Authors: Liang Chen et al. General Comments: The paper has scientific relevance in evaluating the performance of the Noah-MP for boreal forest site. In addition, a parameterization was included in the Noah-MP LSM to represent the vertical heterogeneity in the soil structure, through the introduction of an organic soil layer. Such efforts contribute to the improvement of land surface models. However, the manuscript needs major revisions to be accepted.

The authors need to rewrite the results to correlate them with the proposed objective. For example: 1) Describe the results of the figure X which are relevant to the purpose of the manuscript. 2) Discuss these results.

Thank you for your thoughtful suggestions, which are very helpful to guide us to clarify and improve the manuscript. We much appreciate it! Our responses to your questions are in italics.

According to the figures presented in the manuscript, the OGN simulation is better than the CTL simulation, for the sensible heat flux in spring, the soil temperature at depths of 10-40cm and 40-100cm, and the soil moisture at 40-100cm. The latent heat flux and soil temperature at the topsoil layer from the OGN simulation are very close to the CTL simulation. The soil moisture at the topsoil layer from the OGN simulation presents a worse performance than the CTL simulation, compared with the observations. However, for wet years, there are an improvement (closer to the observations) in the latent and sensible heat fluxes and volumetric liquid water from the OGN simulation in relation to the CTL simulation during spring. So, the authors should review the affirmation below presented in the abstract, and mentioning carefully their principal results.

Excellent suggestions. We have carefully reviewed the results and substantially revised the abstract that reads as:

"A thick top layer of organic matter is a dominant feature in boreal forests and can impact land-atmosphere interactions. In this study, the multi-parameterization version of the Noah land-surface model (Noah-MP) was used to investigate the impact of incorporating a forest-floor organic soil layer on the simulated surface energy and water cycle components with data from a BERMS Old Aspen Flux (OAS) field station in central Saskatchewan, Canada. Compared to the simulation without organic parameterization (CTL), the Noah-MP simulation with an organic-soil (OGN) improved Noah-MP simulated soil temperature profiles and soil moisture at 40-100cm, especially the phase and amplitude of soil temperature below 10 cm. OGN also enhanced simulation of sensible and latent heat fluxes in spring, especially in wet years, which is mostly related to the timing of spring soil thaw and warming. Simulated top-layer soil moisture is better in OGN than that in CTL in summer but worse in winter. The effects of including an organic soil layer on soil temperature are not uniform throughout the soil depth and year, and those effects are more prominent in summer and in deep soils. For drought years, the OGN simulation substantially modified the partition between direct soil evaporation and vegetation transpiration. For wet years, the OGN simulated latent heat fluxes are similar to CTL except for spring season where OGN produced less but closer to observation evaporation. Including organic soil produced more sub-surface runoff and resulted in much higher runoff throughout the season in wet years".

The authors need to improve their discussion about results. For example, the authors did not discuss the positive bias found in both simulations for the SH in summer. The authors

presented the result of the soil temperature at the topsoil layer, but the discussion only appeared in the end of the section.

We added a new paragraph at the end of Section 4.4 to address the high bias in summer SH.

In addition, the authors need to improve the readability to clarify some phrases. Some of mistakes could be fixed by a final reading before submission, such as blanks in the text, two identical phrases in the section 3.1, mistakes in the figure captions.

Specific comments:

- Lines 11-12: ". . ., the most widely used numerical weather prediction and regional climate model in the world." Are there any references about this affirmation?

  Reference added, revised the sentence to read "which is widely used as a numerical weather prediction and regional climate model for dynamical downscaling in many regions world-wide (Chotamonsak et al., 2012)".

- Line 13: As a suggestion, the authors could include the reference Pilotto et al. (2015).

  It is added.

- Lines 13-14: ". . .compared to the legacy Noah LSM. . .". I suggest replacing "legacy Noah LSM" by "earlier versions of the Noah LSM".

