Reply to referee #1.

We thank the referee for the positive comments. Below are our responses (in normal font) to the reviewer's comments and questions (in italics).

General Comments:
This paper presents a reduced order model (ROM) to resolve very fine resolution soil moisture structure in tundra landscapes. The reduced order model is essentially a statistical model developed using principal component analysis, or "principal orthogonal decomposition mapping method" as stated by the authors. They present this ROM method as a way to improve representation of sub-grid scale heterogeneity in land surface models used for climate simulations (grid spacing of O(100km)). Overall, they show the ROM has promise and is able to capture fine scale features from large scale simulations with generally O(10) bases.

The development of this type of sub-grid scale parameterization is novel and likely will have future applications in biogeochemistry modeling. However, the article is somewhat hard to read in places due to usage of the past tense in areas where the present tense should be used.

A: The use of tenses in scientific publications is often a matter of style, sometimes set by the journal. We prefer to use past tense for work that happened in the past, and present tense for discussion and explanation of ideas resulting from work that happened in the past and suggestions for future work. To address this reviewer concern, we have reviewed the manuscript carefully to ensure that these guidelines are followed.

I am also concerned with the ability of this method to be generalized to larger areas and more diverse landscapes.

A: We agree that we have not applied the technique described to areas of size and diversity typical of climate simulations, but re-emphasize that that was not the goal of this paper. As discussed in the manuscript (pg. 2142 line 1), we view the development of relatively small spatial extent ROMs discussed here as only a first step toward representation of 3-dimensional processes in climate-scale land models. To further clarify this restriction of the current work, we have removed the "regional- and climate-scale" from the title and discussed in detail how we intend to extend our current results to larger-scale simulations in Section 3.4: Application to larger-scale hydrological simulation.

Specific Comments:
1) The authors attempt to show transferability and general applicability of this type of ROM to many landscapes, but the development of the generalized ROM is for essentially the same landscape type over a very limited area. The four test sites are all very close together and only cover around 40,000 square meters. It would be very worthwhile to see a general ROM work across landscape types. More discussion of how to develop ROMs for larger areas is needed at a minimum. Would a ROM for each

grid cell from a predetermined grid be made, then applied in a large scale simulation? Would there be only N number of ROMs globally?

A: These questions define research goals that we are currently actively addressing, but are beyond the scope of the current manuscript to answer. To address this reviewer concern, we have added 3 additional paragraphs to discuss how we intend to approach the transferability and general applicability of this ROM approach (Section 3.4). In these additions, we indicate that: (1) it is necessary to construct separate ROMs for different landscape types to ensure the ROMs remain computationally efficient but a single ROM will be applicable to multiple coarse grid cells; (2) based on the parameterization and analysis of the parameter space, one determines scenarios that will be used to generate solutions to construct the ROMs; and (3) we expect there to be a limited number of ROMs required, but the exact number will depend on analysis of the parameter space.

2) The driving simulation is at 8 m resolution which is at least two orders of magnitude smaller than even the highest resolution regional climate simulation using CLM. What resolution would a ROM developed from a large-scale forcing grid of O(1 km) be produced at? Can a ROM even be developed to generate the correct structure at the required resolution for the biogeochemistry simulations from that forcing grid?

A: We have added a paragraph to Section 3.4 to address these questions. We propose a hierarchical approach that involves using POD-MM methods to develop ROMs at multiple scales; scales at which these ROMs are built may critically depend on scales of the different processes we are modeling. For example, lateral and vertical subsurface flow are strongly influenced by the heterogeneous soil properties even though the forcing grid may be at a resolution that is much greater than the resolution at which the subsurface flow is simulated. The heterogeneous soil properties lead to heterogeneous saturation that directly affects biogeochemistry. We do not believe the forcing grid (possibly dictated by the atmosphere model) should limit the resolution at which LSMs are simulated, especially as we develop the next generation of land models. In addition, the test case that we examined demonstrated that our method is capable of reproducing the soil moisture structure at the required resolution for the biogeochemistry simulations. After the ROMs at multiple scales are built, the POD-MM reconstruction procedure is then recursively applied to reconstruct solution at progressively finer scale, starting from the coarsest scale solution (e.g., those computed on the forcing grid).

