Manuscript hess-2015-414 entitled "Modelling evapotranspiration during precipitation deficits: identifying critical processes in a land surface model"

We would like to thank the reviewer for their constructive comments on our manuscript. This document outlines our point-by-point responses to the reviewer comments and the improvements made to the manuscript.

General comments:

The premise of this study, which is "to systematically evaluate the ability of land surface models to simulate [biological and physical] processes [and ecosystem dynamics] during soil moisture deficits" (lines 7-9, pg 10791) is a critical scientific objective for hydrological, ecosystem and climate change research and is a prerequisite for making predictions about how the function and services of different vegetation types will be altered by anthropocentric forcings in the coming century. The subject matter and scope of this study is appropriate for HESS. This study has potential to become a high impact and well cited paper. However, in this current version, this manuscript falls short in making significant advances in model understanding and in convincing me that they have appropriately interpreted their statistics. Therefore, this manuscript needs significant revisions before it should be considered for publication in HESS.

First, it strikes me as strange to use the variation that soil texture produces in model output as a way to evaluate the skill of the model to capture Qe of particular flux towers with a known soil texture. We would expect different soil textures to produce different magnitudes of Qe for different vegetation types. We would also expect different soil textures to produce different patterns of Qe during drying periods for different representations of the soil physics. These two results are indeed shown by the red curves in Figures 6 and 7. But what is not clear to me is how the observations relate to the variation produced by these two alternative formulations and by the contrasting soil type. Without observations of the same vegetation under the same climate, but growing on different soil types, there is no way to tell which model formula is correct. In other words, how would the observations change if the vegetation were growing on a different soil type? The default formula predicts that Qe would behave one way while the new formulation predicts that Qe would behave another way—but which one is correct, you cannot tell from the information presented in this study.

The same case as above can also be made for LAI since the authors prescribed it rather than evaluate the model's ability to predict it. Therefore, the authors should use the known LAI as a constraint on Qe in order to understand other aspects of the model that are poorly constrained. In contrast, how gs is regulated is not known and therefore, this type of comparison to observations does make sense.

We explore the role of key model inputs (LAI and soil) and alternative representations of hydrology and stomatal conductance to further understand biases in CABLE identified in previous studies (De Kauwe et al., 2015b; Li et al., 2012). We concentrate on dry-down periods as model simulations of Q_E are particularly poor during these periods.

We disagree that soil properties or LAI are well constrained at flux tower scales, and even less so at larger scales at which LSMs typically operate (≥0.5 degrees) in coupled models. Data on some soil properties can be obtained for flux tower sites, but even at these scales, soil properties are likely to be variable in space (Koster 2009) and uncertain across the tower footprint (typically around 1 km^2 but dependent on the height of the tower). Similarly, i*n situ* LAI is not a standard measurement provided at flux sites and, consequently, site LAI is typically derived from remotely sensed products (as was done here). Many studies have identified uncertainties in these products for representing both magnitude and timing of LAI variations (De Kauwe et al., 2011 and references therein).

We agree that testing model sensitivity to these inputs, as was done here, does not provide us process-level understanding of Q_E simulations. However, it is a useful exercise to understand the contribution from input uncertainties to the poor simulation of QE.

Using datasets relevant for global-scale applications where soil information is derived from coarse gridded data and highly uncertain at any particular location, we showed that despite varying these inputs, CABLE was unable to capture observed Q_E. This results contrasts previous work demonstrating reasonable simulations of monthly Q_E at large scales (Decker, 2015). This suggests that the representation of processes governing Q_E in CABLE must be insufficient.

Second, it is not clear how to interpret Figures 3, 4, and 8. We would expect NME and MBE to get increasingly large when the model is configured for a soil type that is not consistent with the known soil type for the flux tower. But, that doesn't mean the model is performing poorly ("performance" ln13 pg 10799). Indeed the model may be performing correctly for the vegetation that its soil type is configured for. In fact, I would be alarmed if NME = 0 even though soil texture was changed. Therefore, it should be argued that NME values near 0 or 1 indicate that the model is performing poorly when the soil type in the model does not match the observed soil type. But this distinction is not clear in the manuscript. More importantly, what can we really learn from the reported NME and MBE values when all values are groups together even though there is a mismatch between LAI and soil texture for some of the M_i and O_i values, but not others? How do they inform us in terms of model development when some of the values being put into M_i and O_i are not the same thing?