  Revised the sentence to read "compared to earlier version of Noah LSM"

- Lines 21-22: "Despite continuous evaluation and improvements, Noah-MP has not been evaluated in boreal forest regions." And Yang et al. (2011)?

  Delete the sentence "Despite continuous evaluation and improvements, Noah-MP has not been evaluated in boreal forest regions." The Noah-MP has not been evaluated in specific boreal forest flux sites, Yang et al. (2011) use Noah-MP to test many river basins in the world including boreal forest regions, but the it didn't test the specific boreal flux sites.

- Line 36: I think the word "old" should be removed in this phrase.

  "old" has removed. Revised the sentence to read "especially the version CLASS 2.7"

- Line 61: I would replace "thermal and hydrological components" by "surface components".

  Revised the sentence to read "surface energy and hydrological components,"

- Lines 73-76: I suggest that the soil types at the site should be described with more clarity.

  In my opinion the soil description is already pretty good. Here are some small revisions: The soil is an Orthic Gray Luvisol (Canadian Soil Classification System) with an 8-10 cm deep forest-floor (LFH) organic horizon overlying a loam Ae horizon (0-21 cm), a sandy clay loam Bt horizon (21-69 cm), and a sandy clay loam Ck horizon (69+ cm) . 30% of the fine roots are in the LFH horizon and 60% are in the upper 20 cm of mineral soil.

- Line 100: “Data gaps were filled using a standard procedure.” Reference?

  Add a reference: Data gaps were filled using the Fluxnet-Canada standard procedure. (Amiro et al. 2006).

  Amiro BD, AG Barr, TA Black, H Iwashita, N Kljun, JH McCaughey, K Morgenstern, S Murayama, Z Nesic, AL Orchansky, and N Saigusa. 2006. Carbon, energy and water fluxes at mature and disturbed forest sites, Saskatchewan, Canada. Agric. For. Meteorol., 136: 237-251.

- Lines 125-126: “Noah- MP is a new-generation of LSM, developed to improve major weaknesses of the Noah LSM.” I suggest to change this sentence to something like:“Noah-MP is a new-generation of LSM, which was developed to improve the performance of the Noah LSM.”

  Done.

- Lines 205-206: “They are then treated as the most appropriate combinations for our study site (see Table 3).” This sentence is not clear if the authors used the parameterization options mentioned in the sentence above. I think it should be rewritten.

  These sentences are replaced by “The selected Noah-MP physics options used in this study are similar to Barlage et al. (2015), Gao et al. (2015) and Chen et al. (2014) and are list in Table 1.”.

- Lines 206-208: “The order of the categories based on the IOA scores from the highest to the lowest is. . .”. If the authors kept the comparison between the parameterization options, perhaps this result should be explored and discussed.

  This entire paragraph is deleted and we added the above-mentioned sentence to explain Noah-MP option selection.

- The text does not mention how many soil layers were used in the simulations, and what the depths were used. I believe that the authors have used four layers. Three layers were mentioned in the results: 0-10cm, 10-40cm, and 40-100cm. In the caption of the figure 6, a fourth layer was mentioned as been referring to 100-200cm. Is it correct? Please explain in the methodology.

  It is correct, we use 4 soil layers in Noah-MP: 0~10cm, 10~40cm, 40~100cm, and 100~200cm. Because observations are only available for 0~100cm, in this study we only discuss the results for the top three layers. We revised section 3.1 to reflect this.

- Line 210: I think the authors should create a specific title for the section 4.3, as it was presented in the sections 4.4 and 4.5. In fact, the “evaluation results” also include the sections 4.4 and 4.5.

