Response to Referee #2

We thank R2 for this detailed review. Enclosed please find a detailed explanation of the revisions we made based on R2's comments. For your convenience, comments are in bold and our response is in italic. Revisions we made in the manuscript are presented in italic with grey background.

Druel et al. include a number of new processes and parametrizations into the land surface model ORCHIDEE that are thought to be important in high latitude ecosystems including * parameter optimization of C3 grass, * implementation of a new shrub PFT, and * implementation of a new PFT representing lichens and bryophytes.

Several additional relationships and processes have been also included, such as * shrub-snow interactions, * vertical soil organic matter profile, * moisture dependence of heterotrophic respiration and anoxic conditions, * moss effects on thermal diffusivity.

In general, I fully agree with the importance to advance the LSMs in these respects and I would like to see such important model development published soon. The auhors also use a number of site-level observations and a formal parameter calibration procedure for this model development. However, I have some serious concerns about this manuscript which should be addressed prior to publication.

Most importantly, there are too many different topics treated in this single manuscript which then are themselves mostly only superficially addressed and which even may not have any relation to each other (in the model). I strongly suggest to focus the paper on 1-2 research questions and a reduced amount of new processes added. I would agree with a presentation of new shrub, moss and C3 grass parametrizations. After a thorough model evaluation, some model application could be presented e.g. to understand the relation of their carbon balances to each other and to trees as well as their effects on soil temperature. Still, I believe individual papers for shrub and moss functions and effects would be more clear. If all topics should stay within one paper, then substantial additional text and figures/tables are required in order to i) explain the research question and importance of processes using literature, ii) evaluate new (and

sometimes old if affected) model functions, iii) present and discuss results with recent literature, and maybe apply the model to address a research question.

The organization of the manuscript was an important step ahead of actually writing this article. And as you suggest, we had to decide between isolating in different articles the different boreal vegetation types (PFTs), or writing an article about the global improvement of boreal vegetation (including all PFTs). We chose this second option in light of its submission to GMD, to focus the article on the model implementation at a global scale and not on a model application with in-depth investigation of a scientific question. It must enable users or developers of other LSMs to understand our developments, compare or integrate processes in order to improve global vegetation modelling.

In this context, it seems to us that splitting into different articles may reduce the interest of the study, especially in view of the fact that an overall and comprehensive evaluation (with global data) of the implementation would be difficult to split. Similarly, limiting the number of processes described could preclude the global consideration of the new boreal PFTs, and importantly, prevent reproducibility of our developments – being a venue for comprehensive descriptions of new model developments is an important goal of this journal. Finally, to reduce the size and complexity of this article, we chose to keep the application of implementations you suggest, such as the vegetation dynamics, the impact on soil carbon stocks or climate changes, for later articles.

However, as you suggested, we have added substantial additional content, especially in the introduction to highlight the research question and appropriate references to the existing literature (p.1 l.31-33, p.2 l.7-11,12,15,20-21,23,26-35,38-40 and p.3 l.6-7,33-37), in the results to provide evaluation on other sites (Fig. 9, p.16-17 l. 38 and l.1-4, and p.18-19 l.32-39 and l.1-6), and in the discussion (from p.22 l.17) to compare our results with more recent studies (in particular Porada et al., 2016). Moreover, in order to clarify and reduce the size of the article, we decided to move the results split by continent (ex figs. 8 to 10 and associated texts) into the supplementary material (Figs. S1 to S3), and to substitute them by Artic-wide averages (Fig. 11).

Overall the paper should be considered primarily as a model description with a main focus on non vascular plants and shrubs, while the improvement for C3 grasses reduces to parameter optimization. The evaluation of the new developments at local to continental scales should thus only be considered as a first step to evaluate the potential of a more realistic description of boreal vegetation in a global model and not as an exhaustive evaluation/validation *of the carbon, water and energy balance of these ecosystems. Such exhaustive evaluation is not compatible with an in-depth model description in a single paper and is thus left for a subsequent study. However, we have tried to better justify in the paper our choices for the selected evaluation diagnostics (and not all available observations).*

Some detailed important issues:

0) It is unclear to me how the authors can neglect the recent publication by Porada et al. (2016) which presents a process-oriented and dynamic representation of bryophytes and lichens in the land surface model JSBACH in introduction and discussion.

