We would like to thank all reviewers for their insightful comments, which helped us to improve the manuscript. Our revisions reflect all reviewers' suggestions and comments. For detail, please refer to the responses as follows: reviewer comments in black fonts, responses are in blue fonts.

Reviewer #1

Abstract - This should summarise all the results, so add a sentence about the impact of the root developments as well as the dynamic SOL. Also it would be good to add that initialising with observed SOC gives a better SOC distribution in the simulation than the previous method, just to make this clear.

We implemented suggested rewrites.

Introduction -

Page 3139 Line 5: Please define what you mean by active layer thickness.

We added the following sentences (line 36):

The active layer in permafrost regions is the surficial layer overlying the permafrost, which undergoes seasonal freeze-thaw cycles. Active layer Thickness is the maximum depth of thaw at the end of summer.

Lines 7-8: In fact, Burke et al (2013) does not use a constant carbon density but uses the observations from NCSCD.

We thank the reviewer for pointing it out. We removed Burke et al., (2013) from the list.

Page 3140, Line 5: "Here we describe a fully dynamic SOL to demonstrate the importance of coupling soil biogeochemistry and thermodynamics": In the analysis later you demonstrate the importance of having a soil organic layer of realistic thickness but you do not demonstrate the importance of making it dynamic. I would not suggest that you change this statement but rather that you should do some more analysis/discussion (see later on).

We expanded the discussion section and stressed the importance of dynamic SOL.

Methods -

Is the SiBCASA model used in any coupled GCM/earth system model? This would be good to mention if so.

SiBCASA is not currently coupled to a GCM. SiBCASA's predecessor, SiB, has been coupled to the Colorado State University GCM, although it has since been replaced with a different land surface parameterization. We reference the coupling [Sellers et al., 1996], but decided not to delve into this background in the methods section.

Section 2.1 (Frozen carbon initialization)

It is not clear how you intialise the permafrost carbon - is it at the beginning of the 900 year equilibrium run? If so how do you determine the maximum active layer thickness? Or is it at the end of the 900 year equilibrium run, so your equilibrium run is performed with no permafrost carbon present? Or do you do two equilibrium runs: one without permafrost carbon and a second one when you have initialised it based on the active layer in the previous run? Please describe this procedure more clearly.

We added the following clarification to the Methods section.

'It is important to know the "stable" depth of active layer before initializing frozen carbon. We run the model for several years in order to calculate ALT, and then initialized frozen carbon below the maximum calculated ALT. The frozen carbon was initialized only once during the first equilibrium run cycle. For the next equilibrium run we used the previously calculated permafrost carbon. We defined an equilibrium point when changes in overall permafrost carbon were negligible or almost zero.'

Page 3142, line 18: These equations appear unexpectedly with nothing leading up to them to say what they are. I suggest to add a sentence here along the lines of 'The carbon in each layer is divided into three pools as follows:' and then give a definition of the pools.

We added a sentence as suggested. (L.128)

Section 2.2 (Dynamic SOL)

"SiBCASA already accounted for the effects of organic matter on soil properties like porosity, ...". Please give more details of this! It is important for understanding the work. For example, did the properties vary depending how 'compressed' the organic matter was? Do they vary with depth assuming the organic matter is more compressed at depth? What properties are used and how are they combined with mineral soil properties?

We added the following text (L.149): SiBCASA calculates the soil physical properties as a weighted average of those of organic matter, mineral soil, ice and water (Schaefer et al., 2009). The physical properties include soil porosity, hydraulic conductivity, heat capacity, thermal conductivity, and matric potential. The model calculates the organic fraction used in the weighted mean as the ratio of simulated carbon density to the density of pure organic matter. SiBCASA does not account for the compression of organic matter. Since the prognostic soil carbon pools vary with depth and time, the organic fraction and the physical properties all vary with time and depth. We only summarized these calculations here since the calculations are covered in detail in Schaefer et al. (2009).