  The title and sections have been re-organized: deleting the entire section 4.2; the original Section 4.4 became new Section 4.2 focusing on soil temperature and moisture; changing

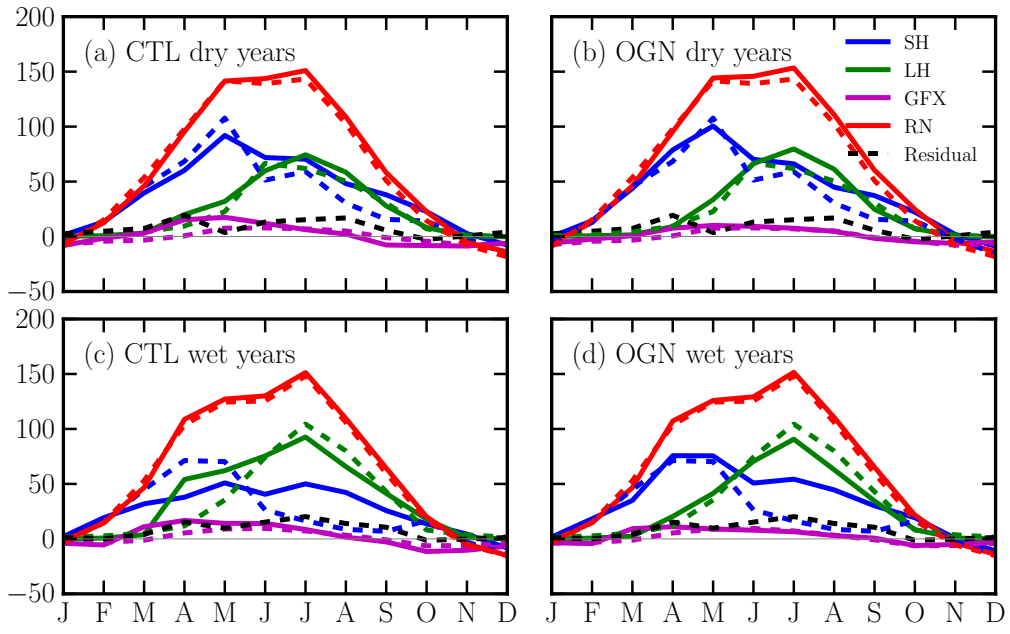
the title of Section 4.3 to read "Seasonal cycles of sensible and latent heat flux"; The titles for Sections 4.4 and 4.5 remain unchanged.

- Perhaps the figures 4 and 5 can become a single figure, as done in the figures 6 and 7. This may help in the analysis of results.

Because Figures 4 and 5 are different types of figures than Figure 6 and 7, we decided to retain Figs. 4 and 5.

- In the second paragraph of the section 4.3, the authors mentioned about the positive sensible heat flux bias simulated by the both simulations in summer. Why does this bias occur? Did you see the field of the net radiation? More interpretation would help.

Thanks for brining this up. We hypothesized that high bias in summer sensible heat flux is partly attributed to energy imbalance in observations. We calculated the energy balance residual term: RN-(SH+LH+G), which is plotted as the black dashed line in the figure below. For summer month (i.e., June, July, and August) in wet years, GFX in CLT and OGN is close to observed values; modeled latent heat flux is underestimated by ~10 W/m2; modeled sensible heat flux is overestimated by ~30 W/m2; and the residual term is ~17 W/m^2. So it is clear that the surface energy imbalance (~17 W/m^2) in observations contribute to a large part of the ~30 W/m^2 high bias in sensible heat fluxes. In dry years, the summer energy imbalance (~15 W/m^2) is nearly equal to the high bias in sensible heat flux (~15 W/m^2). The above explanation is included at the end of Section 3.4 in the revised manuscript.



- Why the RMSE and IOA were not calculated for the soil temperature and moisture?

We calculated the RMSE and IOA for simulated soil temperature and moisture (shown below), but these statistics did not provide additional information than what is already presented in Figs. 4 and 5, so we did not it use in the manuscript.

	CTL			OGN		
	R^2	RMSE	IOA	R^2	RMSE	IOA
SoilT1	0.87	3.71	0.93	0.88	3.85	0.92
SoilT2	0.91	2.96	0.94	0.97	1.16	0.99
SoilT3	0.92	2.80	0.93	0.88	1.77	0.95
SoilW	0.49	0.07	0.69	0.56	0.04	0.84

- Why the simulations with the Noah-MP (independent of the soil type) produce a bias on soil temperature at the topsoil layer in winter? Parameterization?

