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

We would like to thank the reviewers 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.

Response to Reviewer #1

General comments

This paper addresses the important issue of land surface model behaviour during lack of rainfall. Its plots are clear and the statistics appear sound, albeit somewhat basic. However, I felt somewhat empty-handed at the end. Very little, process understanding was gained. Why are the key equations not provided? There are various points in the paper where I get the impression the authors have used the model like a black-box without truly understanding the equations within the model. This is also evident from their description of the model physics, soil physics in particular. This is a missed opportunity and leaves the reader somewhat frustrated. I guess most of the Conclusions could have been drawn without having gone through this considerable modelling exercise. More in-depth explanation of the findings is required using equations presented and explored explicitly, not tentatively (using words such as 'likely', 'multiple explanations are possible', etc.

We have addressed the reviewer's concerns by including key equations in the paper to support the conclusions reached in the paper. We have also added more in-depth discussion to attach broader relevance to the paper's findings, including more detailed examination of each of the model process analysed here (soil, LAI, hydrology and stomatal conductance). We have provided additional supplementary figures to illustrate these points. However, we would like to point out that some explanations remain necessarily tentative due to a lack of observations to test specific model processes but the proposed mechanisms firmly rely on existing CABLE parameterisations.

Specific comments

Page 10792: I think the statements in lines 3-8 are somewhat naïve.

“We use *QE* because it is the variable that links the land surface energy, water and carbon budgets. It is also one of the variables supplied by the land surface to the atmosphere and is therefore important to a climate model. We do not use soil moisture as evaluating soil moisture from LSMs directly is problematic (Koster et al., 2009) due to different soil structures assumptions, storage capacity and timescales inherent in how LSMs represent this variable”.

Soil moisture, and the models underlying hydraulic properties and soil water transfer and root water uptake equations, ultimately determines the latent heat flux; via transpiration and direct soil evaporation. If a model gets QE right, but soil moisture content (considerably) wrong, this is a sign of poor process presentation particularly with regards to soil hydrology and plant water stress parameterisations.

Also, what exactly is meant by ‘soil structure assumptions’. This is unclear terminology.

We agree that evaluating LSMs against soil moisture would be valuable for better understanding hydrological processes in models and acknowledge this in the revised manuscript (section 2.1). However, many discrepancies exist between a real soil column and a model’s representation that make direct comparison extremely difficult, even if appropriate observations existed. Examples in this case include homogeneous soil properties with depth (there are no soil horizons), bedrock distribution and the commensurability of layer discretisation and measurement depths. In addition, few long-term in-situ observations are available for these flux tower sites that record soil moisture changes at appropriate depths (for example CABLE divides the soil into multiple layers with a total depth of 4.6 metres). Furthermore, *in-situ* measurements are highly localized in nature and strongly depend on local soil properties, making direct comparison to larger-scale models difficult (Koster et al., 2009). Remotely-sensed soil moisture products only record the top few centimetres of soil moisture (with the exact depth dependant on vegetation and soil moisture conditions) and typically from a coarser scale than the flux tower point fetch. Thus, these are not helpful for evaluating soil moisture outputs from LSMs in the context of our study. Given the discrepancies between modelled soil moisture and currently available observations, we feel that the uncertainty associated with evaluation against soil moisture means it is of limited value and flux tower measurements of Q_E remain a valuable alternative.

Line 25: soil texture is generally not a model parameter. It is used to derive other parameters from, such as hydraulic conductivity or the water retention curve. On the next page line 13 you use the term soil properties, which would be more appropriate.

The reviewer is right to highlight that in models soil texture is often used to derive other properties, such as hydraulic conductivity or the decomposition rate in CENTURY-type models (and thermal conductivity in CABLE). Nevertheless, it is correct to describe soil texture, the fraction of sand, silt and clay as a model parameter in the CABLE (and other) models. We have replaced soil "texture" with "properties".

Page 10792:

Line 15-17 You say "Where the LSM cannot capture the observations, despite variations in LAI and soil parameters, points to systematic errors in the model's representation of physical processes"

I guess we could consider this roughly to be the case, but this ignores errors in driving variables and energy balance closure errors, or the fact that your parameter range was possibly unsuitable.