As discussed above, our analysis tests the sensitivity of CABLE to these inputs. If the model was sensitive to these parameters and able to capture observations with some combination of LAI and soil parameters, it would be hard to distinguish between poor process representation and uncertainties in model inputs, particularly in larger-scale applications where some inputs are poorly constrained at any particular location (as we discuss in the revised introduction). We show this is not the case: CABLE fails to capture observed Q_E in many cases no matter how the inputs are varied, highlighting deficiencies in process representations. We contrast two possible parameterisations for soil hydrology and stomatal conductance to explore this aspect further and show clear improvements in model performance when parameterising soil hydrology processes differently.

Third, the scope of the study is of limited appeal as it is presented. The manuscript is written as if it were speaking mainly to those interested in modelling the soil boundary condition for a land surface model. The model is just a tool for gaining more detailed understanding (or making predictions) about the system of interest. The study would be appealing to a much broader audience if the authors described what the predictions of the competing hypotheses (i.e. parameterizations) mean in terms how we understand ecology, physiology, and hydrology in a world with a changing climate, and not make the central focus of their discussion simply about model errors. As one example, the authors used two alternative formulas for gs, each representing very different hypotheses about stomatal regulation. Interestingly, the models predicted that the mode for stomatal regulation has very little effect on Qe during periods of water stress for all sites except Howard Springs (Figs. 6 & 7). This is a remarkable result with significant ecological, hydrological and climatological implications that needs to be expanded upon in the Discussion. There are many other example as well. After reading this paper, I did not come away with a clear sense about new hypotheses to test, observations and experiments to make, and model formulas to develop. (See also Specific comments 1c and 1d).

We have added additional sections in the Discussion to discuss each parameterisation separately to clearly identify why (or why not) each of the processes explored improves CABLE simulations of Q_E during dry-down. We have also highlighted model processes that should be explored in future work to resolve existing model biases but could not be constrained form available data at the flux sites analysed here.

Third, there is a considerable amount of information contained in the figures that should be flushed out in order to give greater clarity about the relative contribution each parameterization contributes to the variability in Qe. Take Figure 5 for example (but this comment pertains to all the figures), all of the "alternative LAI, gs, and soil parameterizations" (Fig.5 caption) are all mixed together to show the variation of the time series of Qe. Does one particular parameterization account for most of the variation on either the high end or the low end? If not, say so in the discussion. If so, what does the sensitivity (or lack thereof) to a particular parameterization mean in terms of the ecology, hydrology, physiology, and climatology of the different systems? What are the implications of the predictions of the different parameterizations? Constructing the analysis and discussion in this manner will give much clearer guidance to modellers and empiricists about modelling, experimental and observational needs.

As discussed above, we have extended the discussion to discuss each parameterisation in more detail to give the reader a better understanding of the wider implications of the findings and to clearly identify what processes our study shows are important for simulating Q_E during dry-down.

We have separated the effects of each parameterisation (hydrology, LAI, soil and g_s) in the figures. Figure 5 shows the range in Q_E separating the effects of the default and new hydrological schemes. The effects of LAI, soil and g_s are separated in Figures 6 and 7, separately for both hydrological schemes. We have added additional labels to the Figures to clarify the purpose of each figure.

Specific comments:

1. The message of this paper needs to be tighten-up considerably throughout the existing text and expanded upon in the Methods and Discussion. For example:

a. The Introduction is not particularly focused. It would be helpful if the Introduction were organized around a Problem Statement that is explicitly articulated at the beginning. The Problem Statement should address the culminating result of the study (i.e. lines 4-7, pg. 10804). Unfortunately, the reader has to get all the way to lines 7-9, pg 10796 before they encounter the actual Problem Statement that this analysis attempts to resolve.

We have reorganised the introduction to clearly discuss why we have chosen to explore uncertainties arising from hydrological and gs parameterisations and soil and LAI inputs. A better representation of hydrological processes has been identified as necessary for improving LSM simulations of drought (Tallaksen and Stahl, 2014) but has not been widely explored in previous studies. On the other hand, quantifying errors arising from parameter uncertainties is useful for separating parameter uncertainties from inadequate model parameterizations to identify where the model is unable to capture observations despite ranging key inputs, pointing to likely errors in model mechanisms.

b. Lines 19-25, pg 10791. Why do these models get these results and how do these results relate to the Problem Statement? In other words, what is the rationale for focusing on soil physics instead of biological processes? There is a huge body of literature that suggests we need to emphasize improving our understanding and representation of biological processes such as phenology or plant water-transport, rather than focusing on improving the soil boundary conditions.