  Top layer soil temperature in winter is highly dependent on snow cover, which in the model is related to the forcing conditions. We do not have observations of snow depth or snow water equivalent at this site so we cannot confirm, but only speculate, that the snow is too shallow in the simulations and therefore does not provide enough insulation to the very low atmospheric temperatures in winter. The low snow cover would then not effectively decouple the soil from the atmosphere.

- The authors did not describe the results of the soil temperature at the deeper layers, which show an improvement in the OGN simulation, compared with the CTL simulation. Why?

  The comparison between simulated and observed soil moisture and soil temperature from 0-100 cm are shown in Figs. 4 and 5 and discussed in Section 4.2. There are no observations, so the soil moisture and temperature in the deepest soil (100-200cm) are not shown.

- Lines 247-248: "The inclusion of an organic soil horizon also affects the hydrologic cycle components such as soil water content, runoff, and evaporation (Figure 7)." I think that this phrase should be removed, because it does not represent which was presented in the figures until this moment.

  Removed

- Lines 250-251: ". . .due to the contrasting water retention characteristics of organic and mineral soil." Do you have reference for boreal forest?

  Added the references and revised the original sentence to read "due to the contrasting water retention characteristics of organic and mineral soil (Koven et al., 2009; Rinke et al., 2008; Lawrence and Slater, 2008), the higher porosity in OGN leads to an increase in total soil water content, while the lower topsoil temperatures (Figure 4a) in OGN enhanced the ice content, then decreases the liquid soil water content."

- Figure 12 is called before figure 8. Please verify the number of the figures.

  Figures number verified.

- Lines 275-276: "The OGN-CTL difference is strongest for the drought years 2001, 2002 and 2003." I did not find this result based on figures 4-7.

  Delete this sentence.

- Did you calibrate the parameters used in the model? What were the parameters values used?

  We did not explicitly "calibrate" model parameters. However, we conducted a number of soil parameter sensitivity tests and selected parameter values based on literature (Lawrence and Slater 2008, Letts et al. 2000). This is discussed in Section 3.1.

- I think the results need to be explored further in the section 4.4. This way is confusing to understand. I suggest that the authors should focus in the comparison of the OGN and CTL errors for each season in the drought and wet years. And, include a discussion these results.

  We reorganized several sections and now added more explanations concerning, for instance, the overall impact of adding an organic soil layer in OGN and the high bias in modeled sensible heat fluxes.

- Lines 282-284: "In general, the OGN parameterization improved the simulation of daily daytime SH and LH in terms of both RMSE and IOA (Table 4)." Rewrite this sentence, because the RMSE of the SH from OGN simulation is higher than the CTL simulation in all years (exception for 2005).

  This sentence is revised and reads as "In general, the OGN parameterization improved the simulation of daily daytime LH in terms of both RMSE and IOA, and increased IOA for SH (Table 3). Nevertheless, compared with CTL, OGN increased the bias in SH slightly by ~6% (Table 3). The reason for the general high bias in both CTL and OGN will be explored in Section 4.4".

- Lines 292-293: "OGN overestimates daytime SH compared with observations, while CTL underestimates daytime SH for spring and summer (Figure 8a, b),. . ." The both simulations overestimate the SH in summer.

  Yes, it is corrected to read "OGN overestimates daytime SH compared with observations, while CTL underestimates daytime SH for spring (Figure 8a, b) and both OGN and CTL slightly overestimates SH for summer, autumn and winter (Figure 8b, c, and d)"

- Lines 294-295: Why did not you show the figure with the cycle of the soil temperature? The authors include a description of this result, but they did not show the figure associated. I think this figure should be included in the text.

  The main features of annual cycle of soil temperature are shown in its monthly cycle in Figure 6.