We agree with the reviewer that flux tower measurements themselves contain errors, notably in energy balance closure. We have acknowledged this in the Discussion of the revised manuscript (Section 4.1). However, it appears highly unlikely that the CABLE biases identified here are solely an artefact of erroneous driving/evaluation data due to the large and strongly seasonal nature of Q_E biases. In fact, Haughton et al. (in review) showed based on the PLUMBER results (Best et al., 2015) that problems with energy balance closure in flux tower data do not account for the poor performance of LSMs (including CABLE) when compared against simple benchmarks. We have employed the forcing data and Q_E observations from the PLUMBER study.

We have tested a very large range in soil parameters by varying soil properties from sandy to clay soil type. These represent the extreme soil types available in the standard CABLE soil input dataset but it is not possible to quantify if this soil dataset reflects the full range of soil properties at the flux tower sites. LAI was varied by site depending on the remotely sensed MODIS data. It is possible that the range in LAI does not capture site LAI in some cases. However, this and previous studies (Kala et al., 2014) have pointed to a limited sensitivity of CABLE-simulated total Q_E to LAI. It is thus unlikely that LAI errors could account for the high seasonal biases in Q_E.

Also, it is not sure that the error in QE was related to soil hydrological parameterisations. It could just as well have to do with soil thermal and land surface radiative parameterisations, affecting sensible and soil heat flux.

The nature of the errors in Q_E and the responses of those errors to the parameters and parameterizations strongly points to errors being related to the hydrology. This is supported by Haughton et al. (in review), who found that errors in the partitioning between latent and sensible heat, rather than in the calculation of net radiation and ground heat flux, accounted for the poor performance of LSMs when evaluated against simple benchmarks. Without independent observations of all elements of the surface radiation balance it is of course impossible to be certain but it is a reasonable conclusion based on results across multiple sites that the hydrology is the problem.

Page 10795:

Lines 8-9: “The soil module simulates the transfer of heat and water within the soil and snowpack following the Richards equation”.

This is incorrect: soil heat transfer cannot be determined with the Richards equation. It is generally determined with the Fourier’s law.

The reviewer is of course correct and we have removed this reference to heat transfer.

Page 10796:

Line 4-5: “It does not distinguish between saturated and un-saturated top soil fractions or simulate groundwater dynamics”

This sentence needs elaborating. Distinguish in what way? In the context of surface run-off? At the moment it reads as if model soil moisture plays no role in any soil hydrological process.

We have clarified that the old scheme does not simulate saturated and unsaturated top soil fractions separately but rather treats the top soil layer as one entity, which is either fully saturated or unsaturated. Neither does it take account of water storage in groundwater aquifers, or recharge from those water stores. These are two key differences between the new and default schemes and are therefore highlighted in the text.

Page 10796: Line 24: ..... Table S2.

Table S2 in Supplementary material contains soil physically incorrect terminology. It should be “Soil dry bulk density” not simply “Soil density”.

Also, “suction at saturation” is per definition equal to zero. What you mean is “suction at air entry point”. Furthermore, suctions always have positive values. If you use negative values, as in Table S2, it should be referred to as “matric potential” or rather “matric head” as you are working in length units.

Finally what is meant exactly by soil heat capacity? At air-dry or saturated moisture content? Why are these values the same for all soil texture types?

We have corrected the terminology in Table S2 as per reviewer comments. The dry soil heat capacity has an identical value across the three soil classes as 850 (J/kg/C) was the value provided for the sandy and clay soil types in the standard CABLE soil input data set used in this study (provided with the standard CABLE distribution). The “medium” soil class uses the median value of the sandy and clay soils and thus also uses the same value.

Page 10798:

Lines 1-4: “The default hydrological scheme uses these three soil parameter sets directly, whereas the new scheme employs an empirical approach to calculate the parameters governing water holding and thermal capacities from sand, silt and clay fractions”

I am not sure what is meant here? The default scheme also has values for wilting point moisture content etc. They must also have been derived from texture?

