer for this helpful suggestion and agree that Table 4 can be more accessible by including more general or ecologically interpretable descriptors alongside the ORCHIDEE-specific parameter names. Our primary intention in the current table was to document the parameters that were directly modified in the model, for transparency and reproducibility. At the same time, we recognise that some of these names are specific to the ORCHIDEE model infrastructure and are therefore not immediately transferable to a broader modelling or ecological audience.
However, there is a practical limit to how much additional detail can be included in the manuscript without overcomplicating the presentation or distracting from the main text. For instance, some of the metrics suggested by the reviewer, such as maximum stand density or recruitment density, are not fixed absolute quantities, but depend on other variables such as stem diameter or light availability. For this reason, they cannot always be represented by a single value in the table.

We note that most of the parameters listed in Table 4 already have a direct physical meaning. For example, PIPE_TUNE2 represents plant height (m) at a stem diameter of 1 m, and K_LATOSA represents the ratio of leaf area and sapwood area. The parameters that are less directly interpretable are the self-thinning, recruitment, and GDDNCD_OFFSET parameters. For these, Section 2.2.3 already provides the equations needed to translate them into more general, physically meaningful, and translatable metrics, such as maximum stand density and the temperature-sum threshold for budburst.

To make this interpretation easier for interested readers, we have clarified this point in the text and have added a reference to the calculation of the canopy gap fraction, as well as the default values of gddncd_ref, gddncd_curve, and gddncd_offset in Section 2.2.3 so that these derived quantities can be calculated directly:

> L. 215: "Recruitment was increased by tuning the PFT-specific parameters *(recruitment_alpha, recruitment_beta*) controlling yearly recruits ($d^{ind,rec}$; m^{-2}) as a function of light availability (via $f^{Pgap,trees,season}$, calculated as in Naudts et al. (2015)) and stand density (d^{ind})"
>
> L. 237: "These parameters are not PFT-specific and were originally calibrated against satellite phenology data at global scale (Botta et al., 2000), with the default values *gddncd_ref*: 964, *gddncd_curve*: 0.0058, *gddncd_offset*: 12.8."

8. L315: Newer high-resolution map estimates of tundra plant biomass, divided into woody and non-woody, now exist for the tundra region (e.g. https://arcticdata.io/catalog/view/doi:10.18739/A2NS0M06B). It could be useful to check whether fractional cover assigned to shrub PFT's match these products at least for 2020 to test/justify the latitude-based approach.

We thank the reviewer for this comment and agree that a more detailed evaluation of the prescribed shrub cover distribution would be very highly relevant. In fact, we are currently carrying out a parallel analysis, separate from the present manuscript, that evaluates prescribed shrub cover and simulated biomass using the products of Orndahl et al. (2025). That work also includes a more detailed discussion of available shrub cover products and their limitations for deriving PFT maps for ORCHIDEE. Because we considered this analysis too detailed to include in the present manuscript without detracting from its main focus, we are preparing it as a separate manuscript.

9. Figure 4. The choice of these three shrub sites is somewhat unclear, particularly given they demonstrate pretty same patterns. Incorporating data from other sites, including EC sites that have known distributions of vegetation (that could be mapped to PFT 7-18, e.g. such as vegetation maps and fluxes from Ludwig et al. 2024, https://bg.copernicus.org/articles/21/1301/2024/bg-21-1301-2024.html), seems potentially more useful than three shrub sites.

We thank the reviewer for raising this concern. To address this comment, and similar comments of REFs #1 and #2, we have revisited our selection and included three further sites into the analysis, including recently published data. Now the manuscript includes all sites suitable to represent high-latitude shrub vegetation according to our criteria from the FLUXNET and Ameriflux datasets at the time of writing.

While we agree with REF#3 that a footprint analysis separating the contribution of individual PFTs to the recorded fluxes, as done in Ludwig et al. 2024, would indeed be valuable for model evaluation, this information is not available for the selected FLUXNET and Ameriflux sites, and such an analysis would be outside the scope of this manuscript. In line with our approach to use pan-Arctic datasets for model evaluation, we have carefully selected all sites dominated by shrub tundra from the FLUXNET and Ameriflux datasets.

We have specified the site selection criteria and acknowledged the limitation of this data in L.144 and following, and updated Table 2, and Figures 1, 4, B2 and B3 to include the additional sites.

