

## Reviewer #1: Paul Dirmeyer

General comments:

The authors present an offline model based assessment of connections between soil moisture, surface fluxes and LCL height with a single model in two drainage configurations and a suite of different atmospheric forcings.

I think the authors' conclusions about the dominant role of transpiration over surface evaporation for this model in this study are well demonstrated. I think the authors have a point about the differentiation between surface and root zone SM, and between surface evaporation versus transpiration processes. However, I think the notion of coupling is not adequately demonstrated here; it has not been demonstrated what is the cart and what is the horse. Too much has been presumed in this offline LSM study. Some things can be addressed diagnostically, as recommended below, but without a fully coupled study, and more realistic models, they must be very careful about making conclusions about coupling in nature.

First, the word "coupling" connotes cause and effect. In particular, "land-atmosphere coupling" suggests the return leg of the feedback loop where the land surface state influences the atmosphere. Like correlation, the "Kendall-tau" metric does not prove cause and effect but points out correspondences. This distinction needs to be made clearly in this paper. That phrasing is used a bit in the discussion, but needs to be the central tone of the paper.

It is possible (and has not been demonstrated otherwise here) that the correspondences between LCL and EF or LCL and SM are not cause-effect but effect-effect. Wet season humid conditions driven by moisture advection will lower the LCL without any land surface feedback. In a monsoon, the LCL is at its lowest level during the active phase rainy spells that correspond with adequate soil moisture (caused by the rain), which allow larger evaporation rates in an otherwise moisture-limited, energy-plentiful regime. The weaker correlations for CTRL (which does not drain well) in the wettest areas (sometimes even positive) keeps ET high (fig 2) and thus reduces day to day variability in EF; this could mean ET in CTRL is less responsive to precipitation, as opposed to the LCL being generally responsive to ET.

The possibility that the local water cycle is all atmospherically controlled needs to be eliminated before the existence of coupling can be declared here. Perhaps Kendall taus with daily precipitation and 2m humidity need to be examined as well. What it comes down to, which could be evaluated offline, is whether, for CLM, the ET is controlled by SM or humidity? Since humidity deficit determines both LCL (absolutely) and latent heat flux (partially via both stomatal resistance for transpiration and the humidity C4385gradient term for direct evaporation), it should be that any loss in explained variance between the two, for which humidity is the main factor, should be taken up by the soil moisture availability.

These runs are not coupled; the LSM is only driven by specified meteorology. Thus, any diagnosis of coupling is predicated on assumptions about the processes that have not

already been adequately demonstrated for this place and time of year. Ultimately, in an uncoupled setting, estimates of such metrics must be based on a robust demonstrated process for coupling, which I think has not been demonstrated for monsoon regions in the wet season in general, and definitely not in this study.

Thus, one needs to be careful. This has always been a difficult problem, to establish the effect of land-atmosphere feedbacks in monsoon climates where there is such a strong background of large-scale forcing and circulation. Most such work has focused on India, secondarily on West Africa - that work is not cited here (some studies cited here do address that, e.g., by Ferguson, Taylor, etc., but those aspects are not discussed in this study).

Lastly, some of what the authors uncover are clearly model inadequacies in CLM (see specific comments below) - I would like to see these discussed more in Sec 5.

I would say if the authors would like to maintain the theme of a "coupling" evaluation, major revisions including more analysis are necessary to justify it. On the other hand, if the tone were changed to showing "correspondences" with the focus shifted squarely to the differing role of subsurface soil moisture and transpiration on the demonstrated relationships (which is the current emphasis in the conclusions), then the revisions are more editorial in nature and rather minor.

**Response:**

**We agree with Paul's comments and accept that the use of a statistical measure in offline experiments cannot conclusively prove cause and effect and therefore coupling. While statistical association between SM and LCL is indicative of coupling it doesn't demonstrate the relationship is cause and effect. To remedy this shortcoming we have refocused the manuscript on the association between SM, EF, and the LCL rather than coupling.**

Specific comments:

Throughout: The use of the term "observations" for reanalysis products, GLDAS and the flux estimates (and to a lesser extent the AMSR-E retrievals) is bothersome. "Ob-C4386

servationally based estimates" would be better. There are no direct observations in these data for surface fluxes, and even the state variables are measured sparsely in this region.