It is unclear how the carbon was dealt with in the previous version. The implication is that all carbon entered (and stayed in) the top soil layer and there was none in any deeper layers? Is that true? Please clarify in the text. How did that allow the model to include permafrost carbon (with constant density), as you have mentioned and compared with (Schaefer et al. 2011)?

Each model layer has a complete set of prognostic soil carbon pools. The previous version of the model distributed fine and coarse root growth vertically within the soil column based on observed root distributions. As the roots die, carbon is transferred to the soil carbon pools for that layer. Thus, the maximum rooting depth determined the maximum depth of 'new' or 'active' carbon in the model. Of course, if the maximum rooting depth felt below the permafrost table, the model would accumulate permafrost carbon. The current version of the model initializes the permafrost carbon by assigning carbon to the soil carbon pools below the maximum thaw depth. These frozen pools remained inactive until the layer thaws. As live, above-ground biomass in the model dies, carbon is transferred into the first layer as litter. Without the vertical redistribution we describe here to create a surface organic layer, the top layer of the model tended to accumulate carbon in excess of that expected for pure organic matter.

In equation (4) you multiply C_{max} by f_c. This seems strange given that there was already a factor of f_c in the definition of C_{max} in equation (3). Is there a mistake here? When you say 'C_{max} is 140 kgm^{-3}', perhaps you mean 'ρ_{max}' here? Please check this.

There is a mistake: fc should not appear in equation 4. We corrected the equations accordingly.

What is the purpose of defining OLT_{max}? (Equation 5)

We used OLTmax early in the development phase as a diagnostic tool to analyze our results. However, we now see that it does not provide much additional insight and decided to remove it from the paper.

Section 2.3 (Root growth and soil thermal factor)

I am slightly confused as to how your root growth works. Is the root profile prescribed as exponential? If so, what difference does it make if the roots are only growing in the thawed layers? Does this affect the input of carbon to the different soil layers (as well as the autotropic respiration)? That part was not clear in the text.

We added the following text (L.209): The vertical distribution of new root growth between the soil layers is prescribed using exponential curve fits to observed vertical root distributions. Before the changes we describe here, the maximum rooting depth sometimes exceeded the thaw depth in permafrost soils, resulting in root growth directly in permafrost, which is unrealistic since growing roots cannot penetrate frozen soil. Since the permafrost never thawed, these simulated roots soon died, but never decayed, resulting in an unrealistic buildup of carbon in the upper layers of permafrost. This in turn set up a feedback where the unrealistic increase in organic matter in the simulated permafrost changed the thermodynamic properties and decreased the ALT, resulting in additional carbon buildup. To solve this problem, we kept the original exponential vertical rooting profile, but set maximum rooting depth equal to the thaw depth. This allowed the maximum rooting depth to vary with time and effectively restricted all root growth to within the active layer.

Results -

Page 3147, line 13-14 “so the effect is not as pronounced" - what effect are you talking about here? Not clear.

We clarified it with the following sentence (L.293) “… which allows us to reach a higher heterogeneity between measured and simulated ALTs”

Line 20-21: Slightly confusing to say “a strong peak" when refering to the first plot, since the peak is stronger in the second plot. Probably better to just say “a peak".

We implemented suggested change.

Line 26: This paragraph is not very clear. I’m not sure what you mean by ‘coupling’. It would be better explained along the lines of... ‘In the version without SOL the ALT was generally deep in forest biomes, but in the new version there is a thick SOL (due to high GPP), which leads to a much shallower ALT. This means there is now a significant amount of root growth in the permafrost itself, which puts carbon directly into the permafrost stores. This is unrealistic, and the frozen soil restrictions on GPP and root growth together eliminate this problem.’ (maybe you could write it better but that is the general idea.)

We changed the confusing paragraph to the following (L.305): In the previous version without a dynamic SOL the ALT was generally deep in forest biomes, but in the new version there is a thick SOL (due to high GPP), which leads to a much shallower ALT. Without restricting root growth within only thawed part of the soil the shallower ALT feedback leads to a significant amount of root growth in the permafrost itself, which puts carbon directly into the permafrost stores. This is unrealistic since growing roots cannot penetrate frozen soil, and the frozen soil restrictions on GPP and root growth together eliminate this problem.