Plant responses to drought in CABLE have specifically been explored elsewhere. De Kauwe et al. (2015b) and Li et al. (2012) implemented alternative plant water stress and root water uptake functions into CABLE but did not fully resolve existing biases in CABLE during dry-down periods. This manuscript explores other aspects of Q_E simulations, including soil hydrological processes and stomatal conductance, that are key model processes regulating Q_E fluxes but it is not known from previous studies if they account for underestimations of Q_E during dry-down. We hope the revised introduction addresses the importance of exploring these processes further. That said, we agree that there is a large literature highlighting the need to improve the representation of biological processes, and it is now becoming clear in the climate model literature that this risks an imbalance with the need to improve the hydrological literature. We suggest both are necessary, and reflect on this in the Discussion.

c. The Methods need to include equations for all of the alternative parameterizations examined. The Methods also need to include a Table of parameters and parameter values to maximize the transparency and reproducibility of this study.

We have added equations for the alternative stomatal conductance schemes in the Methods. It is not desirable to reproduce the large number of equations associated with the alternative hydrological schemes, these are fully documented in Decker (2015). However, we have included a number of key equations in the methods as they relate to the discussion later in the manuscript.

Table S2 fully details the soil parameters used in this study and Figure S7 shows the LAI values. Stomatal conductance parameters are available in De Kauwe et al. (2015a), which is freely available. We have referred the reader to this paper in section 2.2.3 of the revised manuscript.

d. The Discussion needs to map out how the equations and parameters (i.e. from 1b above) explicitly link to the different Results illustrated in the Figures. Without doing 1b and 1c, the model remains a bit of a black box, and therefore, it is difficult for modellers to know how to improve the existing formula and what specific parameters are controlling the output. Making these linkages is also important for informing empiricists on which field measurements should be prioritized.

We have added additional sections in the Discussion where we discuss each parameterisation separately, see earlier comment.

e. The influence of the "slope parameter" seems to be a key finding, yet it is given very little attention at the end of the Results and there is no mention of it in the Discussion. The authors state: "The slope appears more critical for simulation of Qe than the other parameterizations investigated here and has strong effect on the magnitude of the **fluxes primarily during dry-down**" (lines 19-21, pg 10803). The authors also state "our goal was to determine whether CABLE can **capture dry-down** associated with rainfall deficits as the components of the model are varied [among which is the hydrology scheme and slope parameter], or whether the model lacks the mechanisms to simulate this phenomenon" (lines 9-11, pg 10808). [Bold type face is the Reviewer's emphasis.] The authors fall short on meeting this goal when they fail to mention the role of one of the most "critical" parameters in the Discussion.

The slope parameter affects the rate of subsurface drainage. A steeper slope parameter increases drainage from lower soil layers, reducing soil moisture and aggravating plant water stress under dry conditions. We discuss this in more detailed in the Discussion of the revised manuscript.

f. Many statements throughout the Discussion need to clearly reference a figure (a few examples are given below). Also, each figure published in the Results section needs to be referenced and discussed in the Discussion section. Otherwise, any figure that is not discussed in the Discussion section should be moved to the Supplement because it is clearly not central to the main message of the study; rather, it is just supporting information.

We now reference each figure in the Discussion as they relate to the statements made, with the exception of Figure 1, which shows the location of study sites.

g.

2. Lines 23 & 28, pg 10800. “Likely due to” This is speculative in both cases. The beauty of using a model is that you can know these two things. By not exploring the output and knowing these for sure, statements like these are not very useful for either modellers or empiricists because they do not unequivocally tell us where to concentrate our efforts (or even worse—speculative statements can lead us down the wrong road). Also, “drying soil” and “compensating errors” both need to be quantified and demonstrated.

We have removed the speculative wording. We have added a supplementary figure showing soil moisture variations during the dry-down period.

3. Lines 21-22, pg 10806 “high soil evaporation **may** result from...” This is speculative. The authors can know this with closer inspection of the canopy turbulence output of their model.

We agree with the Reviewer that this comment was a little speculative and consequently we have removed it.

4. Lines 3-4, pg 10807. “seasonal droughts”. Do you mean dry season? I am not sure what a “seasonal” drought is. Droughts by definition are some type of water-deficit anomaly--be it measured in terms of rainfall, soil moisture, streamflow, etc—and anomalies are not seasonal, they are atypical. This is an important distinction to make because vegetation in areas with dry seasons are adapted for those dry seasons. However, depending on its severity, the plants may not be adapted for a drought that is layered on top of a dry season, which could be an

important ecological filter for certain species as climate changes.

We have corrected this to “seasonal-scale” as used elsewhere in the manuscript.

Technical comments:

Lines 24-26, pg 10790. Awkward sentence. Reorganize as: “LSMs form an integral part of global climate models by controlling how net radiation is partitioned...”

We have reorganised the sentence as per reviewer suggestion.

