Letter of Responses

Authors' note: The original reviewers' comments are *in italic and colored blue*, and our responses follow. All page/line numbers indicated in the responses are those in the revision.

Reviewer 1 (Dr. W. Wieder)

Liang and coauthors present a nice study exploring sensitivities in the E3SM land model to changes in the calculation of soil water potential and subsequently to parametric changes related to plant physiology. They compare simulated results to observations from the MOFLUX site, focusing on carbon fluxes (GPP and soil respiration, SR).

There are a host of changes suggested that generally improve agreement with observed results, but in general it's hard to follow what changes are most important for the improvements. I appreciate the need to keep text and display items simple & digestible for readers, but a bit more complexity would help shed light on the factors responsible for the site-level improvements in the model made here. For example, it looks like the modified soil water potential scheme (Hanson, I think), provides a better fit to GPP, SR and soil moisture (Figs. 2-5), but it remains unclear if the modifications are significantly better (or different) from the Clapp & Hornberger scheme that's been tuned to local edaphic characteristics? It's not that surprising that the parameterization for a global model would not be a good fit to local results, so does the model just need tuning for site-level runs, or are underlying physics and assumptions in the Hanson scheme fundamentally superior to another approach? Addressing this question matters if the long-term aim of this work is to document changes made to ELM from CLM4.5.

Response: We greatly appreciate the valuable comment. Particularly interesting is the question *"if the modifications are significantly better (or different) from the Clapp & Hornberger scheme that's been tuned to local edaphic characteristics"*. In the revised manuscript, we compared the simulated gross primary production (GPP) and soil respiration (SR) when using the Hanson model and the calibrated Clapp & Hornberger model. Generally, both the Hanson model and the calibrated Clapp & Hornberger model improved the simulation of GPP and SR in the E3SM Land Model version 0 (ELMv0), in comparison with the default run with the uncalibrated Clapp & Hornberger model (Fig. S8 and also see below). The ELMv0 with the Hanson model consistently produced higher GPP and SR than that with the calibrated Clapp & Hornberger model. In comparison with the observations, the modeled SR generally fell within the 1 sigma (i.e., standard deviation) range of observations, by using both the Hanson model and the calibrated Clapp & Hornberger model. However, the modeled GPP with the calibrated Clapp & Hornberger model was lower than the observations. Given the goodness-of-fit of the soil water potential (SWP)-volumetric water content (VWC) relationship was default Clapp & Hornberger model < calibrated Clapp & Hornberger model < calibrated Hanson model (revised Table 1 and also see below), these new results further support our conclusion that better representations of SWP can improve the simulations of carbon processes (i.e., GPP and SR here).

In the revised manuscript, we added a new supplementary figure (i.e., Fig. S8) and a paragraph in the text to compare simulations with the Hanson model and the calibrated Clapp & Hornberger model (page 12, lines 10 – 19):

"Moreover, we also explored whether the calibrated Clapp & Hornberger model can lead to similar improvements with the Hanson model (Fig. S8). Generally, both the Hanson model and the calibrated Clapp & Hornberger model improved the simulation of GPP and SR in the ELM, in comparison with the default run (Fig. S8). The ELMv0 with the Hanson model consistently produced higher GPP and SR than that with the calibrated Clapp & Hornberger model. In comparison with the observations, the modeled SR generally fell within the 1 sigma (i.e., standard deviation) range of observations, by using both the Hanson model and the calibrated Clapp & Hornberger model. However, the modeled GPP with the calibrated Clapp & Hornberger model was still lower than the observations. Given the goodness-of-fit of the soil water potential (SWP)-volumetric water content (VWC) relationship was default Clapp & Hornberger model < calibrated Clapp & Hornberger model < calibrated Hanson model (Table 1), these results further support the conclusion that better representations of SWP can improve the simulations of carbon processes. Therefore, throughout the remainder of this manuscript, we used the Hanson model to represent the SWP-VWC relationship."

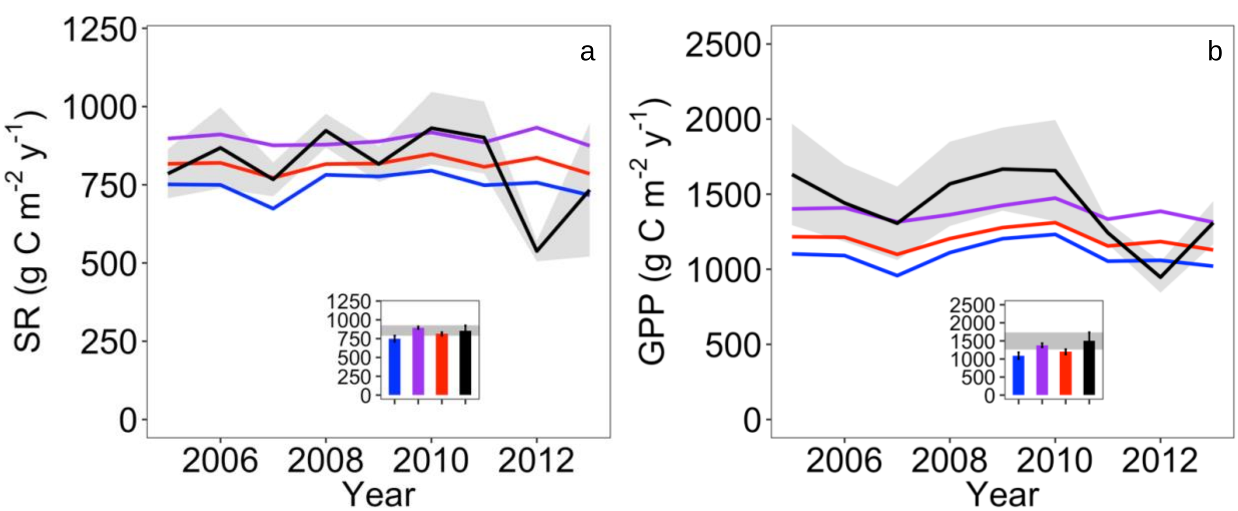

Figure S8: Annual soil respiration (SR) and gross primary production (GPP). Blue lines are the ELMv0 simulations with default parameters ($MOD_{default}$), red lines with the soil water potential improved using the calibrated Clapp & Hornberger model (MOD_{cCP}), and purple lines with the soil water potential improved using the Hanson model (MOD_{H}). Black lines and grey area are the observed (OBS) mean and 1 sigma range, which were calculated from 8 field replications for SR, and from three different net ecosystem exchange partitioning methods for GPP. The inserted bar plots are mean annual average ± 1 sigma across 2005-2011.

The reviewer also asked *"are underlying physics and assumptions in the Hanson scheme fundamentally superior to another approach"*. The short answer is no. Although different approaches for soil water retention curve may have different underlying physics and assumptions, pragmatic models are those which have been well calibrated/parameterized with empirical data. Since the default ELMv0 simulated the SWP poorly at the MOFLUX site (Fig. 3b), one important question we asked in this study was whether better representation of SWP in the model would improve the simulations of carbon processes. To improve the SWP simulation as much as possible, our effort was not limited to tuning the Clapp & Hornberger model since we did not know whether the tuned Clapp & Hornberger model would be good enough to answer the question. Instead, we evaluated a series of soil water retention curve models popularly used in the literature to derive the best-fit model using root-mean-square-error (RMSE) and Akaike Information Criterion (AIC) as suggested by the reviewer in another comment (revised Table 1 and also see below). The Hanson model performed the best, showing the smallest RMSE and AIC values. Both the modeled annual fluxes of GPP and SR fell within the 1 sigma range of observations when using the Hanson model, but not with the calibrated Clapp & Hornberger model as shown in Fig. S8 and above. Thus, we used the Hanson model for all further analyses. In the reminder of this response letter and the manuscript, the improved SWP was simulated using the Hanson model if not otherwise specified, and all changed ELMv0 simulations were compared to the default simulations with the default Clapp & Hornberger model.

Table 1. Root-mean-square-error (RMSE) and Akaike Information Criterion (AIC) of different models in simulating the SWP-VWC relationship for the soil in the MOFLUX site at two depths: 0 to 30 cm and below 30 cm.

	< 30 cm		> 30 cm	
Model	RMSE	AIC	RMSE	AIC
Clapp & Hornberger (default ELMv0)	4.25	157.82	1.33	18.51
Brooks & Corey	3.91	151.05	1.13	13.51
Clapp & Hornberger (calibrated)	0.53	-61.03	0.51	-23.43
Fredlund & Xing	0.51	-63.15	2.43	47.13
Hanson	0.41	-86.07	0.34	-38.98
van Genuchten	0.50	-65.53	0.36	-36.61

Although it needs further exploration as to whether the Hanson model performs the best on the regional and global scales, the default Clapp & Hornberger model used in the ELMv0 performs poorly in simulating SWP on the global scale (See below), which may significantly impact the biogeochemical simulations. In a different (but related) project, we tested the simulated SWP by the default Clapp & Hornberger model used in the ELMv0 against 6928 data points of paired measurements of SWP and VWC across different soil types and ecosystems. Results showed that the default Clapp & Hornberger model used in the ELMv0 was not able to reproduce the observed SWP (see Fig. R1 below). It remains unclear which model will perform best in describing the SWP-VWC relationship on the global scale. More work will be needed to explore the issue, which is beyond the scope of this manuscript.

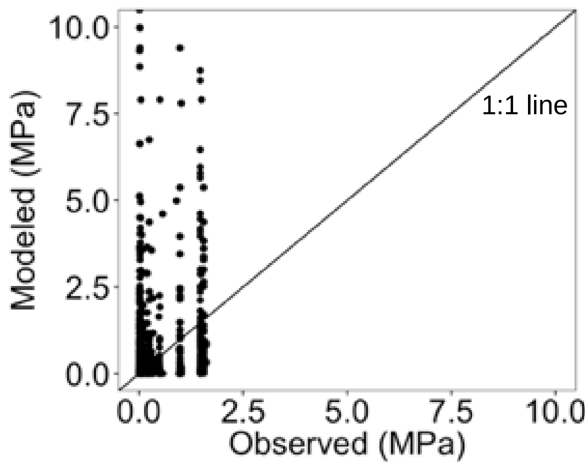

Figure R1: Comparison of observed and simulated soil water potential (-MPa) across different soil types and ecosystems by the default Clapp & Hornberger model in the ELMv0.

Similarly, how important are the suggested parameter changes for capturing the annual cycle of LAI and C fluxes vs. changes to the soil moisture scheme (Fig. 6). Stepping through these changes sequentially in the text and display items will clarify the source(s) of the improvements.

Response: The reviewer provided a great suggestion. Thanks to the reviewer's other suggestion, we are able to more clearly show the importance of the parameter changes and the improved SWP by plotting the mean annual cycle (± 1 sigma) of LAI, GPP and SR (revised Fig. 4 and also see below). With the revised figure, we can step through the changes. Results showed that the ELMv0 with both the default and improved SWP by the Hanson model overestimated the maximum LAI (Fig. 4a). The parameter changes significantly reduced the maximum LAI to better match the observations (Fig. 4a). The parameter changes further increased the simulated GPP and SR during the peak growing season, in addition to the improvement by the adjusted SWP (Fig. 4b, c). However, all modifications of the ELMv0 still overestimated SR during the non-peak growing season, as discussed (previously) in Section 4.2.

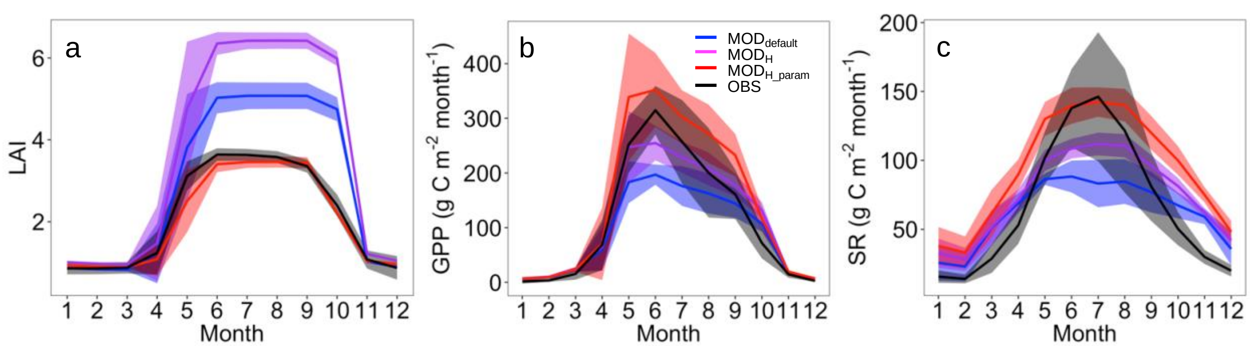

Figure 4 The annual mean cycles of leaf area index (LAI), gross primary production (GPP) and soil respiration (SR). OBS: observation; $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement by the Hanson model; MOD_{H_param}: model output after soil water potential improvement by the Hanson model and parameter adjustments.