It is misleading to suggest that ‘coupling between GPP and ALT’ does not reflect reality, when in fact there is a real coupling between these quantities. Here it is a negative feedback: Increased GPP → increased litter → increased SOL → reduced thawing → reduced GPP. There is also, however, a positive feedback on soil organic carbon as more soil carbon → increased SOL → reduced temperature → reduced respiration → more soil carbon. These feedbacks are potentially important and it is great that your model now includes them, but there is no analysis of these feedback processes, which is a shame. Please cite Koven et al (2009) (which appears in your reference list but not in the manuscript), and include some analysis to demonstrate the coupling in your model. This would significantly improve the worth of the research.

We removed confusing paragraph and added following paragraph to the conclusion:

‘The dynamic SOL and rooting depth strengthens the feedback between GPP and ALT (Koven et al., 2009). Higher GPP produces greater litter fall, which increases the input soil carbon at the surface and results in a thicker SOL. The dynamic SOL changes the properties of the near surface soil, resulting in a shallower ALT and cooler soil temperatures. The dynamic rooting depth accounts for a shallower ALT and modulates GPP accordingly. The cooler soil temperatures slow microbial decay and increase the carbon accumulation rate, which in turn increases the SOL and reduces ALT further. Eventually, this feedback results in the development of a peat bog. The changes we describe here indicate that SiBCASA can simulate the dynamics of peat bog development, but the model does not yet include a dynamic vegetation model to account for conversions between biome types, such as boreal forest to peat bog.'

Page 3148, line 14 onwards. You are discussing the total amount of soil carbon in the simulations. Here it would be good to compare this with the observed total from NCSCD (as well as spatial distribution in Fig. 6).

Prescribing permafrost carbon according to the NCSDC dataset allowed us to better match with the observed pattern in the soil carbon. However, it is does not mean that after the spinup simulated permafrost carbon stocks exactly matched the NCSDC data. During spinup thaw depth varies with time, introducing carbon movement from frozen to thawed pools. In discontinuous zones, if the model simulated permafrost, it tended to produce a deeper active layer depth and thus less permafrost carbon than the NCSCD. The major difference between uniform frozen carbon initialization (Fig 6A, old; Fig 7A,new) and initialization according to the NCSCD (Fig 7B new) is that permafrost exists in more places. However, the NCSCD map (Fig 7C) shows that not all frozen soil contains prescribed by default uniform amount of frozen carbon. Therefore simulating the 'correct' ALT is important and should improve the overall permafrost carbon storage.

Page 3149, line 1 (and end of prev. page): "overestimation of the SOC in Central Siberia occurs due to high SOC at the initial time step" - what do you mean by initial timestep? How could that happen? I think if you explained the spinup procedure more clearly in the methods this would be easier to understand.

We rephrase the sentence for better clarity: The overestimation of SOC in Central Siberia is a result of coupling between GPP and ALT.

Page 3149, line 3: How can the ALT be identical when the soil carbon amounts are different? I thought that the physical soil properties depended on the amount of car- bon? Again perhaps if you had explained in more detail the changes to soil properties in section 2.2 it would be clearer how this could be the case. (See my comments for 'Methods' section, above.)

We included suggested improvements to the method section and added the following text (L.360) for better clarity: This is due to the fact that in both cases soil carbon is added in the permafrost layer below the active layer. Consequently, the amount of soil carbon in the active layer stays does not change between simulations and has the same thermal and carbon dynamics, and thus ALT.