Indeed, Porada el al. (2016) is one of the first descriptions of non vascular plants in an ESM, with a process-based implementation. We missed the paper as it only came out after we had already completed our first draft. We thus added this reference in the introduction when describing the current state of boreal PFT in ESM model p.2 l.31-32: "a first description of lichen and bryophytes was implemented in the JSBACH model (Porada et al., 2013), improve recently with a process-based implementation (Porada et al., 2016)". We also compare our results with those of this latter article (in the discussion and conclusion sections) in order to put into perspective our findings.

1) Mosses have an important function in Boreal forests and the forest ground is usually covered by mosses and lichens. Usually we can expect a NVP cover of more than 50% in Boreal forests and more in tundra (Rapalee et al., 2001; Porada et al., 2016). The approach in this study is to treat NVPs as separate PFT with a separate tile results in minor coverage in most regions. (The color scale in fig 5 is not useful to evaluate the shrub and moss cover, please improve). Hence, there will be a strong bias in moss and lichen effects on the heat balance and biogeochemical ecosystem functions using such model. That limitation should be discussed in detail.

In the version of ORCHIDEE used in this article there is no possibility to take into account and model explicitly the understory vegetation cover (the sum of of all PFTs fraction ≤1). We agree that this poses a severe limitation to fully assess the impact of shrubs and NVPs on ecosystem functioning, and more particularly in boreal landscapes. However, we chose to make a first step with the current model structure, treating NVPs and shrubs as separate PFTs like for the 13 standard PFTS. We should notice that in boreal landscapes the forest cover is relatively sparse with significant gaps, by comparison to temperate or tropical forest cover, thus allowing light to reach the ground *more easily. As a first approximation we can thus estimate that NVPs are only partially controlled by the surrounding trees and that the biotic interactions with the other strata are limited.*

Additionally, treating explicitly the understory vegetation, with a process-based approach, is more complicated as it requires a treatment of the radiation transfer within the canopy that accounts for forest gaps distribution and for the intra-canopy climate. Indeed air humidity and temperature are significantly different above the forest canopy than near the ground. Naudts et al. (2016) made a first crucial step in that direction with the addition in ORCHIDEE-CAN of a 2 streams radiative transfer scheme including a "gap" model and Ryder et al. (2016) further added a multi-layer canopy scheme (for energy, water and carbon fluxes) accounting for intra-canopy climate gradients. Our paper should thus be considered as a first step, describing the main biogeochemical features of NVPs and shrubs (as standalone PFTs), before a more complete and comprehensive integration is made within a vertically discretized canopy model (i.e. the ORCHIDEE-CAN version). We thus decided that the available model structure (at the time of the study) was not sufficient to treat explicitly understory NVPs/shrubs.

In this context we agree that the original land cover maps derived from satellite observations largely underestimate the fractional cover of NVPs and shrubs. However, we made an attempt using existing boreal land cover maps to partly correct for this bias. Note also that Peckham et al. (2009) showed that mosses represent a large cover fraction of burned areas, with thus potentially significant year-to-year variations of NVP cover at regional scale. Overall, it was difficult to increase more substantially the NVP/shrub fractional cover without having unrealistically low tree cover. Our study thus represents a lower estimate of the potential impact of NVPs and shrubs on boreal ecosystem functioning.

Note finally that the colour scale of Fig. 5 has been improved.

Peckham, S. D., Ahl, D. E. & Gower, S. T. Bryophyte cover estimation in a boreal black spruce forest using airborne lidar and multispectral sensors. Remote Sens. Environ. 113, 1127–1132 (2009).

Ryder, J., Polcher, J., Peylin, P., Ottlé, C., Chen, Y., Gorsel, E. V., ... & Valade, A. (2016). A multi-layer land surface energy budget model for implicit coupling with global atmospheric simulations. Geoscientific Model Development, 9(1), 223-245.

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

2) I agree with the authors that the global model can hardly cover small-scale variations in NPP and biomass of shrubs and mosses and lichens. Therefore, I suggest modify Fig 6 such that we see one dot for each climatic zone representing the model and data means but including error bars representing their std. Then one can discuss where the model fails to reproduce natural variance within one climatic zone and natural variance among zones. Fig 7 shows importantly that there is hardly any latitudinal variation in the measurements while the model shows a strong variation. Please, discuss in detail.