In the revised manuscript, we revised the text accordingly to show the effect of the parameter changes on the simulations of LAI, GPP and SR (page 10, line 6 – 10):

"Results showed that the ELMv0 with both the default and improved SWP by the Hanson model overestimated the maximum LAI (Fig. 4a).The adjustment of the aforementioned five parameters (Table 2) significantly reduced the LAI to within a more reasonable range (Fig. 4a). The parameter changes further increased the simulated GPP and SR during the peak growing season, in addition to the improvement by the adjusted SWP (Fig. 4b, c). However, all modifications of the ELMv0 still overestimated SR during the non-peak growing season, resulting in significant overestimation of annual SR fluxes (Fig. S5a). After the parameter adjustments, the annual GPP flux was still within the observed range (Fig. S5b)."

Finally, although the authors claim that improving SWP directly improved soil respiration estimates, it's not clear if this is a direct effect of soil moisture on soil respiration, or merely reflective of the larger plant and soil C stocks simulated as a result of having higher GPP. Concurrently presenting changes to ecosystem C stocks and the soil moisture effect on GPP and heterotrophic respiration (btran and w_scalar, respectively in CLM4.5) will help clarify how / why improvements were made.

Response: We appreciate the insightful comment. The second reviewer had a similar concern. We agree with the reviewers that concurrently presenting changes in ecosystem carbon stocks and the soil moisture effect on GPP (btran) and heterotrophic respiration (ξ_W in equation 9) will help clarify how improvements were made. In the revised manuscript, we analyzed the changes in btran, ξ_W and soil organic carbon (SOC) (Fig. S2-S3 and also see below). The improved soil water scheme using the Hansom model increased both btran and ξ_W during the peak growing season, and reduced ξ_W during the non-peak growing season (Fig. S2). The change in ξ_W was generally consistent with that of SWP (Fig. 3b). While the model simulated SOC with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations increased SOC stocks, matching the reviewer's expectation (Fig. S3). These results, combining with previous results (the original Fig. 4, which was moved as Fig. S1 in the revision as suggested by the reviewer's other comment), indicate that the improved soil respiration by SWP was a joint result of changes in GPP, SOC stocks and the moisture modifier of heterotrophic respiration.

In the revised manuscript, we added these new figures, presented the results and discussed details in the text.

"The changes in annual SR and GPP (i.e., the differences between before and after the improved SWP simulation using the Hanson model) showed a linear relationship (Fig. S1). In addition, the improved soil water scheme using the Hanson model increased both the moisture modifiers of GPP and heterotrophic respiration (i.e., btran and ξ_W) during the peak growing season, and reduced ξ_W during the non-peak growing season (Fig. S2). While SOC when simulated by the model with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations using the Hanson model increased SOC stocks (Fig. S3)."

"Constraining the SWP-VWC relationship with site-specific data and using the Hanson model instead of the ELMv0 default model (Fig. 1) significantly improved the model representation of SWP (Fig. 3) and annual SR (Fig. 2a). The improvements in model fits were due to changes in GPP, SOC stocks, and the moisture modifier on heterotrophic respiration (Figs. S1 – S3). First, the default ELMv0 underestimated GPP (Fig. 2b), as in a recent study where CLM4.5 significantly underestimated GPP at a coniferous forest in northeastern United States (Duarte et al., 2017). GPP affects the substrate supply for SR, as evidenced by the close relationship between changes in SR and GPP (Fig. S1), which is consistent with experimental evidence showing GPP can directly affect the magnitude of root respiration (Craine et al., 1999; Högberg et al., 2001; Wan and Luo, 2003; Verburg et al., 2004; Gu et al., 2008). Second, the changed soil moisture scheme increased the moisture modifier (ξ_W*) on heterotrophic respiration during the peak-growing season, and decreased it during the non-peak growing season (Fig. S2), which is consistent with the trend of changes in SWP (Fig. 3). In addition, the changed soil moisture scheme also increased the simulated SOC stock, the substrate for heterotrophic respiration (Fig. S3). These changes together resulted in the improvement of simulated SR."*

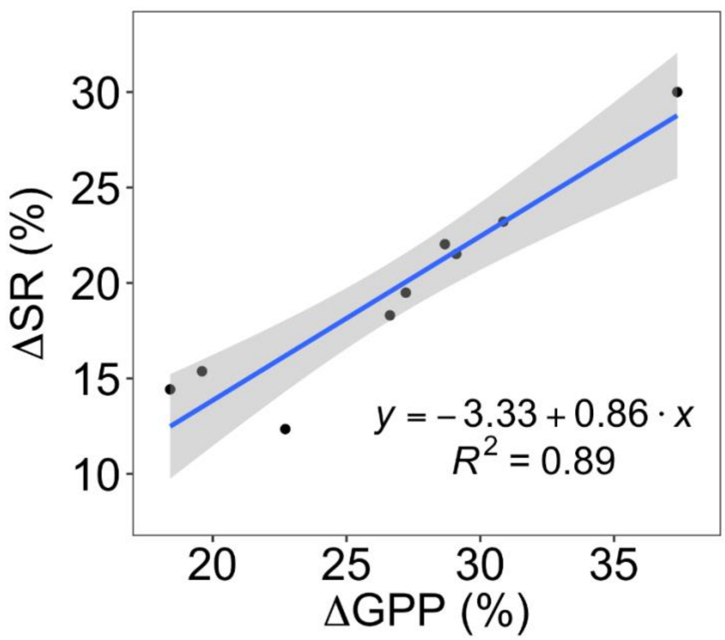

Figure S1: Relationship between changes in simulated annual soil respiration (ΔSR) and gross primary production (ΔGPP) induced by improvement of soil water potential using the Hanson model.

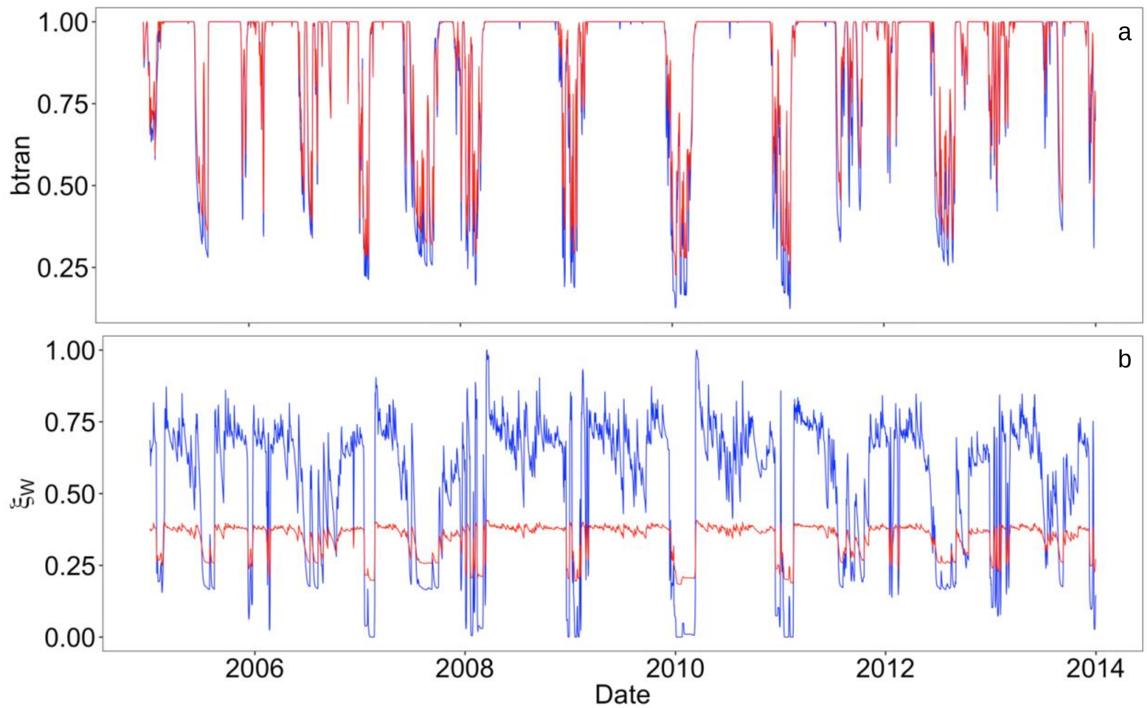

Figure S2 Impact of the changed SWP on the moisture modifiers of GPP (btran, a) and heterotrophic respiration (ξ_W, b). $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model.

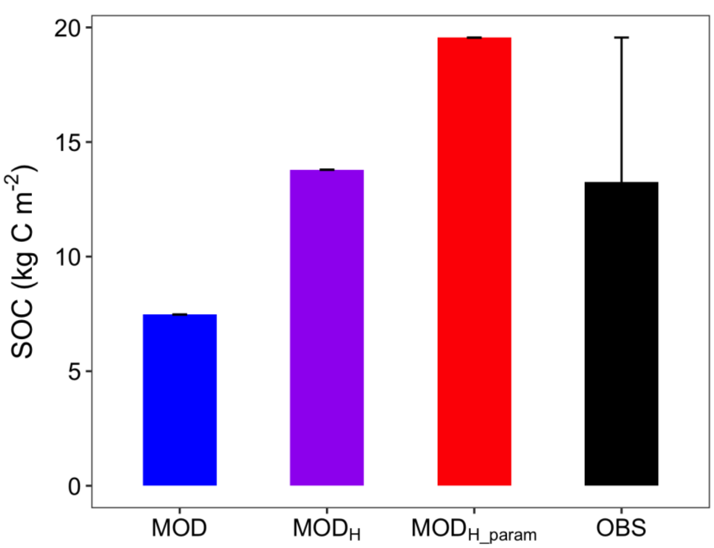

Figure S3 Comparison of the observed and modeled soil organic carbon (SOC) stocks. OBS: observation; MOD: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model; MOD_{H_param}: model output after soil water potential improvement using the Hanson model and parameter adjustments.

Major concerns

I appreciate the effort used to explore alternative formulations for SWP in the model (Fig. 1, Table 1). Two questions come to mind. First, is it worth doing a more thorough model selection process like AIC or BIC that penalizes more complex models for their additional parameters instead of just showing RMSE. Second, as E3SM is intended to be run in global simulations, I wonder what effect alternative formulations for SWP have on water and energy fluxes from the model in site level, and ultimately global, simulations? The GPP results (Fig. 2) are a good start for this, but presumably these changes really modify ET fluxes (and runoff). It seems documenting these changes are likely important (if only in SI)?

Response: The reviewer provided valuable comments in terms of the model selection and the effect of the alternative formulations for SWP on water and energy fluxes. In the revised manuscript, we added Akaike Information Criterion (AIC), in addition to root-mean-square-error (RMSE), for the model selection. Both AIC and RMSE indicated that the Hanson model was the best in simulating the SWP-VWC relationship (i.e., smallest AIC and RMSE values; revised Table 1 and also see below).

Table 1. Root-mean-square-error (RMSE) and Akaike Information Criterion (AIC) of different models in simulating the SWP-VWC relationship for the soil in the MOFLUX site at two depths: 0 to 30 cm and below 30 cm.

	< 30 cm		> 30 cm	
Model	RMSE	AIC	RMSE	AIC
Clapp & Hornberger (default ELMv0)	4.25	157.82	1.33	18.51
Brooks & Corey	3.91	151.05	1.13	13.51
Clapp & Hornberger (calibrated)	0.53	-61.03	0.51	-23.43
Fredlund & Xing	0.51	-63.15	2.43	47.13
Hanson	0.41	-86.07	0.34	-38.98
van Genuchten	0.50	-65.53	0.36	-36.61

In addition, we analyzed changes in simulated evapotranspiration (ET) and runoff as suggested. We plotted the mean annual cycle (± 1 sigma) of both ET and runoff (Fig. S6 and also see below). The change in soil moisture scheme using the Hanson model and parameter adjustments slightly increased ET and decreased runoff. Despite these slight changes, the model-simulated ET generally fell within the observed range, with or without changes in soil water scheme and parameters.