Discussion -

In the second paragraph of the discussion you are talking about Figure 7, which shows the spatial patterns of various physical forcings (air temperature, downward longwave radiation, snow depth, and soil wetness factor). Many claims are made about which factors are contributing to the permafrost distribution and in what way. This analysis is not rigorous and must be improved before the work can be published. For example, "maximum snow depth in South-East Canada is almost half that of West Siberia, which suggests that snow in SE Canada, most likely, is not a major contributor to warm ground temperatures" - well what about the fact that there is a 'critical' snow depth below which the soil is much more sensitive to changes in snow depth (see eg. Ekici et al. 2014a): shallower snow does not necessarily imply that it has much less effect. An actual relationship between two variables can only be asserted based on analysis such as a regression or spatial correlations. If you would like to say which factors have the most influence, either comparing the spatial correlations of the forcing variables with ALT, or perhaps performing a multiple regression of all the variables against ALT, would give you much more definite claims. I suggest you replace this paragraph by a more rigorous analysis of the influencing factors.

We added Figure 9 (shown below) that shows the correlations between ALT and every parameter shown in Figure 7. Also, we expanded the discussion for each of the influencing factors. We added and extra discussion on effect of snow on ALT.

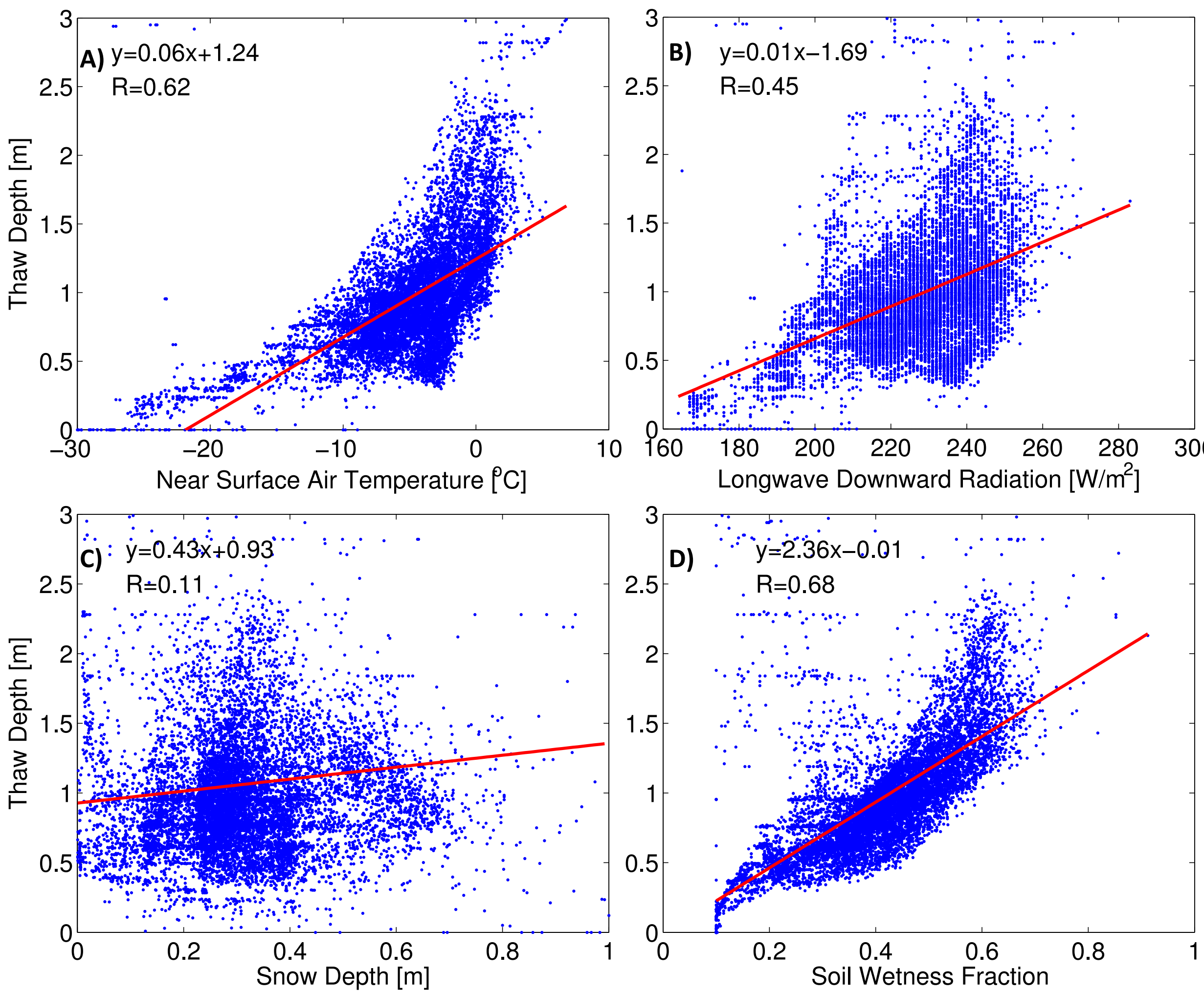

Figure 9. The correlation between ALT and: A) near air temperature for averaged over first two month of the fall season, and B) the down-welling long-wave radiation, averaged yearly over 10

years. C) the maximum snow depth over 10 years for the steady state run, and D) the soil wetness fraction, averaged yearly over 10 years.

Similarly, the final paragraph mentions that the CRUNCEP data may not have the zero degree isotherm in the right place. What makes you think that? Why might it not be a problem with the longwave radiation, for example? Could you perhaps find some in-situ air temperature measurements to support that claim?

We agree that paragraph sounded more like speculations. We substituted the paragraph with the following (L.420): Simulated permafrost vulnerability is tightly coupled with the accurate modeling of the present permafrost distribution, which depends on soil thermal properties. We calculate soil thermal properties based on prognostic soil carbon and soil texture from the Harmonized World Soil database (FAO et al., 2009). Observations indicate that soils in the southeast Canada have high soil carbon as a result of a large number of peat lands (Hugelius et al., 2014). Peat has low thermal conductivity and could preserve permafrost even at NSATs about zero degrees centigrade (Jafarov et al., 2012) even if the surrounding areas do not have permafrost. However, the HWSD input data does not have enough soil carbon in the southeast Canada and southwest Russia, as a result, we could not simulate permafrost in those regions.