We agree with the reviewer that Figure 6 would benefit from grouping the individual measurements within restricted climatic/geographic zones. We have thus followed this advice and grouped them according to the six subzones.

Indeed there is a strong latitudinal variation in the model simulations. However, it seems to us that the latitudinal variation in the measurements is as strong, considering the important variation in the mean as well as in the standard deviation. We therefore regret that we do not understand this comment.

3) It seems, model calibration and evaluation at site level has been performed with the same data. If you have too little data to split the dataset into representative parts for calibration and evaluation, then please repeat the site-level model evaluation with a bootstrap method: iteratively remove data for calibration and evaluate respective model results at these sites.

We agree with the reviewer that optimally we should always split the dataset into a calibration and evaluation parts. However in our case several constraints arose from i) the relatively small size of the initial dataset for such split and ii) the large computing time necessary for the model calibration which complicates any bootstrap approach (i.e. the calibration took several weeks with the Genetic Algorithm that is used). Given these constraints we searched for additional datasets to fulfil several requests from the different reviewers. We thus now apply the following strategy:
1) we keep the original Western Siberia dataset to perform the optimization.
2) we use the observations from two new transects in North America and Eurasia (with appropriate biomass data) to perform the model evaluation.

We added a new figure (Fig. 9) for the model evaluation with associated comments reported below. Note finally that we discuss in the text the *potential shortcomings due to the use of mainly lowland data for the calibration of a global model.*

P.17 l.14-19: "We further compare the simulated biomasses with two other Arctic transects. The first one is the North America Arctic Transect (NAAT). It is situated in a continental area, and includes eight field locations (70°N 149°W to 79°N 100°W) sampled from 2002 to 2006 (Walker et al., 2011b) chosen as representative of zonal conditions. The second, located in a marine-influenced area, is the Eurasian Arctic Transect (EAT). It includes six field locations (58 to 73°N, between 67 to 81°E) sampled from 2007 to 2010 (Walker et al., 2008, 2009a, 2009b, 2011a)."

P.19 l.1-15: "Carbon stock with two Arctic transect

To evaluate the modelled biomass in other Arctic sites (not used in the calibration step), including uplands and lowlands, Fig. 9 shows scatter plots of observed and simulated biomass along two transects: the NAAT (North America) and the EAT (Eurasia) Artic Transect. The NVPs and shrub biomasses are relatively well reproduced by the model (i.e. within the error bars). For both PFTs, the standard deviation of the observations includes the 1:1 line, but the observed biomasses are on average higher than the simulated biomasses. Simulated shrub biomasses are biased low for the NAAT transect but not for the EAT transect.

In contrast, the mean value of observed biomass for boreal C3 grasses (Fig. 9.c) is low compared to the simulated biomasses for both cases. For half of the sites the simulated low biomass is in accordance with the observations, but for the other half the values are much larger (> 300 gC.m^2 whereas the observation do not exceed 54 gC.m^2). Despite the optimization with observations from western Siberia (Fig. 7; leading to a decrease of biomass compared to temperate C3 grasses) there is likely an overestimation of the biomass for boreal C3 grasses, probably associated with an overestimated productivity."

Walker et al, 2011a: Vegetation of zonal patterned-ground ecosystemsalong the North America Arctic bioclimate gradient. Applied Vegetation Science 14, 440–463. Doi: 10.1111/j.1654-109X.2011.01149.x

Walker et al, 2011. 2010 Expedition to Krenkel Station, Hayes Island, Franz Josef Land, Russia, Data Report, Alaska Geobotany Center, Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK. 63 pp.

4) I do not agree that LAI is a valid dataset from remote sensing data which is useful for process model evaluation (and if you like to use it please show in the fig ORCH13-GLASS and ORCH16-ORCH13 in order to understand the previous model bias and improvement).

Possible maps for a landscape-scale model evaluation: fAPAR (JRC), GPP (Jung et al., 2011 or Beer et al., 2010), evapotranspiration (Jung et al., 2010), biomass (Thurner et al., 2014), and inventory-based NPP and biomass data (IIASA; Beer et al., 2006; Quegan et al., 2011). This is important as the fraction of tiles of all PFTs has been modified. In general, it would also really good to evaluate catchment runoff with freely available data of large Arctic rivers.