In the revised manuscript, we added a paragraph in the Results section to describe the changes in ET, runoff and other variables (page 10, line 11 – 16):

"In addition, we analyzed changes in simulated evapotranspiration (ET), runoff, photosynthesis, net primary production, C allocations to fine roots, leaf and woody tissue in response to the changes in the soil water scheme and parameters (Fig. S6, S7). The change in soil moisture scheme and parameter adjustments slightly increased ET and decreased runoff.

Despite these slight changes, the model simulated ET generally fell within the observed range, with or without changes in soil water scheme and parameters (Fig. S6).”

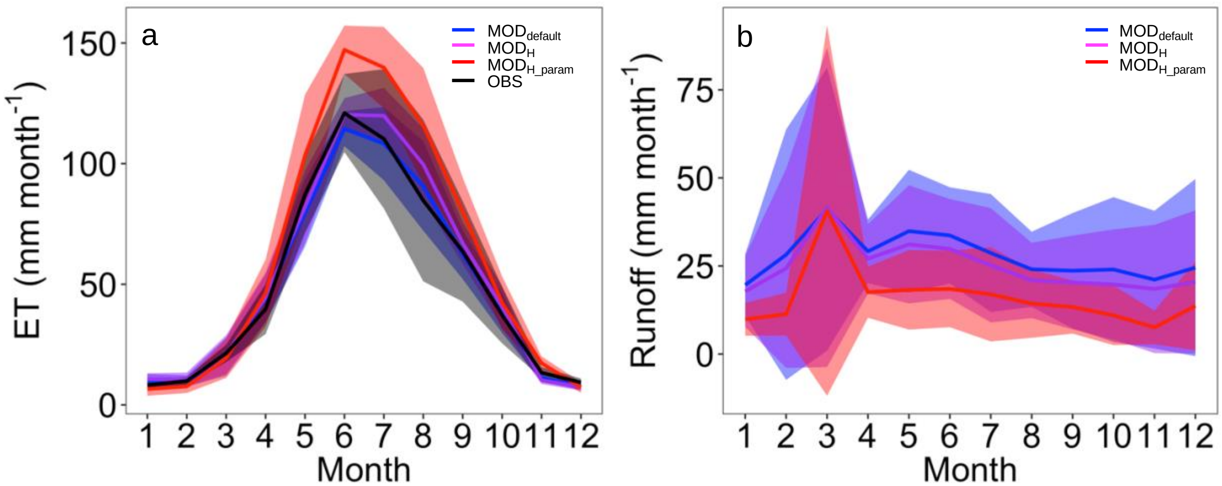

Figure S6 Modeled evapotranspiration (ET) and runoff in response to the improved SWP and parameter adjustments. OBS: observation; $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model; MOD_{H_param}: model output after soil water potential improvement using the Hanson model and parameter adjustments.

If Hanson or van Genuchten formulations are ‘better’ fits to the observations, why aren’t they used for GPP simulations in Fig. 2? What’s the purpose of exploring alternative SWP schemes, if they don’t follow through the C cycle simulations in the model? Reading the text on the bottom of page 7, however, maybe (MODswp) is using the Hanson scheme? If so, does the calibrated Clapp & Hornberger approach provide similar improvement by removing the high bias in the default configuration (Fig. 3). Please clarify in the text and figure captions what’s being shown and why none of the models adequately capture the effect of the 2012 drought.

Response: We apologize for the unclear description in the manuscript. The Hanson model was used through the carbon cycle simulations in the model. As shown in Fig. S2, the changed soil water scheme had impacts on both the moisture modifiers of GPP and heterotrophic respiration. We have revised the text and figure captions to make it clearer.

In response to the question *“does the calibrated Clapp & Hornberger approach provide similar improvement by removing the high bias in the default configuration”*, we conducted an additional analysis as described in the response to the first comment above. We compared the simulated gross primary production (GPP) and soil respiration (SR) when using the Hanson model and the calibrated Clapp & Hornberger model. Generally, both the Hanson model and the calibrated Clapp & Hornberger model improved the simulation of GPP and SR in the ELMv0, in comparison with the default run (Fig. S8 and also see below). The ELMv0 with the Hanson model consistently produced higher GPP and SR than that with the calibrated Clapp & Hornberger model. In comparison with the observations, the modeled SR generally fell within the 1 sigma range of observations, for both the Hanson model and the calibrated Clapp & Hornberger model. However, the modeled GPP with the calibrated Clapp & Hornberger model was significantly lower than the observations. Given the goodness-of-fit of the soil water potential (SWP)-volumetric water content (VWC) relationship was default Clapp & Hornberger model < calibrated Clapp & Hornberger model < Hanson model (revised Table 1 and also see below), these new results further support our conclusion that better representations of SWP can improve the simulations of carbon processes (i.e., GPP and SR here).

In the revised manuscript, we added a new supplementary figure (i.e., Fig. S8) and a paragraph in the text to compare simulations with the Hanson model and the calibrated Clapp & Hornberger model (page 12, lines 10 – 19):

"Moreover, we also explored whether the calibrated Clapp & Hornberger model can lead to similar improvements with the Hanson model (Fig. S8). Generally, both the Hanson model and the calibrated Clapp & Hornberger model improved the simulation of GPP and SR in the ELM, in comparison with the default run (Fig. S8). The ELMv0 with the Hanson model consistently produced higher GPP and SR than that with the calibrated Clapp & Hornberger model. In comparison with the observations, the modeled SR generally fell within the 1 sigma (i.e., standard deviation) range of observations, by using both the Hanson model and the calibrated Clapp & Hornberger model. However, the modeled GPP with the calibrated Clapp & Hornberger model was still lower than the observations. Given the goodness-of-fit of the soil water potential (SWP)-volumetric water content (VWC) relationship was default Clapp & Hornberger model < calibrated Clapp & Hornberger model < calibrated Hanson model (Table 1), these results further support the conclusion that better representations of SWP can improve the simulations of carbon processes. Therefore, throughout the remainder of this manuscript, we used the Hanson model to represent the SWP-VWC relationship."

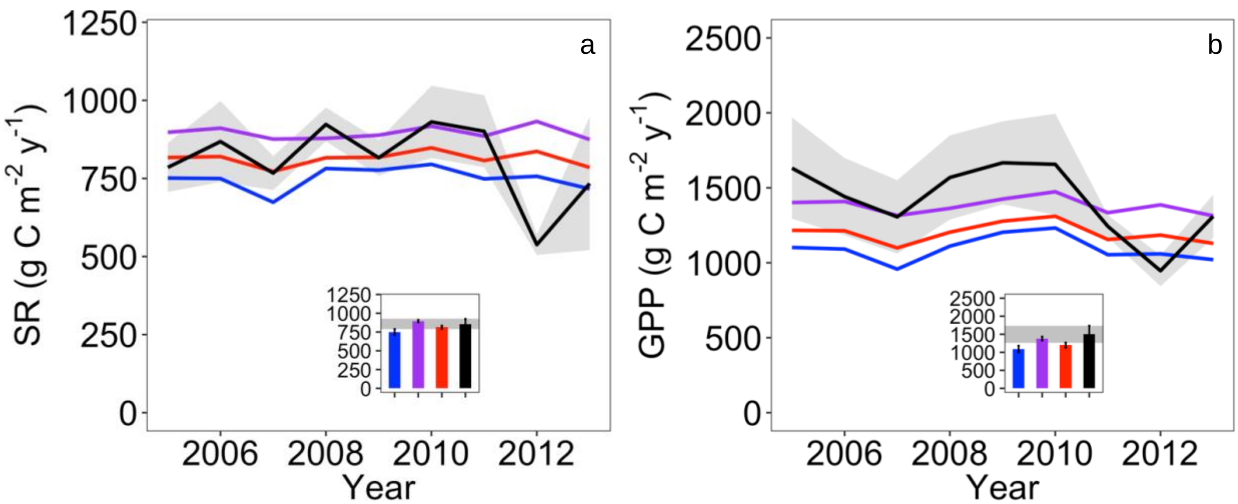

Figure S8: Annual soil respiration (SR) and gross primary production (GPP). Blue lines are the ELMv0 simulations with default parameters ($MOD_{default}$), red lines with the soil water potential improved using the calibrated Clapp & Hornberger model (MOD_{cCP}), and purple lines with the soil water potential improved using the Hanson model (MOD_{H}). Black lines and grey area are the observed (OBS) mean and 1 sigma range, which were calculated from 8 field replications for SR, and from three different net ecosystem exchange partitioning methods for GPP. The inserted bar plots are mean annual average ± 1 sigma across 2005-2011.

In response to the question "*why none of the models adequately capture the effect of the 2012 drought*", we revised Section 4.2 to include more discussion. Failing capturing the effect of the 2012 drought, as well as the underestimated seasonal and interannual variabilities of GPP and SR (Fig. 2, 4), indicate that the current model structure is not sensitive enough to environmental changes. Potential reasons include lacking representations of microbial organisms, macroinvertebrate and other forest floor and soil fauna and root exudates. We discussed in detail in Section 4.2 as:

"Although the simulation of the SWP using the Hanson model improved the representation of both annual SR and GPP, the model continued to overestimate SR during the non-peak growing season (Fig. 4), resulting in significant overestimation of the annual SR fluxes (Fig. S5). In addition, no matter which SWP simulations were used, the ELMv0 had smaller interannual variability than the observations (Fig. 2). Specifically, the model was not able to capture the steep decreases in GPP and SR in the extreme drought year (i.e., 2012). These results indicate that the current model structure is not sensitive enough to environmental changes. A few potential reasons may contribute to the underestimated seasonal and interannual variability. In the ELMv0, heterotrophic respiration contributed a majority proportion (i.e., over 85%) of the total SR during non-growing seasons (Fig. 5), suggesting that the overestimation of SR was primarily due to the biased simulation of heterotrophic respiration. A potential reason for the biased simulation of heterotrophic respiration may be related to the temperature sensitivity (Q_{10}). Theoretically, a higher Q_{10} can result in greater seasonal variability of SR (Fig. S9). Compared to relatively small Q_{10} values, a larger Q_{10} can lead to lower heterotrophic respiration when temperature is below the reference temperature, and greater heterotrophic respiration when temperature is above the reference (Fig. S9). In the ELMv0, the reference temperature is 25 ºC and the Q_{10} of heterotrophic respiration is 1.5 (Oleson et al., 2013). A previous study derived a much greater Q_{10} value (i.e., 2.8) when the parameters were calibrated with data from another temperate forest (Mao et al., 2016). We hypothesized that the Q_{10} value of 1.5 may be too small for the MOFLUX site. We arbitrarily increased Q_{10} from 1.5 to 2.5, but there were minimal effects on the SR simulation (Fig. S10). This indicates that modifying the temperature sensitivity of heterotrophic respiration may not improve the modeled representation of seasonality of SR in the ELMv0.

Another potential reason for the biased simulation of heterotrophic respiration may be that the seasonality of microbial organisms was not adequately represented in the model. Like most ESMs, the ELMv0 represents soil C dynamics using linear differential equations and assumes that SR is a substrate-limited process in the model. However, producers of CO_2 in soils, microbial organisms, have a significant seasonal cycle (Lennon and Jones, 2011). These organisms usually have very high biomass and activity during the peak growing season, with favorable conditions of temperature, moisture and substrate supply, and tend to be dormant under stressful conditions in non-peak growing seasons (Lennon and Jones, 2011; Stolpovsky et al., 2011; Wang et al., 2014; Wang et al., 2015). The seasonality of microbial biomass and activity, in addition to that of GPP and ST, may contribute to the seasonal variability of SR.

Additionally, another reason may be related to the model lacking representation of macroinvertebrate and other forest floor and soil fauna. There is a high density of earthworms at *the MOFLUX site (Wenk et al., 2016). Earthworms can shred and redistribute soil C and change soil aggregation structure, which may alter soil C dynamics and CO_2 efflux to the atmosphere (Verhoef and Brussaard, 1990; Brussaard et al., 2007; Coleman, 2008). Like microbial organisms, earthworms usually have a significant seasonal cycle, showing high biomass and high activity during peak growing seasons and tending to be dormant during non-growing seasons (Wenk et al., 2016). However, a recent review suggests that current experimental evidence and conceptual understanding remains insufficient to support the development of explicit representation of fauna in ESMs (Grandy et al., 2016). Therefore, data collection focused on seasonal variations in fauna and microbial biomass and activity might enable further improvements in the representation of seasonal variation in SR.*

Our analyses also showed that the modeled SR was not able to reach the observed peak in many years during the peak growing season even when the modeled GPP exceeded the observation (Fig. S11). In addition, the parameter modification increased GPP during both peak and non-peak growing seasons, resulting in an even greater overestimation of SR during non-peak growing seasons (Fig. S11). These results suggest that simply increasing GPP may not be adequate to increase the seasonal variability of the simulated SR. A potential reason may be that the current model does not include root exudates. Root exudates are labile C substrates that are important for SR (Kelting et al., 1998; Kuzyakov, 2002; Sun et al., 2017). The root exudate rate is primarily dependent on root growth, showing a seasonal cycle in temperate forests (Kelting et al., 1998; Kuzyakov, 2002). Thus, including root exudates in the model may further increase the model simulated SR during the peak growing season without needing to increase GPP."