**Response:**

**We have replace the term “observations” to be either estimates or observationally derived estimates throughout the manuscript when we are referring to GLDAS, MERRA, or AMSR-E. The manuscript now reserves the term observations for use with the flux tower data which we have added to the paper.**

P10433 L17: How do you mean the word "decadal" here? Decadal usually means multi-year time scales; that is quite a jump from diurnal.

**Response:**

**We removed this sentence as it is out of place considering the context of the manuscript.**

Sec 2.2: Are there any stations or soundings to validate the meteorological data in this region? I would feel a lot better about it if so, especially if they are independent from the assimilation stream. Likewise, are there any flux tower or eddy-covariance measurements of latent heat flux that could be used to validate ET? How about soil moisture measurements to validate the variability and profiles of soil moisture, even at only one location?

**Response**

**We have added measurements from two flux tower data sites to the manuscript. The soil moisture and ET observations are compared directly against the simulations, the AMSR-E data, and the gridded ET products.**

**The two tower sites are also now used to compute Kt for EF-LCL, SM-LCL, and for SM<sub>r2</sub>-LCL for the site that has SM measurements at several depths.**

P10437 L23: MERRA has some well-documented hiccups in its time series when new remote sensing data come into the assimilation stream, especially affecting moisture variables (humidity and precipitation) at lower latitudes. It seems like this might have significant impacts over your study area - impacts that cannot be removed by removing linear trends. Have you examined this?

**Response:**

**The MERRA data was evaluated when the forcing dataset was originally created in 2012. The data were compared against the AWAP data (a gridded daily precipitation product from the Australian Bureau of Meteorology) and no obvious discontinuities were discovered.**

Sec 3.1: It is clear why afternoon LCL is used for the Kendall-tau calculations, but why is morning soil moisture so critical? Is the index really much different if you use afternoon values at the same time as maximum LCL?

**Response:**

**The morning time SM is used for several reasons. The first is that one cannot argue the LCL is controlling the SM when SM is sampled prior to the LCL. This doesn't eliminate the possibility that SM and LCL are controlled by an external factor (and thus both effects with the**

cause unexamined) but prevents direct control of SM by the LCL (through near surface humidity). Secondly, if SM is controlling the LCL one would expect SM to decrease during the day such that the mean afternoon SM will be smaller than the morning time SM. As the hypothesis is that high SM causes low LCL, “observing” SM in the afternoon may not be indicative of the relationship.

P10441 L22 and Fig 2: Normalized by what? Standard deviation? Since your correlation index Kendall-tau is non parametric, why use a normally distributed variance metric?

**Response:**

We have included an explanation that the SM in Figure 2 is normalized using the first two moments. The normalization follows previous work (Koster et al. 2006) that demonstrated how land model simulated SM varies substantially between models unless it is normalized. To highlight the effect of comparing SM with and without normalization we have included a new figure (Figures 4a and 4c) that directly compare SM from the simulations, AMSR-E, and observations at two tower sites.

Fig 2: What is the X axis? Presumably these labels are months, but there is no relationship to the calendar given.

**Response:**

We labeled the X axis to indicate that it is the date.

Fig 2b: How does this compare to the mean? Certainly there must be a lot of spatial variability. And what time of year is shown in Fig 2 - one season, both seasons? Finally, C4387 the color scale is not good - shades on both sides are not well differentiated from each other.

**Response:**

We have included the time period of the comparison in the Figure 2b caption as well as changed the color scale to highlight the spatial pattern of the differences. The differences in the mean SM between the simulations and AMSR-E seen in the new Figures 4a and 4c can be seen in Figure 2b near 12°S 131°E.

Discussion of Fig 3: Also point out DRY does better in the wet season, as CTL fluxes are too vigorous here. I can think of many possible causes; maybe there is too much infiltration in the wet season, the precip forcing could be too smooth in time, or CLM may be tuned to transpire too readily. Are there discharge data in this region to validate

runoff? What is the underlying geology? I imagine there is not much karst there, so standard LSM drainage parameterizations should be able to handle the vadose zone flow adequately.