As mentioned above, the primary objective of the paper is not to provide a complete and comprehensive evaluation of the model with all potential large-scale datasets, but to provide a complete description of the new PFTs (equations and parameters) including only a first step evaluation.

Additionally, the validation of the results by world-scale data is not straightforward and potentially critical. The main problem in proposed global products in that they do not include PFTs (or vegetation) distinctions. Moreover, the biomass, NPP and evapotranspiration are more driven by trees or fire distribution than by the influence of the new PFTs. Comparing these maps with the new vegetation cover could add potentially other sources of bias and thus only little additional information. Moreover, the majority of these data is also derived from satellite observations, with comparable biases to those associated to LAI. The fAPAR product, although less sensitive to saturation issues, also comes with its own issues when comparing to current model outputs. The evapotranspiration product from Jung et al. (2010) may suffer from the small set of eddy-covariance measurements available in the boreal zones.

For the catchment runoff, we have done a summary of the river discharge on the ten main Arctic watersheds (http://www.r-arcticnet.sr.unh.edu/v4.0/main.html) to compare with the runoff + drainage simulated on the same area and the same period, p.21 l.19-23: "Compared to observations (main Artic watershed available at http://www.r-arcticnet.sr.unh.edu/v4.0/main.html), the river discharge simulated indicates a general underestimation in the northern high latitudes, linked to an overestimation of evaporation and sublimation (Gouttevin et al., 2012). Thus, this underestimation with ORC16 is smaller than with ORC13."

Although not ideal, we thus kept the LAI as a first step evaluation. Following the suggestion of the reviewer, we added a map (and a transect) of ORC16-ORC13 in Fig. 11 (the map ODRC13-GLASS was already showed. That shows a significant difference between ORC16 and ORC13, and so the improvement with ORC16: p19. L.21-22: "This improvement with ORC16 is directly due to significant lower LAI values in these regions (north of 55°N) compared to ORC13".

5) The reduction in tree cover results in a reduction of transpiration in your grid cell averages. However, interception loss and evaporation should increase with a layer of mosses and lichens. If the water and energy balance is a topic in your paper, then please show results for all components, not only transpiration in Fig 12.

As explained above, the main focus of this article is to describe the implementation of boreal vegetation and only few key impacts, without a thorough analysis of the water, carbon and energy balances. However, we included additional diagnostics in the supplementary material, Fig. S5., with the main components of the water budget: evaporation (including interception), transpiration, runoff and drainage.

6) In this model version, two modifications affect soil temperature: snow depth and moss&lichen cover. First of all, the model version should be evaluated in terms of snow depth and soil temperature. For soil temperature, you can use GTN-P borehole data from Romanovsky et al. (2010) and Christiansen et al. (2010) available at PANGAEA, and maps of soil temperature and ALT even from your study region from Beer et al. (2013) at PANGAEA. I expect a cooling effect from mosses (Porada et al. 2016) due to higher insulation in summer, and a warming effect due to higher snow depth in areas of high shrub cover (still unclear to me at landscape scale as shrubs accumulate snow from lateral wind transport, so it is just relocated within the grid cell?). In Fig 13 both effects are combined. Is there a way to separate them? In Fig 13 it seems the model overestimates ALT and that is even higher in ORC16? In Fig 13b it seems all three grid cells show higher ALT (red) while in 13c one profile shows warmer temp (red) and the others show cooler temp? I generally suggest concentrating on soil temperature because ALT estimation from modelled temperature is not reliable.

We clearly understand the interest and your questions about soil temperature and water balance, key in the Artic to understand physical processes, e.g. the temporal dynamics of ALT and the evolution of permafrost. The Fig. 13 was made to illustrate small perspectives as a sample of the panel of potential impacts, but not as a comprehensive analysis. Given the current length of the paper, it was not possible to investigate these crucial questions in depth.

Additionally, to be exhaustive and perform proper evaluations of this insulating aspect, a factorial analysis would be needed, which was beyond the scope of this article. A dedicated study, with a different version of the ORCHIDEE model, ORCHIDEE-MICT, (including a description of the permafrost properties) has been conducted (Guimberteau et al. GMP, *submitted). In this context, we have chosen to illustrate only that the combined effect (summer and winter) is often more complex than expected with simplified formulations (although they remain important for understanding complex responses at global scales).*

To represent the specific snow accumulation due to lateral wind transport and due to the lower snow compaction (itself due to branch support), the changes introduced (Section 2.3.2) are, as you suggest, just relocated within a grid cell. This is only applied in the case of the snow height used for the snow protection of shrubs (Equation 15).