Minor concerns
Section 2.3 This is really a broader comment on how author groups working with E3SM intend to articulate the version of the model on which they are working, esp. for readers less familiar with nuances of CLM4.5 development branches and subsequent ELM developments. For example, how is this code different from other publications (e.g. Brunke et al. 2016; Riley 2018)?

Section 2.3. Please justify the decision to use the CLM-CN decomposition module for a paper focused on soil respiration when Bonan and others (2013) clearly demonstrated shortcomings of this model version? It seems like the wrong tool for this job?

Section 2.3. This is also a little confusing, as the opening line of the section states the soil biogeochemistry is vertically resolved, but to my knowledge CLM-CN does not apply vertically resolved soil BGC? Please clarify

Response: We appreciate the detailed comments on the model version. We used the ELM version 0 (ELMv0), which is equivalent to the Community Land Model version 4.5 (CLM 4.5). In ELMv0, the soil biogeochemistry can be simulated with one-layer or multi-layer converging trophic cascade (CTC, i.e., CLM-CN) decomposition model. We used the vertically-resolved CTC decomposition in this study. . Variable thickness of the soil profile in Brunke et al. (2016) and lateral energy and hydrological exchanges in Bisht et al. (2018) (which may be paper (Riley, 2018) the reviewer referred to) were not in the ELMv0.

In response to the comment "*It seems like the wrong tool for this job*", we used the ELMv0, which is structurally equivalent to the CLM 4.5. Recently, the E3SM council completed a comprehensive study of the soil biogeochemistry module by benchmarking different approaches with global (e.g., ILAMB) and Ameriflux datasets. Due to satisfactory overall performance of the CTC approach (i.e., CLM-CN), the council recommended the CTC approach as the default and baseline soil decomposition pathway in future ELM development. Therefore, we decided to use the CTC decomposition pathway in each soil layer for our study.

In the revised manuscript, we revised this section as (page 3, line 27 – 29):

"The ELMv0 used in this study is structurally equivalent to the Community Land Model 4.5 (CLM 4.5), which includes coupled carbon and nitrogen cycles (Oleson et al., 2013). In ELMv0, the soil biogeochemistry can be simulated with one-layer or multi-layer converging trophic cascade (CTC, i.e., CLM-CN) decomposition model. We used the vertically-resolved CTC decomposition in this study. "

Page 4, Line 20. Single point runs (especially with CLM) forced with flux tower measurements have a long history that should be acknowledged here.

Response: We add a sentence to acknowledge the long history of single point runs (page 4, line 20 – 22):

"Single-point runs forced with site-level measurements have a long history to evaluate model representations of phenology, net primary production (NPP), transpiration, leaf area index (LAI), water use efficiency, and nitrogen use efficiency (Richardson et al., 2012; De Kauwe et al., 2013; Walker et al., 2014; Zaehle et al., 2014; Mao et al., 2016; Duarte et al., 2017; Montané et al., 2017)."

Page 7, Line 7. What changes were made to the Clapp and Hornberger parameterization, there are lots of hard coded parameters in eq. 11-13.

Response: For the calibration of the Clapp & Hornberger model, instead of using the hard-coded parameters in Eq. 11-13, we calibrated the three parameters (i.e., Ψ_m, θ_s and Ψ_s) in the Clapp & Hornberger model (Eq. 10). We added the sentence in the **Materials and Methods** section (page 6, line 5 – 7):

"For the calibration of the Clapp & Hornberger model, instead of using the hard-coded (default) parameters in Eq. 11-13, we calibrated the three parameters (i.e., Ψ_m, θ_s and Ψ_s) in the Clapp & Hornberger model (Eq. 10)."

Fig 2. Why are observations shown with a black line and purple bar (inset)? Consistency within and among figures will help readers understand display items more easily? Similarly, using the same color for line of the default model and modified model in Figs. 1 and 2 would be helpful One strength of using flux tower data in single point simulations seems to be examining the seasonal cycle of carbon and energy fluxes. This is somewhat lost in Fig. 6, and I wonder if the display item would be more powerful if simulations are results were averaged over the whole *observation record (e.g. just show 1 year instead of 9, as the interannual variability isn't that obvious (and already shown in Fig. 2)*

Response: We really appreciate the great suggestion on the figure display. In the revised manuscript, we used consistent colors. In addition, we re-plotted Fig. 6 (revised Fig. 4) as the mean annual cycle (with 1 sigma) as suggested here and in a few other comments.

Fig 4 is never really discussed and doesn't add much to the paper in my estimation. Can it be removed from the text? More, it follows that that changes in productivity would have a linear effect on soil C stocks and therefore respiration rates in a first order model like CLM-cn (Todd-Brown refs from the text), so the relationship shown here isn't really surprising.

Response: We moved Fig. 4 to Fig. S1 as discussed above.

Out of curiosity, how do simulated soil (or vegetation) C stocks compare with observed stocks at the site? The focus on fluxes is fine, but given that fluxes are linear related with stocks, do suggested modifications to the model improve estimates of fluxes AND stocks for the site?

Response: This is really an insightful comment. We analyzed the modeled soil organic carbon (SOC) stocks in the revised manuscript (Fig. S3 and also see below). Although the improved SWP simulations increased SOC stocks, the model simulated SOC with different soil water schemes and parameters generally fell within the wide range of observations.

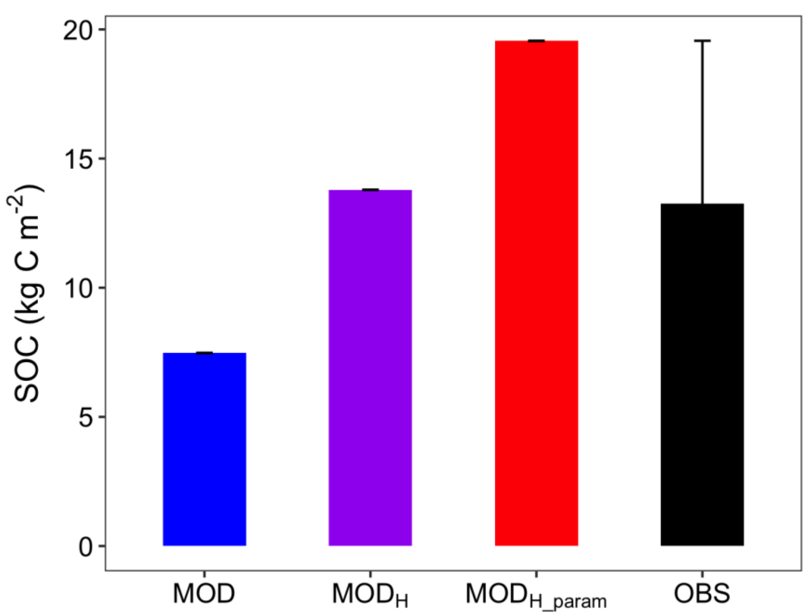

Figure S3 Comparison of the observed and modeled soil organic carbon (SOC) stocks. OBS: observation; MOD: model output before soil water potential improvement; MOD$_H$: model output after soil water potential improvement using the calibrated Hansom model; MOD$_{H_param}$: model output after soil water potential improvement using the calibrated Hansom model and parameter adjustments.

In the revision, we added the description of the result as:

"While the model simulated SOC with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations increased SOC stocks (Fig. S3)."

Seasonal biases in SR and GPP fluxes look pretty bad with default and 'swp' versions of the model (Fig. 5). The parametric changes in Table 2 seem to address some of these seasonal biases (Fig. 6), but it seems like showing the scatter plots on Fig. 5 (maybe with a 3rd color) would be helpful? Along these lines, should both Figs. 5 and 6 show the same 3 simulation ('default', 'swp', and 'swp_param')? Showing the mean annual cycle (+/- 1 sigma on the observations) for all panels in Fig. 6 would help to make this figure easier to digest.

Response: In the revision, we re-plotted Fig. 6 (revised Fig. 4 and also see below) as the mean annual cycle (with 1 sigma) as suggested. We included all 'default', 'H', 'H_param', and 'obs' in the figure. The revised figure can clearly show the seasonal biases in SR and GPP fluxes. Thus, the original Fig. 5 was duplicated by this presentation, so we deleted Fig. 5 in the revised manuscript. Because we moved the original Fig. 4 to Fig. S1, the original Fig. 6 is Fig. 4 in the revised manuscript.

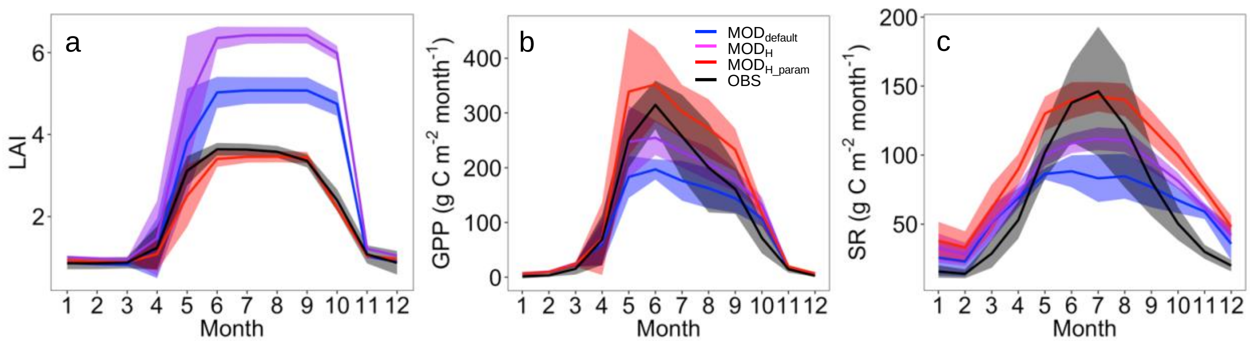

Figure 4 The annual mean cycles of leaf area index (LAI), gross primary production (GPP) and soil respiration (SR). OBS: observation; $MOD_{default}$: model output before soil water potential improvement; MOD_{H}: model output after soil water potential improvement by the Hanson model; MOD_{H_param}: model output after soil water potential improvement by the Hanson model and parameter adjustments.

SLA is something that's measured, maybe not at the site for similar trees to the ones at the site? Is building 3x thicker leaves (Table 2), a reasonable assumption? Similarly, if the authors need to decrease LAI while increasing GPP, flnr necessarily has to increase in the model, but is the 20% increase here supported by databases like TRY, or are these parameter changes just illustrating big nobs in the model that are poorly constrained by observaions?

Response: We appreciate the detailed suggestion on parameter values. Unfortunately, SLA was not measured at the site. The parameter adjustments were based on a surrogate based global optimization using measurements of C and energy fluxes at the site (Lu et al., 2018). The TRY database showed that the SLA for broadleaved deciduous forest ranges from < 0.0005 to > 0.005

m^2 g^{-1} C, with mean values of 0.015 m^2 g^{-1} C (Kattge et al., 2011). Thus, the adjustment of the parameter *slatop* fell within the range of observations.

Page 10, line 12 please report statistics to support claims being made. Visually, the red line looks closer to the observations than the blue one (Fig. 6b,c). How do the annual totals look?

As with comment above, how do changes in annual fluxes or total stocks compare with observations following parametric changes suggested in Table 2?

Response: We replotted Fig. 6 (Fig. 4 in the revision) to show the annual cycle ± 1 sigma considering all 9 years of data. With the revised figure, we can methodically step through the changes. Results showed that both the default and improved SWP using the Hanson model overestimated the maximum LAI (Fig. 4a). The parameter changes significantly reduced the maximum LAI to better match the observations (Fig. 4a). The parameter changes further increased the simulated GPP and SR during the peak growing season, in addition to the improvement by the adjusted SWP (Fig. 4b, c). However, the ELMv0 still overestimated SR during the non-peak growing season.