**Response:**

**We have included the sentence “The overestimation of DJF ET compared to the gridded product is much more pronounced for the CTRL simulations (Figure 3a) than the DRY simulations (Figure 3c).” to highlight that DRY is closer to the gridded product in DJF than CTRL.**

**The new Figures 4c and 4d show that the wet season ET from DRY and CTRL is not consistently over estimated at either flux tower site and is underestimated by 30-40  $\text{Wm}^{-2}$  for the Adelaide River site.**

Sec 4.2: Just an aside comment: I would love to see someone actually measure soil moisture profiles here. Worldwide there are very few such measurements in monsoon regimes.

**Response:**

**We have now included SM observations from two flux tower sites. Unfortunately, while the Howard Springs site now measures SM at various depths, measurements only go back a few years and are not continuous.**

P10443 L14-15 "...indicating that the surface evaporation is the dominant ET mechanism." How do you reach this conclusion? Please elaborate. You explain the transpiration argument below but not the surface evaporation argument here.

**Response:**

**A previous version of the manuscript included a figure that showed the fraction of ET from transpiration for DRY and CTRL from SON and DJF. Both DRY and CTRL show ET comprised of 10-35% transpiration during DJF with the result being soil or canopy evaporation.**

**We added the following to explain this point “The mean DJF ET is similar between CTRL and DRY, with differences between the two only 10-20  $\text{Wm}^{-2}$ , corresponding to roughly 10-20% of the mean value. The fractional contribution of transpiration to the total ET during DJF is roughly 10-30% for both DRY and CTRL (figure not shown) indicating that the surface evaporation is the dominant ET mechanism. “**

P10443 L16-17: Two "however"s in a row.

**Response:**

**Removed the first however.**

P10446 L1: 90% of water uptake capacity in the top 1m is almost certainly unrealistic, especially in a wet/dry regime where the woody species must have deep roots to survive the dry season. They will tap the shallow moisture during the wet season when it is easy. Such dynamic root responses are not part of CLM or most other LSMs, and are a shortcoming for simulating transpiration in semi-arid and seasonally arid biomes like this.

**Response:**

**We agree that the actual depth of water removal in the simulated system is not realistic, as some Eucalypt species have been shown to have rooting depths in the tens of meters. We have further acknowledged this fact and added the following to the manuscript. “Applying Equation (4) using  $SM_{rz}$  imposes a different set of problems, as the rooting depth is model dependent and generally only approximately known. There is substantial evidence that eucalypts have rooting depths exceeding 20 meters, however neither CLM4 or the direct observations in this study extend that deep. Due to these limitations,  $SM_{rz}$  is computed as the weighted mean of the SM observations at 10, 40, and 100cm for the Howard Springs site. We assume that the  $SM_{rz}$  consists of the soil layers between the surface and a depth of 1m, as greater than 90% of the prescribed roots in CLM4 are within 1m of the surface (Oleson et al. 2010). This assumed rooting depth is consistent with the model formulation but not realistic given the rooting depths of eucalypts.”**

P10446 L19-26: Fig 8 seems tacked on; the figure is merely described but the consequences are not explained.

**Response:**

**We have added discussion of the consequences of the previous Figure 8 (now Figure 9). We now explicitly state that the use of multiple estimates of the LCL may cause  $K_t$  to be uncertain so the uncertainty is examined.**

Sec 5: I would say the unrealistically wet SM profile in dry season in CTRL makes up for the overly shallow rooting profile in CLM for this biome; the right answer is reached C4388

for the wrong reasons. By removing one of the two compensating errors (in the DRY case) the results deteriorate.

**Response:**

**We agree that the rooting depth within the model is much more shallow than what the rooting depth physically is. To acknowledge this point we have added “Within the model, the soil column-groundwater interactions parameterized in CTRL inhibit the large, ET limiting  $SM_{rz}$  reduction present in DRY. In reality the inability of DRY to maintain ET**

**during SON may result from the shallow rooting depths assumed in CLM4. The depths are substantially more shallow than the rooting depths of eucalypts. Realistic rooting depth profiles (reaching 20 meters) and corresponding soil layer depths may negate the impact of the parameterized soil column-groundwater impacts current in CLM4.“**

P10447 L13-14: The similar Kendall-tau between EF-LCL despite model configuration, ET or SM is an indicator that the atmosphere (humidity) is in control, not the land surface state.