Conclusions -

Again you claim that the dynamics of the SOL are crucial. In fact, as far as I can tell it is the presence of the (right kind of thickness) SOL that is important. For example, this could be achieved with a static method based on soil carbon observations (such as in Chadburn et al. (2015)). You should show in the paper why your method is better or at least discuss the findings of Koven et al. (2009), which showed the impact of the dynamic coupling between soil carbon and soil properties.

We agree and rewrote the conclusion section.

Lines 9-10. I'm not sure what this sentence means. Particularly, how did the change to plant root growth improve the ALT? And how did it improve the soil carbon? Was this because the carbon was no longer input to the permafrost? I think this needs to be clarified here or better explained in the analysis.

We rewrote the conclusion section.

Line 14-15: "The initialised soil carbon respired during spinup due to abundance of permafrost within the top 3m." This does not make sense physically. Please check.

We agree and rewrote the conclusion.

Figures -

Figure 7, please define 'water stress factor'. Also, do you really mean 'non-dimensionless'? I guess it should just be 'dimensionless'?

We substituted the water stress factor with more meaningful description soil wetness factor. We corrected the typo.

A final suggestion I have which is currently missing from the paper: Some observed data to support the changes in GPP that have resulted from your changes to the model (eg. those shown in Figure 4). Does simulated GPP improve, and if not, why not - or how bad is it?

The major goal was not compare the simulated GPP with measured GPP, but rather to indicate the effect of soil thermal conditions on GPP. Our goal was to illustrate the reduction in GPP when the soil thermal physics is coupled with plant photosynthesis. Preliminary comparisons indicate the SiBCASA simulated GPP may be too high globally, but the cause of this bias is the remotely sensed phenology used as input and not the soil biogeochemistry and thermodynamics. Including such a comparison here would 'open a can of worms' and introduce ideas that are completely unrelated to SOL dynamics. This would greatly dilute the emphasis of the paper of SOL dynamics, so we decided not to include it here. We are planning a rigorous comparison under the Multi-scale Synthesis and Terrestrial Model Intercomparison project (MsTMIP, nacp.ornl.gov).