The additional sites show similar patterns as already described in the manuscript, both in the flux exchange magnitude and seasonality, and the degree of agreement with the ORCHIDEE flux estimations at the respective locations. Their patterns and fits with the simulated fluxes are added to the results in L.346, without changing the message. We have added a comparison between average GPP from all six sites and across the entire tundra region for all ORCHIDEE shrub PFTs.

References:

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. Progress in photosynthesis research: volume 4 proceedings of the VIIth international congress on photosynthesis providence, Rhode Island, USA, august 10–15, 1986,

Belke-Brea, M., Domine, F., Barrere, M., Picard, G., & Arnaud, L. (2020). Impact of Shrubs on Winter Surface Albedo and Snow Specific Surface Area at a Low Arctic Site: In Situ Measurements and Simulations. *Journal of Climate*, *33*(2), 597-609. https://doi.org/https://doi.org/10.1175/JCLI-D-19-0318.1

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., & Bopp, L. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010.

Box, E. O. (1996). Plant functional types and climate at the global scale. *Journal of Vegetation Science*, *7*(3), 309-320.

Chen, Y. Y., Gardiner, B., Pasztor, F., Blennow, K., Ryder, J., Valade, A., Naudts, K., Otto, J., McGrath, M. J., Planque, C., & Luyssaert, S. (2018). Simulating damage for wind storms in the land surface model ORCHIDEE-CAN (revision 4262). *Geosci. Model Dev.*, *11*(2), 771-791. https://doi.org/10.5194/gmd-11-771-2018

de Rosnay, P., & Polcher, J. (1998). Modelling root water uptake in a complex land surface scheme coupled to a GCM. *Hydrology and earth system sciences*, *2*(2/3), 239-255.

Ducoudré, N. I., Laval, K., & Perrier, A. (1993). SECHIBA, a new set of parameterizations of the hydrologic exchanges at the land-atmosphere interface within the LMD atmospheric general circulation model. *Journal of Climate*, 248-273.

Farquhar, G. D., von Caemmerer, S. v., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta*, *149*(1), 78-90.

Forbes, T., Rossiter, D., & Van Wambeke, A. (1987). *Guidelines for evaluating the adequacy of soil resource inventories* (Vol. 4). Cornell University, Department of Agronomy.

Hugelius, G., Ramage, J., Burke, E., Chatterjee, A., Smallman, T. L., Aalto, T., Bastos, A., Biasi, C., Canadell, J. G., & Chandra, N. (2024). Permafrost region greenhouse gas budgets suggest a weak CO2 sink and CH4 and N2O sources, but magnitudes differ between top-down and bottom-up methods. *Global Biogeochemical Cycles*, *38*(10), e2023GB007969.

Kattge, J., Knorr, W., Raddatz, T., & Wirth, C. (2009). Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology*, *15*(4), 976-991.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., & Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1). https://doi.org/https://doi.org/10.1029/2003GB002199

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., & Wohlfahrt, G. (2010). Separation of net ecosystem exchange into assimilation and

respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology*, *16*(1), 187-208.

López-Blanco, E., Exbrayat, J.-F., Lund, M., Christensen, T. R., Tamstorf, M. P., Slevin, D., Hugelius, G., Bloom, A. A., & Williams, M. (2019). Evaluation of terrestrial pan-Arctic carbon cycling using a data-assimilation system. *Earth System Dynamics*, *10*(2), 233-255.

López-Blanco, E., Lund, M., Williams, M., Tamstorf, M. P., Westergaard-Nielsen, A., Exbrayat, J.-F., Hansen, B. U., & Christensen, T. R. (2017). Exchange of CO 2 in Arctic tundra: impacts of meteorological variations and biological disturbance. *Biogeosciences*, *14*(19), 4467-4483.

Ludwig, S. M., Schiferl, L., Hung, J., Natali, S. M., & Commane, R. (2024). Resolving heterogeneous fluxes from tundra halves the growing season carbon budget. *Biogeosciences*, *21*(5), 1301-1321.