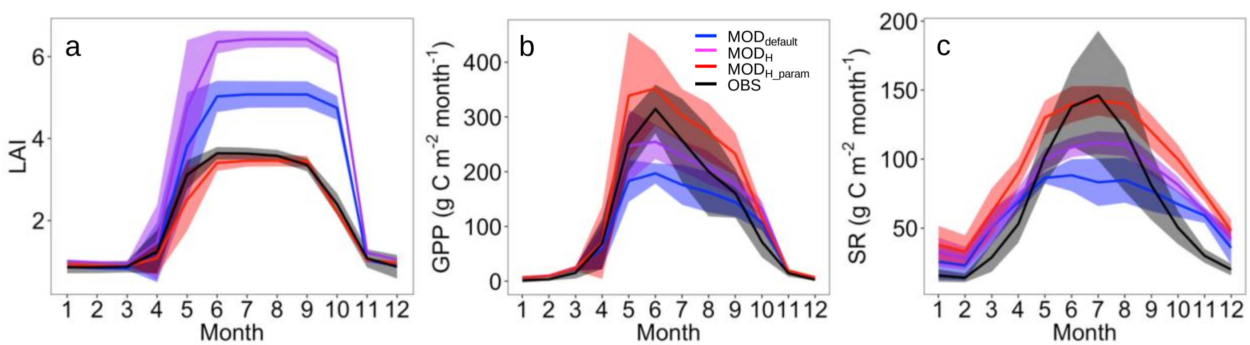

Figure 4 The annual mean cycles of leaf area index (LAI), gross primary production (GPP) and soil respiration (SR). OBS: observation; $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement by the Hanson model; MOD_{H_param}: model output after soil water potential improvement by the Hanson model and parameter adjustments.

To answer the question "*how do the annual totals look*", we analyzed the mean annual fluxes of GPP and SR (Fig. S5 and also see below) and SOC stocks (Fig. S3 and also see below). After the parameter adjustments, GPP was still within the observed ranges, while SR was significantly overestimated due to the overestimation of SR during the non-peak growing season. In Section 4.2, we discussed the potential reasons, including Q_{10} and representations of microbial organisms, macroinvertebrate and other forest floor and soil fauna.

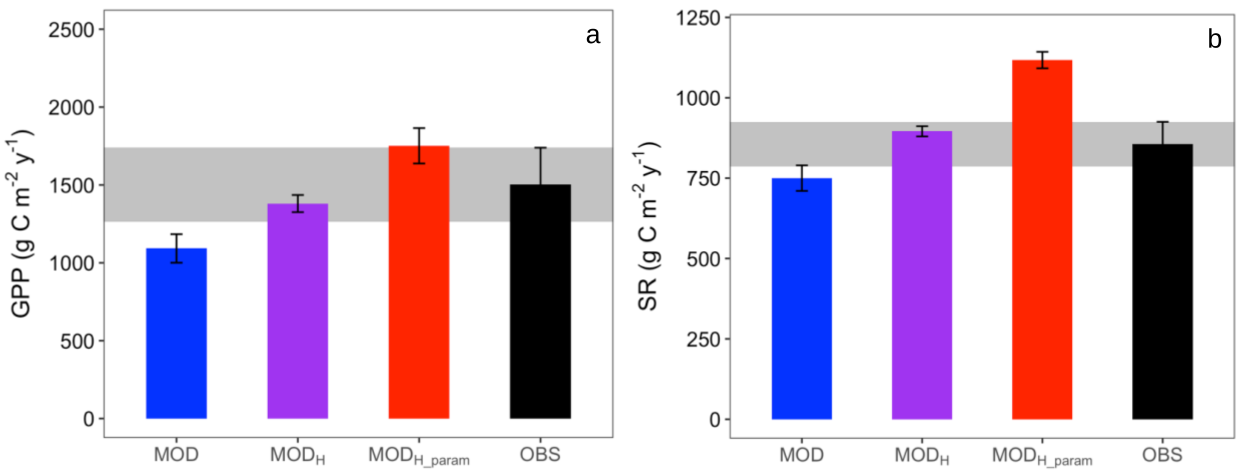

Figure S5 Comparison of the observed and modeled gross primary production (GPP) and soil respiration (SR). OBS: observation; MOD: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the calibrated Hansom model; MOD_{H_param}: model output after soil water potential improvement using the calibrated Hansom model and parameter adjustments.

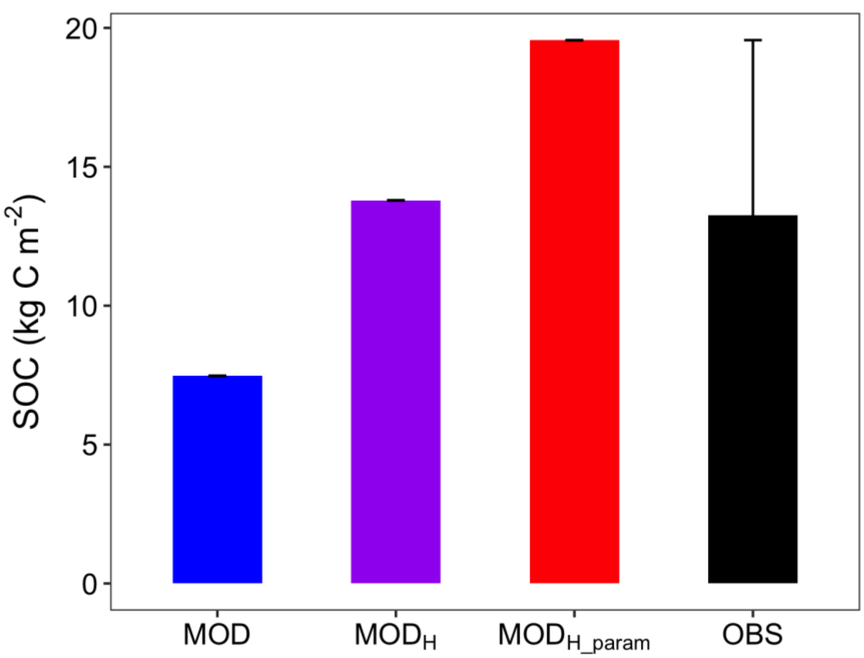

Figure S3 Comparison of the observed and modeled soil organic carbon (SOC) stocks. OBS: observation; MOD: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the calibrated Hansom model; MOD_{H_param}: model output after soil water potential improvement using the calibrated Hansom model and parameter adjustments.

In the revision, we added the new results (page 10, line 6 – 10):

"Results showed that the ELMv0 with both the default and improved SWP by the Hanson model overestimated the maximum LAI (Fig. 4a).The adjustment of the aforementioned five parameters (Table 2) significantly reduced the LAI to within a more reasonable range (Fig. 4a). The parameter changes further increased the simulated GPP and SR during the peak growing season, in addition to the improvement by the adjusted SWP (Fig. 4b, c). However, all modifications of the ELMv0 still overestimated SR during the non-peak growing season, resulting in significant overestimation of annual SR fluxes (Fig. S5a). After the parameter adjustments, the annual GPP flux was still within the observed range (Fig. S5b)."

We also discussed the potential reasons in Section 4.2:

"Although the simulation of the SWP using the Hanson model improved the representation of both annual SR and GPP, the model continued to overestimate SR during the non-peak growing season (Fig. 4), resulting in significant overestimation of the annual SR fluxes (Fig. S5). In addition, no matter which SWP simulations were used, the ELMv0 had smaller interannual variability than the observations (Fig. 2). Specifically, the model was not able to capture the steep decreases in GPP and SR in the extreme drought year (i.e., 2012). These results indicate that the current model structure is not sensitive enough to environmental changes. A few potential reasons may contribute to the underestimated seasonal and interannual variability. In the ELMv0, heterotrophic respiration contributed a majority proportion (i.e., over 85%) of the total SR during non-growing seasons (Fig. 5), suggesting that the overestimation of SR was primarily due to the biased simulation of heterotrophic respiration. A potential reason for the biased simulation of heterotrophic respiration may be related to the temperature sensitivity (Q_{10}). Theoretically, a higher Q_{10} can result in greater seasonal variability of SR (Fig. S9). Compared to relatively small Q_{10} values, a larger Q_{10} can lead to lower heterotrophic respiration when temperature is below the reference temperature, and greater heterotrophic respiration when temperature is above the reference (Fig. S9). In the ELMv0, the reference temperature is 25 ºC and the Q_{10} of heterotrophic respiration is 1.5 (Oleson et al., 2013). A previous study derived a much greater Q_{10} value (i.e., 2.8) when the parameters were calibrated with data from another temperate forest (Mao et al., 2016). We hypothesized that the Q_{10} value of 1.5 may be too small for the MOFLUX site. We arbitrarily increased Q_{10} from 1.5 to 2.5, but there were minimal effects on the SR simulation (Fig. S10). This indicates that modifying the temperature sensitivity of heterotrophic respiration may not improve the modeled representation of seasonality of SR in the ELMv0.

Another potential reason for the biased simulation of heterotrophic respiration may be that the seasonality of microbial organisms was not adequately represented in the model. Like most ESMs, the ELMv0 represents soil C dynamics using linear differential equations and assumes that SR is a substrate-limited process in the model. However, producers of CO_2 in soils, microbial organisms, have a significant seasonal cycle (Lennon and Jones, 2011). These organisms usually have very high biomass and activity during the peak growing season, with favorable conditions of temperature, moisture and substrate supply, and tend to be dormant under stressful conditions in non-peak growing seasons (Lennon and Jones, 2011; Stolpovsky et al., 2011; Wang et al., 2014; Wang et al., 2015). The seasonality of microbial biomass and activity, in addition to that of GPP and ST, may contribute to the seasonal variability of SR.

Additionally, another reason may be related to the model lacking representation of macroinvertebrate and other forest floor and soil fauna. There is a high density of earthworms at the MOFLUX site (Wenk et al., 2016). Earthworms can shred and redistribute soil C and change soil aggregation structure, which may alter soil C dynamics and CO_2 efflux to the atmosphere (Verhoef and Brussaard, 1990; Brussaard et al., 2007; Coleman, 2008). Like microbial organisms, earthworms usually have a significant seasonal cycle, showing high biomass and high activity during peak growing seasons and tending to be dormant during non-growing seasons (Wenk et al., 2016). However, a recent review suggests that current experimental evidence and conceptual understanding remains insufficient to support the development of explicit representation of fauna in ESMs (Grandy et al., 2016). Therefore, data collection focused on seasonal variations in fauna and microbial biomass and activity might enable further improvements in the representation of seasonal variation in SR.

Our analyses also showed that the modeled SR was not able to reach the observed peak in many years during the peak growing season even when the modeled GPP exceeded the observation (Fig. S11). In addition, the parameter modification increased GPP during both peak and non-peak growing seasons, resulting in an even greater overestimation of SR during non-peak growing seasons (Fig. S11). These results suggest that simply increasing GPP may not be adequate to increase the seasonal variability of the simulated SR. A potential reason may be that the current model does not include root exudates. Root exudates are labile C substrates that are important for SR (Kelting et al., 1998; Kuzyakov, 2002; Sun et al., 2017). The root exudate rate is primarily dependent on root growth, showing a seasonal cycle in temperate forests (Kelting et al., 1998; Kuzyakov, 2002). Thus, including root exudates in the model may further increase the model simulated SR during the peak growing season without needing to increase GPP."

Page 12, line 4. This doesn't seem like a fair statement or comparison, as results from the tuned Clapp & Hornberger scheme are never presented.

Response: In the revision, we compared the ELMv0 simulations with the Hanson model and the calibrated Clapp & Hornberger scheme as described above. The performance of the Hanson model was better than the calibrated Clapp & Hornberger model. Thus, we used the Hanson model to improve the SWP simulations.

Page 12, line 10, given the dominance of Rh in contributions to soil respiration (Fig 7). I'd suspect that changes in SR have more to do with larger SOM stocks than they do links between substrate supply through GPP, as suggested here, but no data are presented along these lines?