The default scheme uses all soil parameter values in Table S2 as inputted, whereas the new scheme only uses inputted sand, silt and clay fractions and calculates the other eight parameters from these texture parameters using pedotransfer functions (see comment immediately below). We have clarified this in the text.

In both cases using so called ‘pedotransfer functions’? Use this word instead of “empirical approach”. Which ones were used? By the looks of it Cosby et al., seeing you are using the Clapp and Hornberger B parameter? You need to state this explicitly.

Also, these soil hydraulic parameters govern more than just water holding capacity. They govern soil water transfer via Darcy’s law and Richard’s equation (with Ks embedded in them).

Finally: you are using heat capacity, but is thermal conductivity not required in Fourier’s law?

The new and old schemes use Clapp and Hornberger (1978) relations to relate soil moisture, soil matric potential, and hydraulic conductivity. The old scheme uses a look up table to relate the inputted soil class (e.g. sand or clay) to the parameters in the Clapp and Hurnberger formulation. The new scheme, however, follows the work of Cosby et al. (1984) and uses pedotransfer functions to relate the sand, silt and clay content to the parameters in the Clapp and Hornberger equations, thereby negating the need for a look up table. We have clarified this in the text and replaced “water holding capacities” with “hydraulic properties”.

CABLE calculates thermal conductivity from soil texture parameters (sand, clay and silt fractions) and thermal conductivity is thus not an input parameter.

Line 6-7:” Leaf area index (LAI) plays an important role in the surface energy balance in CA- BLE (Kala et al., 2014)”.

Can this sentence be elaborated upon by 1-2 follow-on sentences? In what way? By upscaling from leaf to canopy scale conductance? In light interception?

We have added further details on the use of LAI in CABLE. It is used to calculate aerodynamic resistances, partitioning the absorbed radiation flux between sunlit and shaded leaves and to scale fluxes from the leaf to the canopy.

Page 10799:

Line 1-3: "...... the dry periods were defined based on precipitation as this allowed the use of available observations, but we note the simulated fluxes will also depend on other processes such as soil moisture availability".

The simulated fluxes will also very much depend on the driving variables that determine evaporative demand. That has been overlooked in this definition of 'dry periods'.

As we point out in the text, there are various ways to define a "dry" period. We have added evaporative demand as one such definition. We have chosen to use precipitation as this relies directly on available observations, but acknowledge that many other indices could have been chosen. As the dry periods used in this study coincide with the hottest part of the year in most cases, an alternative definition of a dry period taking into account evaporative demand should broadly coincide with the time periods used here. Given the large biases in CABLE that clearly coincide with periods of low rainfall (Fig. S6), we do not believe that the choice of drought definition compromises the results.

Line 8-10. "The dry-down period generally coincides with the maximum and the following minimum observed latent heat flux during the one-year period but has been adjusted for some sites to best capture typical model behaviour (Fig. S6)".

What is meant by this exactly?

We mean that usually the periods to be analysed are chosen via an automated procedure but on occasions we adjust the period chosen where there is a better example that is not identified automatically. We have modified the text to elaborate this.

I feel what is missing from the paper is a basic description of plant water stress, e.g. along the lines of the beta function in the Jules model (see Egea et al. 2011). I understand CABLE uses the same approach?

We have added a description of the plant water stress function in CABLE in the Methods. CABLE uses a similar beta function to JULES and a number of other LSMs.

Page 10800 Lines 23-24: “This is likely due to overly rapid drying of top soil layers, which strongly control *QE* in CABLE (De Kauwe et al., 2015c)” But which process is being affected (mostly) here? Transpiration or evaporation?

In both cases soil moisture content is a key variable, yet we do not get any insight into how well this variable is predicted by the model.

We have clarified in the text that transpiration is affected and elaborated on the reasons behind the rapid drying. We have also added a supplementary figure showing the individual components of evapotranspiration (soil evaporation and transpiration) during the dry-down periods.