Macander, M. J., Frost, G. V., Nelson, P. R., & Swingley, C. S. (2017). Regional Quantitative Cover Mapping of Tundra Plant Functional Types in Arctic Alaska. *Remote Sensing*, *9*(10), 1024. https://www.mdpi.com/2072-4292/9/10/1024

Marie, G., Jeong, J., Jactel, H., Petter, G., Cailleret, M., McGrath, M. J., Bastrikov, V., Ghattas, J., Guenet, B., Lansø, A. S., Naudts, K., Valade, A., Yue, C., & Luyssaert, S. (2024). Simulating Ips typographus L. outbreak dynamics and their influence on carbon balance estimates with ORCHIDEE r8627. *Geosci. Model Dev.*, *17*(21), 8023-8047. https://doi.org/10.5194/gmd-17-8023-2024

Murphy, B. A., Sulman, B. N., Yuan, F., Salmon, V. G., Yang, D., Kumar, J., Dengel, S., Herndon, E., Fettrow, S., Brown, C., Torn, M. S., Chafe, O. E., Pegoraro, E. F., & Iversen, C. M. (2025). Integrating Characteristic Arctic Vegetation in a Land Surface Model Improves Representation of Carbon Dynamics Across a Tundra Landscape. *Journal of Geophysical Research: Biogeosciences*, *130*(12), e2025JG009039. https://doi.org/https://doi.org/10.1029/2025JG009039

Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., Bellasen, V., Berhongaray, G., Bönisch, G., & Campioli, M. (2015). A vertically discretised canopy description for ORCHIDEE (SVN r2290) and the modifications to the energy, water and carbon fluxes. *Geoscientific Model Development*, *8*(7), 2035-2065.

Orndahl, K. M., Berner, L. T., Macander, M. J., Arndal, M. F., Alexander, H. D., Humphreys, E. R., Loranty, M. M., Ludwig, S. M., Nyman, J., Juutinen, S., Aurela, M., Mikola, J., Mack, M. C., Rose, M., Vankoughnett, M. R., Iversen, C. M., Kumar, J., Salmon, V. G., Yang, D., . . . Goetz, S. J. (2025). Next generation Arctic vegetation maps: Aboveground plant biomass and woody dominance mapped at 30 m resolution across the tundra biome. *Remote Sensing of Environment*, *323*, 114717. https://doi.org/https://doi.org/10.1016/j.rse.2025.114717

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., & Granier, A. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, *11*(9), 1424-1439.

Salmon, V., Breen, A., Rogers, A., Ely, K., Kumar, J., Sulman, B., & Iversen, C. (2025). *NGEE Arctic Phase 4 plant functional type framework for pan-arctic vegetation* Next-Generation Ecosystem Experiments (NGEE) Arctic. https://doi.org/10.15485/2529470

Shinozaki, K., Yoda, K., Hozumi, K., & Kira, T. (1964). A quantitative analysis of plant form-the pipe model theory: I. Basic analyses. *Japanese Journal of ecology*, *14*(3), 97-105.

Shuman, I. N. C., Serbin, S. P., Erb, A. M., Schaaf, C. B., & Yang, D. (2025). Fine-scale vegetation composition and structure shape spatiotemporal variation in surface albedo across a low Arctic tundra landscape. *Environmental Research: Ecology*, *4*(4), 045001. https://doi.org/10.1088/2752-664X/ae04ef

Vuichard, N., Messina, P., Luyssaert, S., Guenet, B., Zaehle, S., Ghattas, J., Bastrikov, V., & Peylin, P. (2019). Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): Multi-scale evaluation of gross primary production. *Geoscientific Model Development*, *12*(11), 4751-4779.

Yin, X., & Struik, P. C. (2009). Theoretical reconsiderations when estimating the mesophyll conductance to CO2 diffusion in leaves of C3 plants by analysis of combined gas exchange and chlorophyll fluorescence measurements. *Plant, Cell & Environment*, *32*(11), 1513-1524.

Yue, C., Ciais, P., Cadule, P., Thonicke, K., Archibald, S., Poulter, B., Hao, W. M., Hantson, S., Mouillot, F., Friedlingstein, P., Maignan, F., & Viovy, N. (2014). Modelling the role of fires in the terrestrial carbon balance by incorporating SPITFIRE into the global vegetation model ORCHIDEE – Part 1: simulating historical global burned area and fire regimes. *Geosci. Model Dev.*, *7*(6), 2747-2767. https://doi.org/10.5194/gmd-7-2747-2014

Yue, C., Ciais, P., Luyssaert, S., Li, W., McGrath, M. J., Chang, J., & Peng, S. (2018). Representing anthropogenic gross land use change, wood harvest, and forest age dynamics in a global vegetation model ORCHIDEE-MICT v8.4.2. *Geosci. Model Dev.*, *11*(1), 409-428. https://doi.org/10.5194/gmd-11-409-2018