- Lines 305-307: "Note that the OGN simulation also improves surface heat fluxes significantly in drought years, because the snowmelt process dominates during spring months." In drought years, the OGN simulation did not improve the SH, compared with the CTL simulation in spring. Note that the bias of the SH from OGN simulation is higher than the CTL simulation in spring.

Modified the original sentence to read "Note that the OGN simulation also improves latent heat fluxes in drought years, because the snowmelt process dominates during spring months".

- Do the curves of the diurnal cycle of the figures 8 and 9 represent the daytime, nighttime or mean?

These results are seasonally averaged diurnal cycles of heat fluxes.

- Section 4.5: Why did not the authors show the figure with the annual cycle of the soil temperature?

The main features of annual cycle of soil temperature are shown in its monthly cycle Figure 6.

- Section 4.5: It is interesting that the authors mention that the annual cycle shows that there has been an improvement (closer to the observations) in the latent and sensible heat fluxes and volumetric liquid water from the OGN simulation in spring for wet years, in relation to the CTL.

- Conclusions: The authors repeated the results. The conclusions should contain the principal results found and the suggested hypothesis or explanations associated to these results. As I mentioned before, I think the authors should focus the improvement of the OGN simulation based on the observations and the CTL simulation.

Good suggestion! We revised it to remove redundancy and to reflect main results.

- Lines 369-370: "The incorporation of an organic layer at the top of the soil helps improve the nighttime sensible heat flux for all seasons." The authors did not mention about the nighttime sensible heat flux in their results. I think the authors should mention it in their results or they should remove this sentence of the conclusions.

Deleted this sentence. Due to uncertainties in nighttime flux measurements, we focused our analysis on daytime observation data.

Technical corrections:

- Line 7: ". . .multi-parameterization. . ." Niu et al. (2011) and Yang et al. (2011) use multiparameterization. - Lines 49-50: "…(Letts et al. 2000, Beringer et al. 2001, Molders and Romanovsky 2006, Nicolsky et al. 2007, Lawrence and Slater 2008, etc.)." I think it is would be better "...(e.g., Letts et al. 2000, Beringer et al. 2001, Molders and Romanovsky 2006, Nicolsky et al. 2007, Lawrence and Slater 2008)."

Revised and reads as "Niu et al. (2011) and Yang et al. (2011) use the Noah LSM with multi-parameterization options (Noah-MP) discussed the seasonal and annual cycles of snow, hydrology, and vegetation."

Revised and reads as by "...(e.g., Letts et al. 2000, Beringer et al. 2001, Molders and Romanovsky 2006, Nicolsky et al. 2007, Lawrence and Slater 2008)."

- Lines 72-73: “The forest regenerated after a natural fire in 1919 and had a 1998 stand density of 830 stems ha-1.” I think this sentence is confused, it could be replaced by "The forest was regenerated after a natural fire in 1919, and in 1998 it had a stand density of 830 stems ha-1.”

This sentence is revised and reads as “The forest was regenerated after a natural fire in 1919, and in 1998 it had a stand density of ~830 stems ha-1”

- Line 94: I think the authors should include in the manuscript the meaning of the variable theta.

Replace “resulting in high VWC values that may not be characteristic of the flus footprint. Theta is also measured at 2.5- and 7.5-cm depth in the forest-floor LFH layer using two profiles” by “resulting in high Volumetric Water Content (VWC) values that may not be characteristic of the flux footprint. VWC is also measured at 2.5- and 7.5-cm depth in the forest-floor LFH layer using two profiles”

- Line 101: “The net radiation flux density Rn was calculated. . .” The authors should correct this phrase for “The net radiation flux density (Rn) was calculated. . .” or “The net radiation flux density, Rn, was calculated. . .”

This sentence is revised and reads as “The net radiation flux density, Rn, was calculated.”

- Lines 128-130 and 135-137 are the same sentence. - Review the figure captions, especially the figures 6 and 7.

Rewrote section 3.1