In addition, we have added a supplementary figure showing soil moisture variations for the new and default hydrological schemes for each of the six soil layers during the one-year periods. Unfortunately soil moisture observations do not exist for soil depths used in CABLE (4.6 m) or are not made freely available at these sites. We agree that the model should ideally be evaluated against soil moisture as an additional constraint but this was not possible and, in common with many other studies (e.g. De Kauwe et al., 2015; Li et al., 2012; Whitley et al., 2015), we have instead relied on Q_E measurements.

Lines 23-24: “This is particularly evident during warmer summer months when fluxes are more strongly moisture-limited”

One may assume that this is indeed the case, but with no information on soil moisture content, nor how SMC affects evapotranspiration, this remains speculative.

We have added a supplementary figure showing soil moisture variations (see comment immediately above). We have also added the plant water stress and soil evaporation functions in the Methods to show how soil moisture limits both components of ET.

Lines 27-28: “While encouraging, this is likely due to compensating errors, such that early season overestimations in *QE* are counteracted by underestimations during the dry-down periods.“ Remove the word likely. It either is or it isn’t. You have the data in front of you.

We have removed ‘likely’.

Page 10801 Line 8: “The model dries down too quickly”. The model itself is not drying...

We have rephrased the sentence.

Line 12-14: “These characteristics of CABLE are not dependent on the choice of LAI, *g*s, or soil parameters; the range in *QE* fails to overlap the observations irrespective of how these properties are varied.”

I have not been able to find anywhere in the paper between what values these variables have been ranged. Nor am I sure what is meant by soil variables. Are these the soil hydraulic variables? Or texture percentages?

The range in LAI is shown in Figure S5 and referenced clearly in section 2.2.5, which also details how the LAI time series were generated and varied between model runs. Soil parameters are discussed in detail in section 2.2.4 and the parameter values of each of the three soil classes that were used as CABLE inputs are displayed in Table S2. The soil parameters were varied between the values of three soil classes. We have reworded the sentence to state that the results are not dependent on the choice of LAI or soil inputs or g_s parameterisations.

Page 10802

Line 19-23: “Both hydrological schemes are sensitive to soil parameters during the dry- down period but show smaller variations due to soil during other parts of the year (see Amplero, Blodgett, Howard Springs and Palang in Figs. 6 and 7). This transition from low to high sensitivity occurs as soil moisture stores begin to deplete and *QE* becomes increasingly limited by moisture supply”.

This last sentence seems a very obvious statement. Does it need stating? How meaningful is it anyway, if we are not told (at least approximately) how CABLE deals with plant water stress?

As stated in a previous comment, we have added an equation for plant water stress and a figure showing soil moisture variations during the dry-down period.

Line 23-25: “The new hydrological scheme uses a narrower range of parameter values for water holding capacity and conductivity (Table S2) and thus results in a smaller range of uncertainty due to soil parameters.”

As far as I can see Table S2 does not give a range per soil parameter. It gives one value only for each soil texture type. Also, what is meant by water holding capacity? Water holding capacity is defined as the total amount of water a soil can hold at field capacity. Did you mean ‘water retention curve’. Furthermore, ‘conductivity’ needs to read ‘hydraulic conductivity’

The model was ran using three alternative soil classes as per Table S2, generating a range of 3 parameter values in simulations. Table S2 details the parameter values for each soil class. The sentence in question was removed from the text.

Page 10803

“The slope parameter affects the rate of subsurface drainage and represents a key difference between the new and default schemes”

If this is the case, then why the emphasis on water retention? These parameters are generally related to plant water availability.

We are not sure what the Reviewer means here. We hope that the additional edits we have implemented will satisfy the Reviewer but are open to further changes if their comment can be clarified.

Why is the slope parameter not in Table S2?

The slope parameter is not a soil parameter but was varied separately in additional simulations between 0 and 5 degrees using the "medium" soil class and new hydrology only, as detailed in Section 2.2.2.

What meaning does this parameter have anyway at the site scale?