3) *Since a ROM is not physically based, how will changes in climate or hydrologic response in time be captured? The authors note that proper sampling of the full forcing phase space will help increase generality, but what about situations occurring in the future that may not be in the observed record?*

A: We have discussion of these issues in Page 2146, line 16-25, but enhance the manuscript to more thoroughly address this reviewer concern. For a more diverse parameter space, relying solely on historical climate forcings is insufficient. Statistical or adaptive sampling techniques should be used to sample the parameter space to ensure we take into account future conditions that are not represented by historical data. Accurately defining the extent of the parameter space is thus crucial. In addition, just as with any data assimilation technique, the ROM must be updated when new information is available, or the forcing moves outside of the phase space under which the ROM is developed. For example, if we are using the ROM at a parameter point that is outside the convex hull of the parameter space used to construct the ROM, it indicates that the ROM needs to be updated to reflect the change in the extent of the parameter space.

4) *The tense throughout the article is incorrect. The authors use past tense to describe work in this paper, which gets somewhat confusing. For example: Page 2131, line 5: "To that end, we described, tested and applied" This is referring to work in this paper? If it is, the tense should be present: "To that end, we describe, test and apply". Many pages describing the work have this issue.*

A: Really, it should be 'To that end we describe below tests and applications of the model …'. As we mention above, we will carefully review the manuscript for correct use of tenses.

5) *I wonder about the last statements of the article regarding topographic differences between the study sites. What are the units in Figure 1? If they are meters, there is only a 1 m range in the color bar, indicating very little topographic variation. What about a region of steep, complex topography where there may be 20 m elevation change over a 100 m plot, then the next 100 m plot is flat? How would that impact a site independent ROM?*

A: The units in Figure 1 are in meters. The overall topographic relief for the BEO is low, but the four NGEE study sites have distinct microtopographic features that include low-centered, high-centered, and transitional polygons. Contrasting polygon types are indicative of different stages of permafrost degradation and were the primary motivation behind the choice of study sites for the NGEE-Arctic project.

For the scenario that you have described, the site-independent ROM built only with sites with steep, complex topography will perform poorly for a site with flat topography. The topography of the subdomains must be included in the parameter space from which we sample the solutions. Proper parameterization and subsequent sampling of the topography will allow the ROM to represent any topography within the larger domain. As such, similar to answers to Question (1) and (3), proper sampling of the parameter space is essential to the representativeness of the ROM. We have added one paragraph to Section 3.3 to clarify this reviewer concern.

Technical corrections:
Page 2127, line 23: Change to: "wetland biogeochemistry and occurs at scales"

A: The above error has been changed in the revised manuscript.

Page 2132, line4: What does the "O" stand for in BEO? Should it be "Barrow Environmental Observatory?"

A: The above error has been changed in the revised manuscript.

Page 2132, line 14: Change to: "the majority of precipitation falling during the"

A: The above error has been changed in the revised manuscript.

Page 2134, line 27: Change to: "Eq. (3) in Sect. 3"

A: The above error has been changed to "results based on (3) are presented in Sect. 3".

Page 2141, line 14: Should it be: "e_(Δx_g)^(POD-MM)" the second time on this line?

A: Yes. We have restructured the manuscript and the definition of the above now appears under Sect. 2.2.6.

Page 2141, line 16: Should it be: "Δx_g = 0.5 m, e_(Δx_g)^(POD-MM2)" You may want to re-check super and subscript notations elsewhere.

A: Yes. The above error has been changed in the revised manuscript. We have rechecked the super and subscript notations as well.

Figure 1: Units are needed on color bar. Also would be nice to have X and Y distances on the axes to have an understanding of how big the study sites are.

A: We added both the unit for the color bar, and X and Y distances on the axes. We have also explicitly stated the spatial extent of the sites in the caption.

Figure 3: Y-axis is labeled as a PDF, it may be better to normalize the histogram counts so it is truly a PDF.

A: We normalized the y-axis in the revised manuscript.

Figure 4: Units for X-axis

A: We added m^3/m^3 to the x-axis.

Figure 9: Units needed

A: We added m^3/m^3 to the colorbar.

Figure 10: Units needed. Also, color range doesn't capture range of values well for most panels, should change.

A: We added m^3/m^3 to the colorbar. The large regions with constant red color in several of the panels reflect the fact that the solution is saturated in these regions. As such, changing the ranges of the color will not change how these panels look. We have added additional comment in the caption.

Figure 15: Using the same Y-axis range for both panels would be helpful.

A: We used the same Y-axis range in the revised manuscript.

Figure 16: Again, normalize the histogram for the PDF.

A: We normalized the y-axis in the revised manuscript.

Figure 17: Units needed.

A: We added m^3/m^3 to the colorbar.