Response: As described above, the improved SR was a joint result of changes in GPP, SOC stocks and the moisture modifier of heterotrophic respiration. We revised this part (page 11, line 1 – 11) as:

"The improvements in model fits were due to changes in GPP, SOC stocks, and the moisture modifier on heterotrophic respiration (Figs. S1 – S3). First, the default ELMv0 underestimated GPP (Fig. 2b), as in a recent study where CLM4.5 significantly underestimated GPP at a coniferous forest in northeastern United States (Duarte et al., 2017). GPP affects the substrate supply for SR, as evidenced by the close relationship between changes in SR and GPP (Fig. S1),

which is consistent with experimental evidence showing GPP can directly affect the magnitude of root respiration (Craine et al., 1999; Högberg et al., 2001; Wan and Luo, 2003; Verburg et al., 2004; Gu et al., 2008). Second, the changed soil moisture scheme increased the moisture modifier (ξ_W) on heterotrophic respiration during the peak-growing season, and decreased it during the non-peak growing season (Fig. S2), which is consistent with the trend of changes in SWP (Fig. 3). In addition, the changed soil moisture scheme also increased the simulated SOC stock, the substrate for heterotrophic respiration (Fig. S3). These changes together resulted in the improvement of simulated SR."

Page 12, line 20. This statement may be true, but it's not clear that changes to VMC proposed here had much of an effect on the Rh component of the model. To show this, it seems like showing the soil moisture effect (w_scalar) on soil decomposition rates from different configurations of the model would be needed. Otherwise, I'd suspect that improvement to SR (Fig 2, 5) are predominantly driven by larger soil C stocks (via higher GPP), but not from direct improvement in the SWC on soil biogeochemistry, as suggested in the Powell paper referenced.

Response: The soil moisture modifier (w_scalar, which is ξ_W in Eq. 9) of heterotrophic respiration was shown in Fig. S2 (and also see below) in the revision. In addition, the SOC stocks under different schemes were shown in Fig. S3. As discussed above, the improvement of SR was a joint result of changes in GPP, SOC stocks and the moisture modifier.

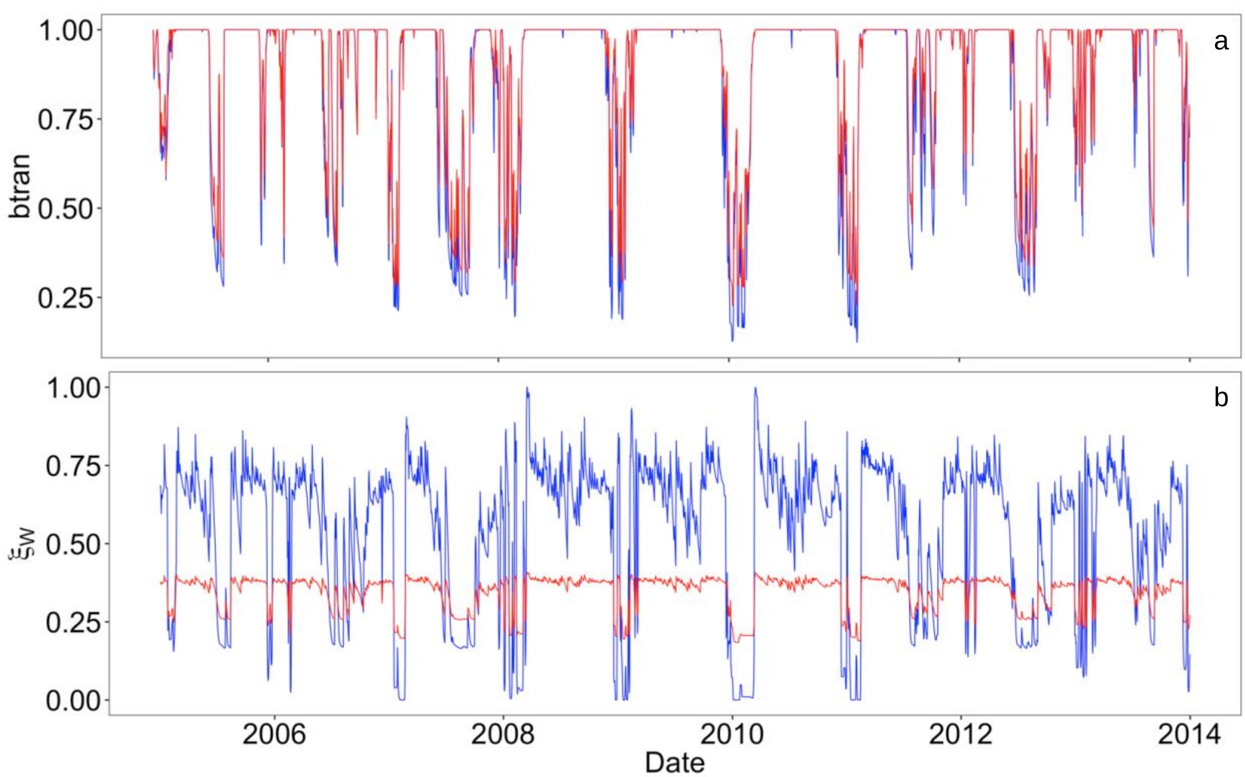

Figure S2 Impact of the changed SWP on the moisture modifiers of GPP (btran, a) and heterotrophic respiration (ξ_W, b). $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model.

Page 12, line 20 what is SWP-VMC, should be VWC?

Response: Yes and revised.

Page 13, line 2. Again this claim is poorly supported by the data presented. Yes, the tuned Clapp and Hornberger model is not the 'best' model in Table 1, but are results for GPP or SR markedly different than the Hansen results shown?

Response: Yes, the results for GPP and SR were markedly different when using the Hanson model and the calibrated Clapp & Hornberger model as discussed above. The ELMv0 with the Hanson model consistently produced higher GPP and SR than that with the calibrated Clapp & Hornberger model. In comparison with the observations, the modeled SR generally fell within the 1 sigma (i.e., standard deviation) range of observations, by using both the Hanson model and the calibrated Clapp & Hornberger model. However, the modeled GPP with the calibrated Clapp & Hornberger model was lower than the observations.

In response to the reviewer's concern, we deleted the sentence in the revision.

Page 14. The q10 analysis is nice, but I wonder if a more ecological explanation is relevant here- specifically highlighting the role of root exudates in supplying labile C substrates that are important for SR? The land model here doesn't consider these ecologically important C fluxes that likely have an important control over the seasonal dynamics of soil respiration and microbial biomass already discussed?

Response: The reviewer provided a great potential reasons for the modeled biases of SR. In the revision, we added the discussion of root exudates as:

"Our analyses also showed that the modeled SR was not able to reach the observed peak in many years during the peak growing season even when the modeled GPP exceeded the observation (Fig. S11). In addition, the parameter modification increased GPP during both peak and non-peak growing seasons, resulting in an even greater overestimation of SR during non-peak growing seasons (Fig. S11). These results suggest that simply increasing GPP may not be adequate to increase the seasonal variability of the simulated SR. A potential reason may be that the current model does not include root exudates. Root exudates are labile C substrates that are important for SR (Kelting et al., 1998; Kuzyakov, 2002; Sun et al., 2017). The root exudate rate is primarily dependent on root growth, showing a seasonal cycle in temperate forests (Kelting et al., 1998; Kuzyakov, 2002). Thus, including root exudates in the model may further increase the model simulated SR during the peak growing season without needing to increase GPP."

References: Bonan, et al (2013) Evaluating litter decomposition in earth system models with long-term litterbag experiments: an example using the Community Land Model version 4 (CLM4), Global Change Biology, 19, 957-974.
Brunke et al. 2016. "Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5)." Journal of Climate 29(9): 3441-3461, doi:10.1175/JCLI-D-15-0307.1.

Riley, W. 2018. "Impacts of Microtopographic Snow Redistribution and Lateral Sub- surface Processes on Hydrologic and Thermal States in an Arctic Polygonal Ground Ecosystem: A Case Study Using ELM-3D v1.0." Geoscientific Model Development 11(1): 61-76, doi:10.5194/gmd-11-61-2018.

Response: We appreciate the information!

Reviewer 2

This paper reports an effort of tuning an Earth system model, E3SM, to fit observed leaf area index (LAI), gross primary production (GPP, derived from eddy flux data), and soil respiration at a temperate deciduous forest site. The authors specifically tested different empirical relationships between volumetric water content (VWC) and soil water potential (SWP), and found tuning soil water potential improve the simulation of soil respiration. So, they concluded that "modelling soil respiration can be significantly improved by better model representations of the soil water retention curve." I agree with the authors that the well data-constrained model, Hanson model, increased the prediction of soil water potential, and may improve the simulation of GPP, which have been shown by the results (Figs. 3 and 7). But for the improvement of soil respiration, I think it's just a coincidence. From the Fig. 5a (page 9), we can see the new VMC-SWP relationship (i.e., Hanson model) increases soil respiration rate overall, but it does NOT change the pattern. This means the performance of soil respiration modeling is not improved. The authors also pointed out that the original model underestimates GPP and soil respiration (Line 13, page 7, and Fig. 2). So, the improvement of soil respiration prediction was not due to the improvement of SWP simulation, but because increases in GPP. The increases in GPP may increase carbon allocation to roots or total soil carbon, and therefore increase soil respiration. And, according to Fig. 7, the most possible reason for underestimating soil respiration is that the root respiration is not high enough in growing season, which also leads to the seasonal pattern that does not fit the observations because root respiration is usually high in growing season and very low in non-growing season.

Response: We greatly appreciate the valuable comment. The reviewer criticized that "*the improvement of soil respiration prediction was not due to the improvement of SWP simulation, but because increases in GPP*". A similar comment was raised by the first reviewer who commented that "*it's not clear if this is a direct effect of soil moisture on soil respiration*". We appreciate both the reviewers pointed the issue out. In response to the comments, we analyzed the changes in the soil moisture modifiers of GPP (btran) and heterotrophic respiration (ξ_W) and soil organic carbon (SOC) (Fig. S2-S3 and also see below), as suggested by the first reviewer. The improved soil water scheme using the Hanson model increased both *btran* and ξ_W during the peak growing season, and reduced ξ_W during the non-peak growing season (Fig. S2). While the model-simulated SOC with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations increased SOC stocks (Fig. S3). These results, combining with previous results (revised Fig. S1), indicate that the improved annual fluxes of SR by SWP was a joint result of changes in GPP, SOC stocks and the moisture modifier on heterotrophic respiration.

In the revised manuscript, we added two new supplementary figures (Fig. S2, S3), moved Fig. 4 as Fig. S1, presented the results (page 8, line 1 – 5) and discussed details (page 11, line 1 – 11) in the text.

"The changes in annual SR and GPP (i.e., the differences between before and after the improved SWP simulation using the Hanson model) showed a linear relationship (Fig. S1). In addition, the improved soil water scheme using the Hanson model increased both the moisture modifiers of GPP and heterotrophic respiration (i.e., btran and ξ_W*) during the peak growing season, and reduced* ξ_W *during the non-peak growing season (Fig. S2). While SOC when simulated by the model with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations using the Hanson model increased SOC stocks (Fig. S3)."*

"Constraining the SWP-VWC relationship with site-specific data and using the Hanson model instead of the ELMv0 default model (Fig. 1) significantly improved the model representation of SWP (Fig. 3) and annual SR (Fig. 2a). The improvements in model fits were due to changes in GPP, SOC stocks, and the moisture modifier on heterotrophic respiration (Figs. S1 – S3). First, the default ELMv0 underestimated GPP (Fig. 2b), as in a recent study where CLM4.5 significantly underestimated GPP at a coniferous forest in northeastern United States (Duarte et al., 2017). GPP affects the substrate supply for SR, as evidenced by the close relationship between changes in SR and GPP (Fig. S1), which is consistent with experimental evidence showing GPP can directly affect the magnitude of root respiration (Craine et al., 1999; Högberg et al., 2001; Wan and Luo, 2003; Verburg et al., 2004; Gu et al., 2008). Second, the changed soil moisture scheme increased the moisture modifier (ξ_W*) on heterotrophic respiration during the peak-growing season, and decreased it during the non-peak growing season (Fig. S2), which is consistent with the trend of changes in SWP (Fig. 3). In addition, the changed soil moisture scheme also increased the simulated SOC stock, the substrate for heterotrophic respiration (Fig. S3). These changes together resulted in the improvement of simulated SR."*

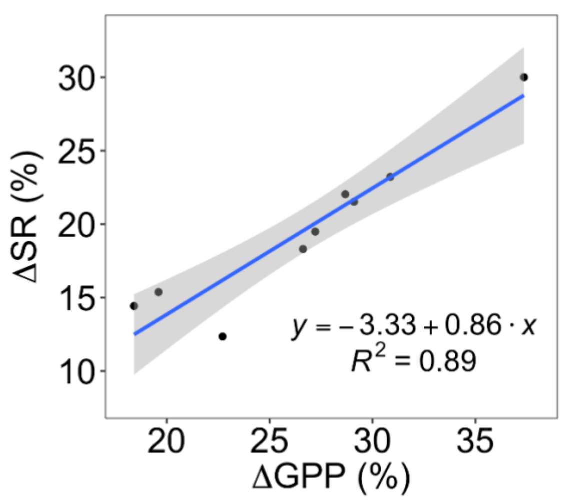

Figure S1: Relationship between changes in simulated annual soil respiration (ΔSR) and gross primary production (ΔGPP) induced by improvement of soil water potential using the Hanson model.