The slope parameter can be obtained from high-resolution elevation data at scales relevant for flux tower sites. In this case, the slope reflects the average slope of a 3x3 km area centred around the tower. The surface slope derived with this method only reflects the slope at a scale larger than the footprint of the site and, as we acknowledge in the manuscript, it is a first order approximation to the features affecting subsurface drainage such as bedrock slope and lateral heterogeneity in subsurface properties. However, as the model was shown to be sensitive to the choice of slope, we believe it is important to discuss the effect of the slope parameter on Q_E simulations.

Page 10804

Seems this section makes a lot of rather obvious statements that could have been made without the work conducted in this paper.

The Reviewer demonstrably has a deep understanding of the areas covered by this paper. We suggest that other readers may not see all the statements as obvious and in any case, demonstrating statements supported by evidence is useful. We do agree that some of our discussion and conclusion is fairly straightforward for the expert, but we are also writing for readers who may not have the awareness of the Reviewer.

Page 10805

Lines 5-7: “The reason for the overestimation of peak fluxes is not clear but is not resolved by the new hydrological scheme despite this parameterising many of the relevant processes differently”.

This sentences underlines the frustrating nature of this paper: CABLE is used like a black-box without the authors truly having investigated the inner workings of the model. I doubt this is acceptable to the journal and its readers.

Why/how would changes in a slope parameter affect soil evaporation?

We have added equations to show how soil evaporation is treated in each hydrological scheme. We also demonstrate that these changes do not resolve the biases. As several of the processes that can be responsible for the excessive soil evaporation cannot be constrained from observations at these flux sites, it is not possible to present definitive reasons behind these biases. Such reasons include errors in LAI. This is a systematic bias in CABLE and warrants a separate body of work to fully understand the model errors relying on alternative data streams. We simply highlight this excessive spring Q_E as a possible mechanism for depleting soil moisture stores earlier than observed but without an alternative scheme available to test, it is not possible to fully quantify this in the present study.

In terms of the slope parameter, if slope increases such that runoff increases soil evaporation is affected by water balance constraints.

Line 9: “..multiple potential causes to this excessive QE??” There are only a limited number of causes and they are all embedded in the equations that determine soil evaporation.

Multiple causes was used to highlight that there are a number of processes affecting simulation of soil evaporation, which the section goes on to discuss. We have reworded this sentence. However, soil evaporation can be affected by causes unrelated to the equations that describe soil evaporation via changes in state variables.

Page 10806

Line 9-11 “Other model processes, particularly vegetation response to drought, have been identified as critical for capturing drought processes and shown to improve CABLE performance during droughts but were not explored here.”

I do not understand this? Wilting point and field capacity are two key parameters in the CABLE model plant water stress factor (empirical scalar beta, see De Kauwe et al. 2015).

So implicitly you have explored vegetation response to drought?!

The Reviewer is correct that we vary the wilting point and field capacity parameters, which are part of the water stress function. However, we do not contrast alternative formulations for plant water stress as this has been done elsewhere (De Kauwe et al., 2015b; Li et al., 2012; Egea et al. 2011). Instead, we have explicitly explored hydrological processes and errors in model inputs as alternative explanations for poor simulation of Q_E during dry-down, which has not been fully resolved in previous studies.

One can also argue that the water stress function in its current form does not reflect vegetation drought responses as we understand them from experiments, as it takes no account of plant adaptations to drought in dry environments (De Kauwe et al., 2015). It mostly reflects properties of soil, and has little empirical support (Medlyn et al., in review). This section discusses the limitations of this approach and discusses alternatives presented in other studies as a means to improve drought simulations. We have reworded parts of this section to clarify this.

References:

Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. a., Dong, J., Ek, M., Guo, Z., Haverd, V., van den Hurk, B. J. ., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. A., Stevens, L. and Vuichard, N.: The plumbing of land surface models: benchmarking model performance, J. Hydrometeorol., 16, 1425–1442, doi:10.1175/JHM-D-14-0158.1, 2015.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14, 601–604, 1978.

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, 2015.

Haughton, N., Abramowitz, G., Pitman, A. J., Or, D., Best, M., Johnson, H. R., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P., Dong, Z., Haverd, V., van der Hurk, B., Nearing, G., Pak, B., Santanello, J. A., Stevens, L. E. and Vuichard, N.: The plumbing of land surface models: is poor performance a result of methodology or data quality?, J. Hydrometeorol., in review.