**Response:**

**The figure isn't shown however the SM-EF Kendall-tau is statistically significant in both DJF and SON for DRY and CTRL. Therefore the SM is always imparting significant control on ET in the modeling system. Similarly, the Kendall tau of SM<sub>rz</sub>-LCL demonstrates that SM is significantly associated with LCL, just not always SM<sub>1</sub>.**

P10447 L18: Here the term "coincidence" is used - this kind of neutral verbiage should be used throughout unless "coupling" can be more rigorously demonstrated.

P10448 L1-2: Very little area looks "positive" to me. This might have to do with the lack of magnitude dependence in the Kendall-tau, discussed by the authors. That is why indices like the terrestrial coupling index were developed (Guo et al. 2006, Dirmeyer 2011).

**Response:**

**To aid in the readers ability to follow the text we have included two locations as examples of both positive and negative association.**

**While several different terrestrial coupling indices exist, the purpose of this manuscript is to simply highlight that measures of association must be treated carefully as the results are dependent on how one specifies the SM term.**

P10448 L5-8: The changes from SON to DJF are still the same for SM<sub>rz</sub> as for SM<sub>1</sub>, just a bit weaker.

**Response:**

**We have explicitly added the location of a region where the results from SM<sub>rz</sub> and SM<sub>1</sub> differ.**

P10448 L12: The study of Jasechko et al. (2013) has been strongly refuted by several subsequent papers (e.g., Coenders-Gerrits et al. 2014, Sutanto et al. 2014, Wang-

Erlandsson et al. 2014) and they have subsequently backed off from their original claim (Schlesinger and Jasechko 2014). Also, Haverd et al. (2013) estimate half of Australia's ET is bare soil evaporation.

**Response:**

**We have removed the Jasechko et al. (2013) reference and now cite Coenders-Gerrits et al. 2014 and Schlesinger and Jasechko 2014. The manuscript was originally written prior to Jasechko et al. 2013 being so thoroughly refuted and is now updated to include the new papers.**

P10448 L16-17: I would say this study is likewise limited. Referring also to Eq 1, this study neglects that  $\Delta PBL$  can also occur due to large scale non-local influences, which strongly drive the  $E_{Fatm}$  term in monsoon regimes.

**Response:**

**We were trying to explicitly point out limitations of our method. We have altered the opening sentence of the paragraph to state that we are discussing our study as well as others that utilize similar methods.**

**We have added discussion regarding the model dependence of our results. We now discuss (Section 4.6) the physically unrealistic rooting depths in CLM4 and how this effects the results. We also discuss the model dependence of the impact of the groundwater on dry season ET (Section 5).**

**To bolster our argument we have added measurements from two flux tower sites within the study domain. We have included a new Figure 4 that shows the agreement between the simulated soil moisture and ET at the two sites. More importantly we have used the tower sites to calculate the EF-LCL and SM-LCL Kt and included the results in Figures 6, 7, and 8. The tower sites support the results we see with the model, in that EF-LCL are statistically associated during both DJF and SON, while the site disagree about the SM-LCL association during SON. The comparison with the simulated results isn't perfect due to the differing SM depths between the data and the model.**

**Reviewer #2:**

Received and published: 16 October 2014

(Page and line numbers in this review are from the printer-friendly version of the manuscript.)

In this study, the authors use offline simulations from a land surface model (CLM4) over Northern Australia to study land-atmosphere coupling during both the dry (SON) and wet (DJF) seasons – more specifically, they investigate:

- i) Whether including root-zone soil moisture (SM), versus surface layer SM, in the statistical metric they use to evaluate soil moisture-atmosphere coupling, matters for the diagnosed coupling;
- ii) Whether the mean background soil moisture content in the root-zone (varied between two simulations using different configurations of CLM4) matters for the diagnosed coupling.