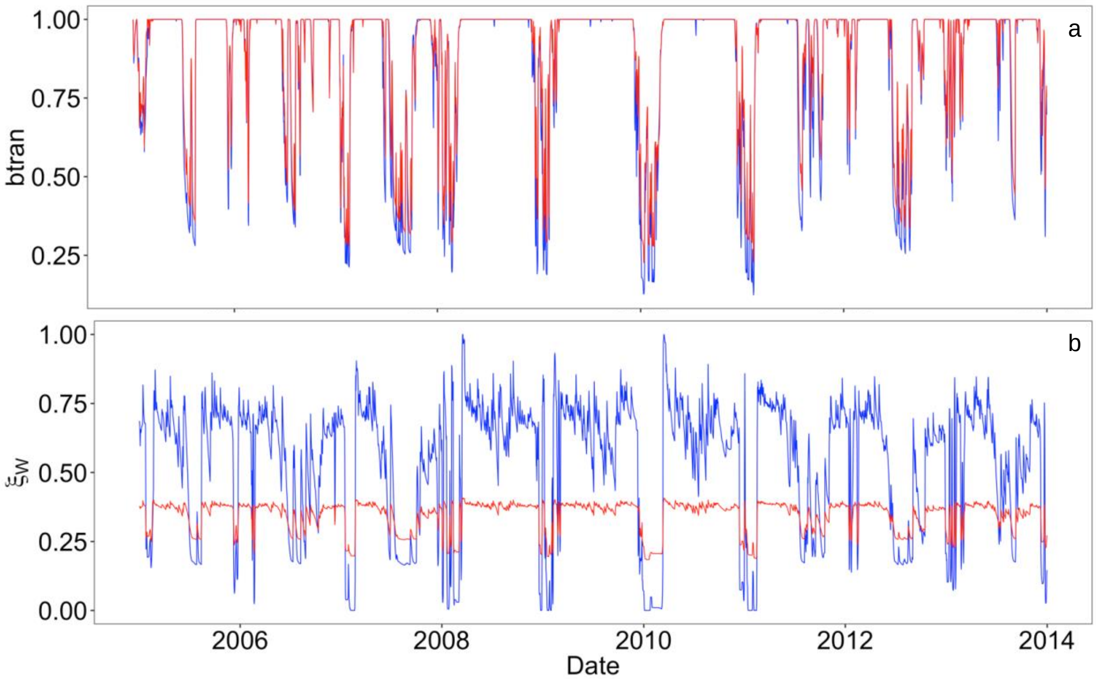

Figure S2 Impact of the changed SWP on the moisture modifiers of GPP (btran, a) and heterotrophic respiration (ξ_W, b). $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model.

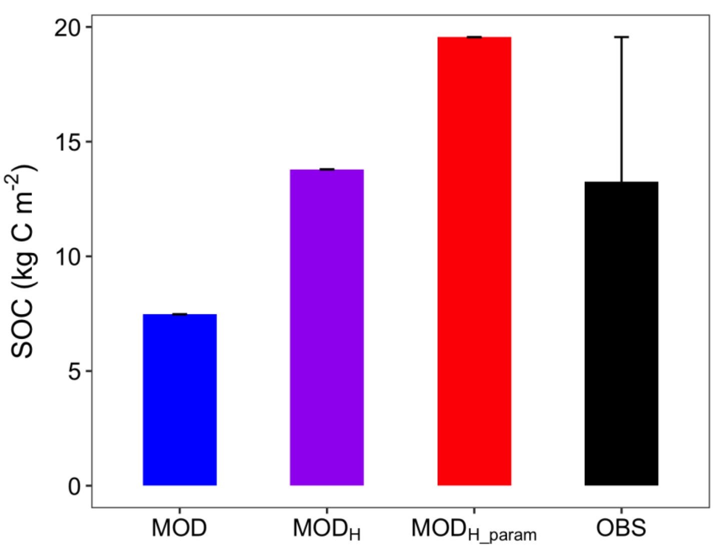

Figure S3 Comparison of the observed and modeled soil organic carbon (SOC) stocks. OBS: observation; MOD: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model; MOD_{H_param}: model output after soil water potential improvement using the Hanson model and parameter adjustments.

A detailed report on the tuning of an ESM is valuable even if no new mechanisms were added. It helps to understand model performance and the thoughts behind the model development. For improving simulation of soil respiration, the authors had looked at the sensitivity to temperature, LAI, GPP, and relative contributions of roots and soil carbon, and tuned a bunch of parameters (Table 2 in page 5). A detailed analysis of the successes and fails of these tunings would be interesting. For example, I'd like to see how the improvement of SWP prediction affects plant physiology, photosynthesis, allocation, NPP (because NPP=Rh at equilibrium). These variables may change soil respiration.

Response: The reviewer provided an insightful suggestion. In the revised manuscript, we added the analyses of the effects of the improved SWP on photosynthesis, NPP, and carbon allocation to fine root, leaf and woody tissue (Fig. S7 and also see below). Results showed that the improved SWP generally increased all photosynthesis, NPP and carbon allocations to different tissues during the growing season. In addition, parameter adjustments further increased them.

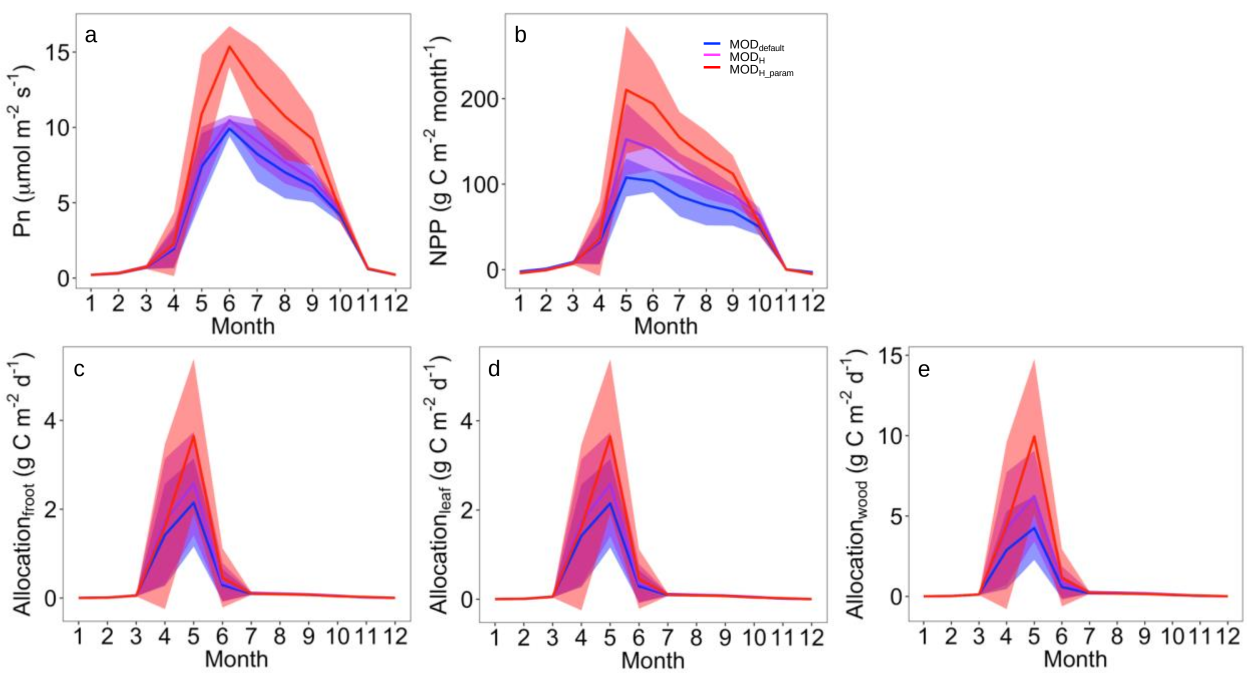

Figure S7 The annual mean cycles of photosynthesis (Pn), net primary production (NPP) and C allocations to fine root (Allocation$_{froot}$), leaf (Allocation$_{leaf}$) and woody tissue (Allocation$_{wood}$). MOD$_{default}$: model output before soil water potential improvement; MOD$_{H}$: model output after soil water potential improvement using the Hanson model; MOD$_{H_param}$: model output after soil water potential improvement using the Hanson model and parameter adjustments.

We added the description of these results in the revised manuscript as (page 10, line 11 – 16):

"In addition, we analyzed changes in simulated evapotranspiration (ET), runoff, photosynthesis, net primary production, and C allocation to fine roots, leaf and wood in response to the changes

in the soil water scheme and parameters (Fig. S6, S7). The change in soil moisture scheme and parameter adjustments slightly increased ET and decreased runoff. Despite these slight changes, the model simulated ET generally fell within the observed range, with or without changes in soil water scheme and parameters (Fig. S6). The improved SWP and parameter adjustments generally increased all photosynthesis, NPP and carbon allocations to different tissues during the growing season (Fig. S7)."

Specifically, for water effects on soil heterotrophic respiration, the model uses two equations to link volumetric water content to heterotrophic respiration: VMC–>SWP and SWPRh. The second equation (SWPRh, Eq 9 in page 4) is much more critical than the first one for modeling heterotrophic respiration. It represents the knowledge of how soil moisture affects microbial physiology. It needs to be explored in detail if the goal of this research is to improve the simulation of soil respiration.

Response: We agree with the reviewer that the moisture modifier (ξ_W) is a critical factor in the model in determining how soil moisture affects microbial physiology. In the revised manuscript, we analyzed the effect of the improved SWP on ξ_W (Fig. S2b). The improved SWP increased ξ_W during the peak growing season, and reduced ξ_W during the non-peak growing season (Fig. S2), which was consistent with the changes in SWP (Fig. 3b). The changes in ξ_W, as well as GPP and SOC stocks, jointly determined the effect of the improved SWP on SR, as discussed above and in the revised manuscript.

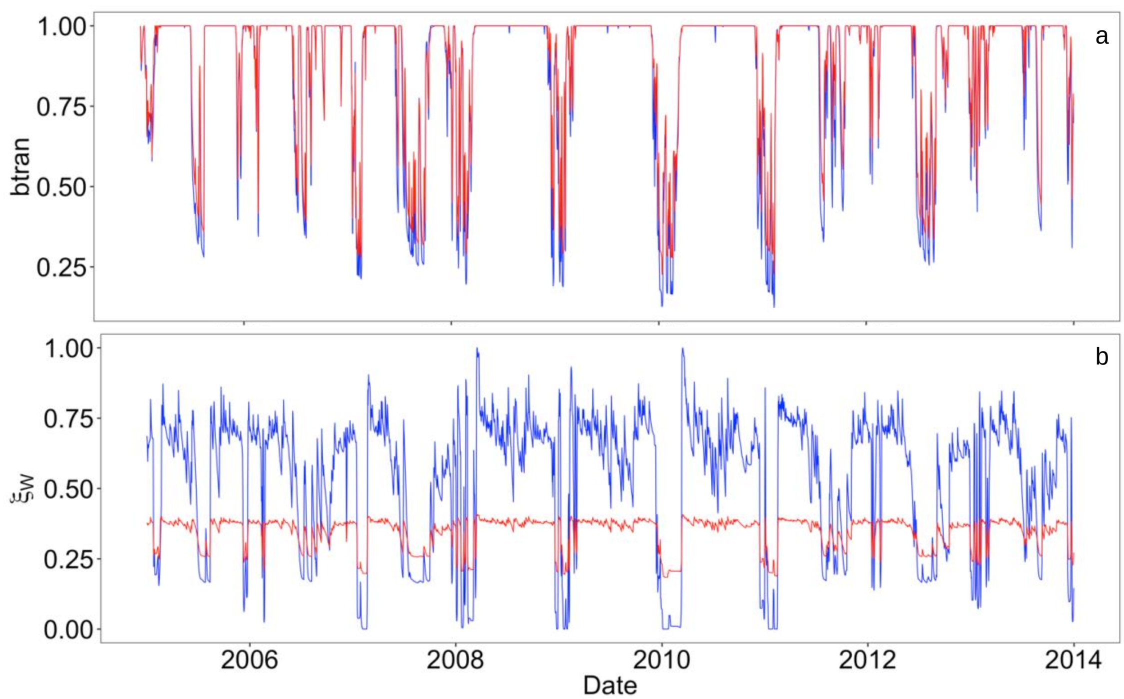

Figure S2 Impact of the changed SWP on the moisture modifiers of GPP (btran, a) and heterotrophic respiration (ξ_W, b). $MOD_{default}$: model output before soil water potential improvement; MOD_H: model output after soil water potential improvement using the Hanson model.