Lines 22 & 23, pg 10797 “87%” and “66%” These do not match Table S2.

We have corrected this in the manuscript.

Line 2, 10798. “empirical approach” What is this? Elaborate.

We have clarified this (see comment to Reviewer #1).

Lines 16-17, pg 10800. “Overall, both hydrological...” This sentence is not really true for all sites. E.g. see Harvard Forest or Umich.

This behaviour is typical of most sites, including Harvard Forest for some parameter choices (see Figure S1). The sentence does not suggest this applies to all sites (note “overall”).

Lines 9-10, pg 10802. “due to compensating biases” What are these?

The sentence goes on to explain this: “early season overestimations in Q_E are counteracted by underestimations during the dry-down periods”.

Line 3, pg 10804. “Have shown” Needs to reference a figure.

We have added reference to Figure 5.

Line 4, pg 10804. “have also shown” Needs to reference a figure.

We have added reference to Figure 6.

Line 19, pg 10804. “showed” Needs to reference a figure.

We have added reference to Figures 6 and 7.

Line 23, pg 10804. “the contribution of LAI (Fig. Xa), gs (Fig. Xb), and soil parameterisations (Fig. Xc)” Each parameterization needs to reference their respective figures.

We have added reference to corresponding figures in the text.

Line 2, pg 10805. “We identified” Needs to reference a figure.

We have added reference to the corresponding figure in the text.

Lines 6-8, pg 10805. Last sentence of the paragraph is not true for all sites. This sentence needs to include a qualifier at the end (before (Fig. S7)). For example, insert “for most sites”.

We have revised the sentence as per reviewer suggestion.

Lines 17-19, pg 10805. Which sites in Figure S7. Clearly reference the figure at the end of the sentence e.g. (Fig. S7a,b,d,f).

We have referenced specific sites.

Lines 9-10, pg 10807 and elsewhere in the text. “monthly climatology”. LAI is a vegetation property, not a property of the climate. Therefore, it strikes me as confusing when LAI is referred to as being part of the climatology.

Referring to mean monthly LAI (or other land surface property) as a monthly climatology is standard terminology employed in other studies (e.g. Oleson et al., 2008).

All Figures. The font size is way too small.

We have increased the font size of all figures.

Figure 8. Legend needs labels. And, caption needs to state what colours go with each of the parameters.

We have modified the figure and caption as per reviewer suggestions.

Table S2. Check numbers on “medium soil”. Should be decimals?

We thank the Reviewer for spotting this. We have corrected the silt, clay and sand fractions for the medium soil type.

References:

De Kauwe, M. G., Disney, M. I., Quaife, T., Lewis, P. and Williams, M.: An assessment of the MODIS collection 5 leaf area index product for a region of mixed coniferous forest, Remote Sens. Environ., 115, 767–780, doi:10.1016/j.rse.2010.11.004, 2011.

De Kauwe, M. G., Kala, J., Lin, Y.-S., Pitman, A. J., Medlyn, B. E., Duursma, R. A., Abramowitz, G., Wang, Y.-P. and Miralles, D. G.: A test of an optimal stomatal conductance scheme within the CABLE land surface model, Geosci. Model Dev., 8, 431–452, doi:10.5194/gmd-8-431-2015, 2015a.

De Kauwe, M. G., Zhou, S.-X., Medlyn, B. E., Pitman, A. J., Wang, Y. P., Duursma, R. A. and Prentice, I. C.: Do land surface models need to include differential plant species responses to drought? Examining model predictions across a latitudinal gradient in Europe, Biogeosciences, 12, 12349–12393, doi:10.5194/bgd-12-12349-2015, 2015b.

Decker, M.: Development and evaluation of a new soil moisture and runoff parameterization for the CABLE LSM including subgrid-scale processes, J. Adv. Model. Earth Syst., 7, doi:10.1002/2013MS000282.Received, 2015.

Li, L., Wang, Y.-P., Yu, Q., Pak, B., Eamus, D., Yan, J., van Gorsel, E. and Baker, I. T.: Improving the responses of the Australian community land surface model (CABLE) to seasonal drought, J. Geophys. Res., 117, G04002, doi:10.1029/2012JG002038, 2012.

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, a. and Qian, T.: Improvements to the community land model and their impact on the hydrological cycle, J. Geophys. Res. Biogeosciences, 113, doi:10.1029/2007JG000563, 2008.

Tallaksen, L. M. and Stahl, K.: Spatial and temporal patterns of large-scale droughts in Europe: model dispersion and performance, Geophys. Res. Lett., 41, 429–434, doi:10.1002/2013GL058954.Received, 2014.