The study addresses an important issue: the dependence of diagnosed SM-atmosphere coupling on season, background land surface state (here, mean SM) and SM depth definition. In particular, the latter matters as satellite-based SM retrievals of SM, and thus associated diagnoses of SM-atmosphere coupling, often only correspond to the top-surface layer. It is thus important to understand the impact of this limitation on the estimated coupling (in particular when satellite-derived diagnoses differ from other observation-based assessments). However, I have some significant concerns with the study as it stands, which I believe warrant quite major revisions. My concerns have to do with the methodology used, as well as the presentation and interpretation of results.

#### — Methodology

The authors drive an offline land model with different atmospheric datasets, and then essentially correlate the simulated soil moisture (at different depths), as well as evaporative fraction (EF), with an estimate of the lifting condensation level (LCL) derived from the atmospheric forcing. I could see this framework being used to evaluate observed SM-atmosphere coupling, using hourly model-simulated SM and EF as surrogates to observations (given that such observations are not widely available). This would assume, though, that the land model simulates “perfect” (given the forcing) soil moisture and land-atmosphere fluxes – this would certainly need to be discussed. But the authors go beyond that, and assess coupling in different land model configurations. I don’t see how this offline framework can be used to evaluate the sensitivity of the coupling to model configuration – here, the parameterization of groundwater and resulting mean deep soil moisture levels – since the atmosphere is always the same in every simulation and does not “see” the fluxes produced by the land model in different configurations. Given the difference in soil moisture between the two simulations (figure 4), I would expect the daily “sequences” of simulated surface fluxes (or deep soil moisture) to be different between both runs. In real life these different sequences would be associated with different atmospheric “sequences” (of LCLs), but here they are associated, by design, with the same atmosphere. I don’t see how land-atmosphere coupling can then be assessed in a relevant way. Unless the authors can explain otherwise, I fail to see how this experimental set up is suitable for investigating the question the authors want to address (i.e. the impact of model configuration on coupling).

#### **Response:**

**In response to these comments as well as the other reviewers we have reframed the**

**manuscript and discussion in terms of the association between SM,EF and LCL rather than the coupling between them. We acknowledge that statistical association cannot demonstrate the cause and effect relationship that coupling denotes. The manuscript makes no effort to discern whether coevolution of SM, EF, and LCL are cause and effect or simply effect-effect (from a different unexamined forcing). However statistically significant association found for SMrz-LCL but not for SM1-LCL indicates that future fully coupled experiments should account for SM beyond the first layer.**

Another methodological issue, potentially, is that, to evaluate the SM (or EF)- LCL coupling, the authors use a Kendall correlation coefficient, following Ferguson et al. (2012). I am not familiar with the latter study, but I have the following concern: it looks like the authors are correlating absolute values of SM (or EF) and LCL – not anomalies. I appreciate that Kendall correlation coefficients are better suited than, e.g., Pearson correlation coefficients, for non-linear relationships. However, I am concerned that, in a region with strong seasonality like the monsoon region of Northern Australia, correlations between absolute values are mostly going to capture the seasonally-forced co-evolution of the corresponding variables (i.e., SM, EF and atmosphere). This co-evolution happens without land surface feedbacks on the atmosphere. The strong correlations on figures 5-7, and the fact that there is overall very little difference between the CTRL and DRY simulations on figures 5-7 - despite, like I indicated above, probably different daily sequences of surface fluxes and SM- suggests to that this might be the case (i.e., seasonality dominating the signal). I strongly recommend the authors address and discuss this point.

**Response:**

**We agree that the seasonal cycle must be accounted for to prevent the seasonality from controlling the derived statistical association. The calculation of Kendall Tau (discussed in Section 3.1) is performed after detrending the data over each season. Detrending the data over each season removes the large seasonal cycle that would otherwise dominate the results. To make our procedure more clear to the reader we added the following sentence in Section 3.1 “The data are necessarily detrended separately over each season prior to deriving  $K\tau$  to prevent the strong seasonal cycle (Figures 2 and 3) from controlling the statistical relationship. “**

In short, I think the authors cannot really investigate question ii) (see first paragraph) in the present framework. I think they should either drop this part of the analysis and produce a more restricted paper on the difference between SM1-EF and SMrz-EF, or use fully coupled simulations instead.