The related revisions are shown below and in the text (page 8, line 1 – 5; page 11, line 1 – 11).

"The changes in annual SR and GPP (i.e., the differences between before and after the improved SWP simulation using the Hanson model) showed a linear relationship (Fig. S1). In addition, the improved soil water scheme using the Hanson model increased both the moisture modifiers of GPP and heterotrophic respiration (i.e., btran and ξ_W*) during the peak growing season, and reduced* ξ_W *during the non-peak growing season (Fig. S2). While SOC when simulated by the model with different soil water schemes generally fell within the wide range of observations, the improved SWP simulations using the Hanson model increased SOC stocks (Fig. S3)."*

"Constraining the SWP-VWC relationship with site-specific data and using the Hanson model instead of the ELMv0 default model (Fig. 1) significantly improved the model representation of SWP (Fig. 3) and annual SR (Fig. 2a). The improvements in model fits were due to changes in GPP, SOC stocks, and the moisture modifier on heterotrophic respiration (Figs. S1 – S3). First, the default ELMv0 underestimated GPP (Fig. 2b), as in a recent study where CLM4.5 significantly underestimated GPP at a coniferous forest in northeastern United States (Duarte et al., 2017). GPP affects the substrate supply for SR, as evidenced by the close relationship between changes in SR and GPP (Fig. S1), which is consistent with experimental evidence showing GPP can directly affect the magnitude of root respiration (Craine et al., 1999; Högberg et al., 2001; Wan and Luo, 2003; Verburg et al., 2004; Gu et al., 2008). Second, the changed soil moisture scheme increased the moisture modifier (ξ_W*) on heterotrophic respiration during the peak-growing season, and decreased it during the non-peak growing season (Fig. S2), which is consistent with the trend of changes in SWP (Fig. 3). In addition, the changed soil moisture scheme also increased the simulated SOC stock, the substrate for heterotrophic respiration (Fig. S3). These changes together resulted in the improvement of simulated SR."*

Short comment 1:

Dear authors,
in my role as Executive editor of GMD, I would like to bring to your attention our Editorial version 1.1:

http://www.geosci-model-dev.net/8/3487/2015/gmd-8-3487-2015.html

This highlights some requirements of papers published in GMD, which is also available on the GMD website in the 'Manuscript Types' section:

http://www.geoscientific-model-development.net/submission/manuscript_types.html

In particular, please note that for your paper, the following requirement has not been met in the Discussions paper:

• "The main paper must give the model name and version number (or other unique identifier) in the title."

Even if this is not a strict requirement for evaluation paper, we like to encourage authors to provide also the version number of the evaluated model, as usually evaluation results depend on model version.

Additionally, please not, that GMD is encouraging authors to provide a persistent access to the exact version of the source code used for the model version presented in the paper. As explained in https://www.geoscientific-modeldevelopment.net/about/manuscript_types.html the preferred reference to this release is through the use of a DOI which then can be cited in the paper. For projects in GitHub (such as thee E3SM Land Model) a DOI for a released code version can easily be created using Zenodo, see https://guides.github.com/activities/citable-code/ for details.

Yours,
Astrid Kerkweg

Response: We appreciate the comments. We revised the title to include the model name and version number. The revised title is *"Evaluating the E3SM Land Model version 0 (ELMv0) at a temperate forest site using flux and soil water measurements"*. We also provided the link to the model.

References:

Bisht, G., Riley, W. J., Wainwright, H. M., Dafflon, B., Yuan, F. M., and Romanovsky, V. E.: Impacts of microtopographic snow redistribution and lateral subsurface processes on hydrologic and thermal states in an Arctic polygonal ground ecosystem: a case study using ELM-3D v1.0, Geosci Model Dev, 11, 61-76, 2018.

Brunke, M. A., Broxton, P., Pelletier, J., Gochis, D., Hazenberg, P., Lawrence, D. M., Leung, L. R., Niu, G. Y., Troch, P. A., and Zeng, X. B.: Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5), Journal of Climate, 29, 3441-3461, 2016.

Brussaard, L., de Ruiter, P. C., and Brown, G. G.: Soil biodiversity for agricultural sustainability, Agr Ecosyst Environ, 121, 233-244, 2007.

Coleman, D. C.: From peds to paradoxes: Linkages between soil biota and their influences on ecological processes, Soil Biol Biochem, 40, 271-289, 2008.

Craine, J. M., Wedin, D. A., and Chapin, F. S.: Predominance of ecophysiological controls on soil CO_2 flux in a Minnesota grassland, Plant and Soil, 207, 77-86, 1999.

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Hickler, T., Jain, A. K., Luo, Y. Q., Parton, W. J., Prentice, I. C., Smith, B., Thornton, P. E., Wang, S. S., Wang, Y. P., Warlind, D., Weng, E. S., Crous, K. Y., Ellsworth, D. S., Hanson, P. J., Seok Kim, H., Warren, J. M., Oren, R., and Norby, R. J.: Forest water use and water use efficiency at elevated CO_2: a model-data intercomparison at two contrasting temperate forest FACE sites, Global Change Biology, 19, 1759-1779, 2013.

Duarte, H. F., Raczka, B. M., Ricciuto, D. M., Lin, J. C., Koven, C. D., Thornton, P. E., Bowling, D. R., Lai, C. T., Bible, K. J., and Ehleringer, J. R.: Evaluating the Community Land Model (CLM4.5) at a coniferous forest site in northwestern United States using flux and carbon-isotope measurements, Biogeosciences, 14, 4315-4340, 2017.
Grandy, A. S., Wieder, W. R., Wickings, K., and Kyker-Snowman, E.: Beyond microbes: Are fauna the next frontier in soil biogeochemical models?, Soil Biol Biochem, 102, 40-44, 2016.
Gu, L., Hanson, P. J., Mac Post, W., and Liu, Q.: A novel approach for identifying the true temperature sensitivity from soil respiration measurements, Global Biogeochem Cy, 22, 10.1029/2007GB003164, 2008.
Högberg, P., Nordgren, A., Buchmann, N., Taylor, A. F. S., Ekblad, A., Hogberg, M. N., Nyberg, G., Ottosson-Lofvenius, M., and Read, D. J.: Large-scale forest girdling shows that current photosynthesis drives soil respiration, Nature, 411, 789-792, 2001.
Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., and Wright, I. J.: TRY–a global database of plant traits, Global change biology, 17, 2905-2935, 2011.
Kelting, D. L., Burger, J. A., and Edwards, G. S.: Estimating root respiration, microbial respiration in the rhizosphere, and root-free soil respiration in forest soils, Soil Biol Biochem, 30, 961-968, 1998.
Kuzyakov, Y.: Separating microbial respiration of exudates from root respiration in non-sterile soils: a comparison of four methods, Soil Biol Biochem, 34, 1621-1631, 2002.
Lennon, J. T., and Jones, S. E.: Microbial seed banks: the ecological and evolutionary implications of dormancy, Nature reviews. Microbiology, 9, 119, 2011.
Lu, D., Ricciuto, D., Stoyanov, M., and Gu, L.: Calibration of the E3SM Land Model Using Surrogate Based Global Optimization, J Adv Model Earth Sy, 0, doi:10.1002/2017MS001134, 2018.
Mao, J., Ricciuto, D. M., Thornton, P. E., Warren, J. M., King, A. W., Shi, X., Iversen, C. M., and Norby, R. J.: Evaluating the Community Land Model in a pine stand with shading manipulations and $^{13}CO_2$ labeling, Biogeosciences, 13, 641-657, 2016.
Montané, F., Fox, A. M., Arellano, A. F., MacBean, N., Alexander, M. R., Dye, A., Bishop, D. A., Trouet, V., Babst, F., and Hessl, A. E.: Evaluating the effect of alternative carbon allocation schemes in a land surface model (CLM4. 5) on carbon fluxes, pools, and turnover in temperate forests, Geosci Model Dev, 10, 3499-3517, 2017.
Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S., Thornton, P. E., Bozbiyik, A., R., F., Heald, C. L., Kluzek, E., Lamarque, J. F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S. P., Ricciuto, D. M., Sacks, W. J., Sun, Y., Tang, J., and Yang, Z. L.: Technical Description of Version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-503+STR, Boulder, Colorado, 2013.
Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G. S., Chen, J. M., Ciais, P., Davis, K. J., Desai, A. R., Dietze, M. C., Dragoni, D., Garrity, S. R., Gough, C. M., Grant, R., Hollinger, D. Y., Margolis, H. A., McCaughey, H., Migliavacca, M., Monson, R. K., Munger, J. W., Poulter, B., Raczka, B. M., Ricciuto, D. M., Sahoo, A. K., Schaefer, K., Tian, H. Q., Vargas, R., Verbeeck, H., Xiao, J. F., and Xue, Y. K.: Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis, Global Change Biology, 18, 566-584, 2012.

Stolpovsky, K., Martinez-Lavanchy, P., Heipieper, H. J., Van Cappellen, P., and Thullner, M.: Incorporating dormancy in dynamic microbial community models, Ecol Model, 222, 3092-3102, 2011.
Sun, L. J., Ataka, M., Kominami, Y., and Yoshimura, K.: Relationship between fine-root exudation and respiration of two Quercus species in a Japanese temperate forest, Tree Physiology, 37, 1011-1020, 2017.
Verburg, P. S. J., Arnone, J. A., Obrist, D., Schorran, D. E., Evans, R. D., Leroux-Swarthout, D., Johnson, D. W., Luo, Y. Q., and Coleman, J. S.: Net ecosystem carbon exchange in two experimental grassland ecosystems, Global Change Biology, 10, 498-508, 2004.
Verhoef, H. A., and Brussaard, L.: Decomposition and Nitrogen Mineralization in Natural and Agroecosystems - the Contribution of Soil Animals, Biogeochemistry, 11, 175-211, 1990.
Walker, A. P., Hanson, P. J., De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Asao, S., Dietze, M., Hickler, T., Huntingford, C., Iversen, C. M., Jain, A., Lomas, M., Luo, Y. Q., McCarthy, H., Parton, W. J., Prentice, I. C., Thornton, P. E., Wang, S. S., Wang, Y. P., Warlind, D., Weng, E. S., Warren, J. M., Woodward, F. I., Oren, R., and Norby, R. J.: Comprehensive ecosystem model-data synthesis using multiple data sets at two temperate forest free-air CO2 enrichment experiments: Model performance at ambient CO2 concentration, J Geophys Res-Biogeo, 119, 937-964, 2014.
Wan, S. Q., and Luo, Y. Q.: Substrate regulation of soil respiration in a tallgrass prairie: Results of a clipping and shading experiment, Global Biogeochem Cy, 17, 2003.
Wang, G., Mayes, M. A., Gu, L., and Schadt, C. W.: Representation of dormant and active microbial dynamics for ecosystem modeling, PloS one, 9, e89252, 2014.
Wang, G., Jagadamma, S., Mayes, M. A., Schadt, C. W., Steinweg, J. M., Gu, L., and Post, W. M.: Microbial dormancy improves development and experimental validation of ecosystem model, The ISME journal, 9, 226, 2015.
Wenk, E. S., Callaham, M. A., O'Brien, J. J., and Hanson, P. J.: Soil Macroinvertebrate Communities across a Productivity Gradient in Deciduous Forests of Eastern North America, Northeast Nat, 23, 25-44, 2016.
Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo, Y. Q., Wang, Y. P., El-Masri, B., Thornton, P., Jain, A., Wang, S. S., Warlind, D., Weng, E. S., Parton, W., Iversen, C. M., Gallet-Budynek, A., McCarthy, H., Finzi, A. C., Hanson, P. J., Prentice, I. C., Oren, R., and Norby, R. J.: Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO2 Enrichment studies, New Phytologist, 202, 803-822, 2014.