7) Parameter estimation: Please show a priori and a posteriori parameter distributions in the appendix.

We added the corresponding supplementary: Table S2.

8) Please include a discussion section in which you interpret the results using literature in order to learn something. Parts of your summary section can be used if enhanced by literature. The conclusions and outlook section should be much reduced.

We acknowledge that the long "summary and conclusion" section (section 4) was maybe not the best choice to highlight the results of the study and replace them in the context of recent findings with similar models. We have chosen to follow the reviewer's advice and to split section 4 into a "discussion" section and a "conclusion" section (from p.22 l.17). The discussion now provides few interpretation of the results; however given the above-mentioned main objective of the paper (a model description), we do not provide a comprehensive interpretation of all carbon, water and energy related results. The conclusion has thus been reduced to the main key points of the paper, with an outlook of the next steps.

9) Several new methods are described but their importance, evaluation, and application is unclear: * Section 2.2.6: anoxic conditions are not simulated, soil organic matter dynamics are no topic of the paper. Please remove. Or was the intension to evaluate GPP and NEE at eddy covariance sites? * Why is shrub allometry important and why not only assume smaller trees? * Shrub-snow interactions are not evaluated or analyzed. What do we learn from these additional functions? * Effects on albedo: Has been albedo improved when comparing to satellite products?

We agree that there is probably a lack of evaluation of the new implementations described in the paper. The main reason comes from the need to keep the paper at a reasonable size and that a full evaluation including also a wider range of scientific applications has been left for a subsequent study. On the contrary we tried to represent the ecological complexity of vegetation, because biogeochemical and biophysical processes are interwoven.

Although we did not intend to evaluate the NEE at eddy covariance sites in this paper, we chose to include the modification linked to soil organic matter dynamics in order to provide a comprehensive model (for gross and net carbon fluxes), including the major processes that needed to be improved for subsequent biogeochemical applications.

Specifically for lowlands/peatlands, the maximum decomposition rate simulated with a maximum water content (i.e. in anoxic conditions) is not physically coherent and thus needed revision.

For shrubs, change in allometry (compared to trees) is the key process implemented for their representation: i) the initial tree allometry equation did not allow trees smaller than 10 meters, ii) this allometry impacts directly the mean and maximum values of biomass, which can be accumulated, iii) the height of the vegetation (and particularly the shrubs) is very important to take into account the snow temperature and protection (to maintain biomass in winter). The shrub-snow interaction is not precisely evaluated or analyzed as we believe the first priority is to evaluate whether the shrub biomass (i.e. including height, number of individuals,..) is realistically simulated.

The same concerns apply for the albedo, knowing that only the processes controlling the albedo of the snow were updated, and that the albedo of each new PFT has been kept to that of the original PFT (as a first approximation). Additional work is needed to fully characterize the albedo of the new PFTs and for NVPs its dependence to moisture conditions. This work is beyond the scope of the paper and we thus decided not to focus on a global evaluation of the albedo with existing satellite products.

In conclusion, it would have been orthogonal to the main objective of the paper to neglect key processes controlling the biogeochemical and biophysical functioning of the new boreal PFTs. But the evaluation and application of all of these aspects is impossible in one (already too long) article.

Minor issues:

Fig 10: not used in results but only in summary and that there also the fig does not support the sentence.

The Fig. 10 was use and directly mentioned in the result (Section 3.2., in the first submitted version from p.19 l.20 to p.19 l.34). However, to be more clear and concise we have decided to move this figure, as well as the figures 8 and 9, to the supplementary (Fig. S4).

CO2 conductance in non-vascular plants depends strongly on its moisture and not on stomatal conductance. If that concept is not used here, then please discuss this limitation and related potential biases in detail.

We agree and it is for this reason that we have modified the constant value of the variable named "stomatal conductance" (Section 2.2.1., Eq. 1 and 2.) to reduce its dependence to active stomata and increase its dependence to moisture.

Page 16, line 35: I do not understand.

This was a description of the list of optimized parameters. As you suggested, it is now more explicit with the appendix (Table S2). In addition, these lines have now been moved at the end of Section 2.6.2.