— Presentation, significance and interpretation of results

Beyond this first order comment, I also take issue with how the authors are describing their results. A lot of the paper consists in qualitative comparisons maps of correlations in different seasons (SON and

DJF) and different model configurations (CTRL and DRY). The characterization of these differences is sometimes not consistent throughout the paper. For instance, p.10444 l.5 “DRY is generally more strongly coupled than CTRL during DJF” and p.10447 l.13 “the coupling between EF-LCL is similar in both model configurations”; also, p.10447 l.3 “The ET from CTRL and DRY are similar”, p.10447 l.14: “despite the mean ET. . . differing considerably between CTRL and DRY”. Sometimes differences are mentioned, but then deemed insignificant whereas they appear about as large as differences deemed significant (e.g., p.10444 lines 15-16). This all makes it look as though the analysis lacks objectivity and coherence. To clarify things and let readers evaluate differences objectively, result presentation and statistical significance should be made clearer. In particular on figures 5-6-7, the authors should i) indicate significance levels on the maps, either by whitening out non-significant points or maybe with contours; ii) present map of differences between runs CTRL and DRY (also for figure 3); iii) indicate significance levels (for differences) on these difference maps.

### **Response:**

**We have altered the description of the results to provide the reader with the exact locations that we are discussing. By explicitly stating the regions we discuss we allow the reader to more easily follow the argument. It may not have been clear, however the results in (now) Figures 6, 7, and 8 are greyed out when they are not statistically significant (as described in the figure captions). We have added text to explain that the insignificant results are greyed out.**

Another concern, which falls a little bit along the same lines as the one above, has to do with the role of transpiration (Tr) in total evapotranspiration (ET). Here as well the authors appear to contradict themselves several times: p.10443 l.14 “DJF. . . indicating surface evaporation is the dominant ET mechanism” p.10445 l.21: “acknowledging the importance of transpiration during the wet season” p.10446 l.7: “DJF...despite evaporation dominating the simulated ET” p.1046 l.14: “during DJF ... transpiration is partly governed by the water availability within the root zone” which implies that Tr plays a role in DJF ET and coupling. These seemingly contradicting statements reflect a lack of clarity in the corresponding processes that, it seems to me, could easily be alleviated by showing the different components of ET in the CLM outputs: soil evaporation, interception, Tr. In particular, the authors need to show this to back up their claim that ET in DJF is mostly soil evaporation, which in the manuscript rests essentially on Figure 3 and the claim that DJF ET is similar in both simulations despite different root-zone soil moistures (although  $\sim 10 \text{ W.m}^{-2}$  differences can be noted).

### **Response:**

**A previous incarnation of the manuscript included a separate Figure showing the fraction of ET that comes from transpiration for each season. To clarify the amount of ET owing to transpiration without adding more figures we have added the following text:**

**“The fractional contribution of transpiration to the total ET during DJF is roughly 10-30% for both DRY and CTRL (figure not shown) indicating that the evaporation is the dominant ET mechanism.”**

Another point where I thought the manuscript could be improved, was in discussing the physical processes diagnosed in the correlations: for instance, the sign of the EF-LCL coupling, or SM-LCL. This could easily be discussed in the manuscript. Similarly, there is no suggested explanation for why SM1-LCL coupling is positive in SON (significant positive correlation) while EF-LCL and SMrz-LCL are negative: what is happening in terms of SM1-EF, SM1-SMrz, etc.? This should be analyzed, so readers can have a better sense of why the SMrz-atmosphere coupling differs from the SM1-atmosphere one: how, why do SM1 and SMrz become uncoupled?

— Other comments

The use of several datasets to drive CLM4, although explained in detail in the Methods section, is never really exploited in the analysis. Figure 8 and text p.10446 lines 19-26 start to address inter-ensemble member differences, but do not draw any conclusion: what is to be concluded from figure 8? Differences are not very clear and, here as well, statistical significance should be addressed.

**Response:**

**We have added discussion as to why we analyze the standard deviation among the ensemble members. We also now discuss that the low standard deviation relative to the median values indicates that the sign of the Kendall tau results is robust despite the ensemble sampling multiple LCLs.**

Finally, I found the introduction to be long and lacking focus. I recommend the authors identify the problem they want to address and “zoom in” on it more clearly and rapidly. As it is now the issues addressed and the goal of the study do not stand out clearly.

**Response:**

**We have shortened the introduction to eliminate unnecessary text.**

—Comments along the text:

p.10432 line 24: “temperature” should go with atmospheric states. Figure 1: panels b c d are not discussed, thus should be removed.

**Response:**

**We have removed panels b,c, and d from Figure 1.**

Section 2: presentation of datasets and methods felt a bit backwards, as model validation is discussed before model and simulations: I would recommend reorganizing as: forcing datasets; model; obs; methods.

**Response:**

**We agree and have reorganized the manuscript per the reviewer suggestion.**

C4505

p.10441 line 13: "GLDAS. . . MERRA.BT" should be named earlier

**Response:**

**We have added the names to section 2.1.**

p.10441 line 18; "ensemble" the four members from 4 forcing datasets ?

**Response:**

**Yes.**

Section 4.1

Figure 2a: what time span is the X-axis ? What is the time resolution (monthly?) ? Figure 2b: would be more informative if showing % instead.

**Response:**

**We have added labels to the X axis. We have also changed the colors of Figure 2b to better show the differences between the AMSR-E product and the simulations.**

Figure p.10442 l.25: this statement feels awkward after a whole paragraph discussing differences between simulations. If SON ET is much lower in DRY than in CTRL, how can they both agree with observations?

**Response:**

**We have changed the last paragraph to read**

**“ The results from Figures 2,3 and 4 demonstrate that CLM4 simulates the monthly and seasonal first layer soil moisture and evapotranspiration reasonably. While the details of the**

**model performance vary depending on which site, season, and ensemble member is used for validation, overall the spatial and temporal patterns of ET and SM are generally captured by the modeling system. The accuracy of the estimated land surface states and fluxes therefore enables the use of the simulated variables in the diagnoses of the land-atmosphere association strength during the SON and DJF seasons.”**

Section 4.3: the physical meaning of these correlations should be explained (higher EF, lower LCL, etc.).

p.10444 1.21: “negative coupling” should be explained.

**Response:**

**We added the following text to the method section describing Kendall-tau:**

**“ The physical meaning of a negative SM-LCL  $K\tau$  association is as follows. A high value of SM will cause a larger ET flux, moistening the lower atmosphere, causing a lower LCL. Thus we hypothesis that in regions where the land-atmosphere are coupled the SM-LCL  $K\tau$  should be negative. If SM has no association with LCL than  $Kt$  is expected to be statistically insignificant. Similarly, if ET is negatively associated with LCL ( $Kt < 0$ ), it means that high ET maybe moistening the lower atmosphere again leading to a lower LCL. “**

p.10444 1.22-24: this statement should be “unpacked”, it is a bit unclear.

**Response:**

**We have reworded the statement as**

**“ The similarity in  $SM_1$ -LCL correspondence between CTRL and DRY during both DJF and SON implies a similar temporal variability of  $SM_1$  as related to the LCL. From Figure 3, the mean ET fluxes are considerably different during SON. The similar temporal behavior relative to the LCL for both DRY and CTRL indicates that the  $SM_1$  variability is physically independent of the season mean ET fluxes.”**

p.10445 1.1 : “slightly higher” where ?

**Response:**

**We added the following explanation:**

**“Some regions (17°S 126°E) exhibit increases in the magnitude of  $K\tau$  in CTRL relative to DRY in DJF (Figures 7a and 7c) although the differences are statistically insignificant over most of the domain”**

p.10446 1.4 : probably “figures 6 and 7”.

**Response:**

**The figures have been renumbered.**

Figure 5 caption: “morning time EF”: the text indicates it is afternoon EF (p.10439 1.4).

**Response:**

**Fixed this error as it was mislabelled in the caption.**