Kala, J., Decker, M., Exbrayat, J.-F., Pitman, A. J., Carouge, C., Evans, J. P., Abramowitz, G. and Mocko, D.: Influence of Leaf Area Index Prescriptions on Simulations of Heat, Moisture, and Carbon Fluxes, J. Hydrometeorol., 15, 489–503, doi:10.1175/JHM-D-13-063.1, 2014.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K. and Puma, M. J.: On the nature of soil moisture in land surface models, J. Clim., 22, 4322–4335, doi:10.1175/2009JCLI2832.1, 2009.

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.

Medlyn, B., De Kauwe, M., Walker, A., Duursma, R., Luus, K., Mishurov, M., Pak, B., Smith, B., Wang, Y.-P., Yang, X., Crous, K., Drake, J., Macdonald, C., Norby, R., Power, S., Tjoelker, M. and Ellsworth, D. S.: Using models to guide experiments: a priori predictions for the CO2 response of a nutrient- and water-limited mature Eucalypt woodland, Glob. Chang. Biol., in review.

Whitley, R., Beringer, J., Hutley, L., Abramowitz, G., De Kauwe, M. G., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y.-P., Williams, M. and Yu, Q.: A model inter-comparison study to examine limiting factors in modelling

Australian tropical savannas, Biogeosciences Discuss., 12, 18999–19041, doi:10.5194/bgd-12-18999-2015, 2015.

Response to Reviewer #2

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.

Modelling evapotranspiration during precipitation deficits: identifying critical processes in a land surface model

A.M. Ukkola[1], A.J. Pitman[1,2], M. Decker[1,2], M.G. De Kauwe[3], G. Abramowitz[1,2], J. Kala[1,4] and Y.-P. Wang[5]

[1]ARC Centre of Excellence for Climate System Science, University of New South Wales, Kensington, NSW 2052, Australia

[2]Climate Change Research Centre, University of New South Wales, Kensington, NSW 2052, Australia

[3]Department of Biological Sciences, Macquarie University, Balaclava Road, North Ryde, NSW 2109, Australia

[4]Murdoch University, School of Veterinary and Life Sciences, Environmental and Conservation Sciences, Murdoch, 6150, WA, Australia

[5]CSIRO Ocean and Atmosphere, Aspendale, Victoria 3195, Australia

Correspondence to: A.M. Ukkola (amukkola@gmail.com)

For submission to *Hydrology and Earth System Sciences*

Abstract

Surface fluxes from land surface models (~~LSM~~LSMs) have traditionally been evaluated against monthly, seasonal or annual mean states. The limited ability of LSMs to reproduce observed evaporative fluxes under water-stressed conditions has been previously noted, but very few studies have systematically evaluated these models during rainfall deficits. We evaluated latent heat ~~flux~~fluxes simulated by the Community Atmosphere Biosphere Land Exchange (CABLE) LSM across 20 flux tower sites at sub-annual to inter-annual time scales, in particular focusing on model performance during seasonal-scale rainfall deficits. The importance of key model processes in capturing the latent heat flux ~~are~~was explored by employing alternative representations of hydrology, leaf area index, soil properties and stomatal conductance. We found that the representation of hydrological processes was critical for capturing observed declines in latent heat during rainfall deficits. By contrast, the effects of soil properties, LAI and stomatal conductance ~~are shown to be~~were highly site-specific. Whilst the standard model performs reasonably well at annual scales as measured by common metrics, it grossly underestimates latent heat during rainfall deficits. A new version of CABLE, with a more physically consistent representation of hydrology, captures the variation in the latent heat flux during seasonal-scale rainfall deficits better than earlier versions but remaining biases point to future research needs. Our results highlight the importance of evaluating LSMs under water-stressed conditions and across multiple plant functional types and climate regimes.

1 Introduction

~~Land surface models (LSMs) simulate the exchange of water and energy between the land surface and the atmosphere (Pitman, 2003). They control how net radiation is partitioned between sensible and latent heat (Q_E), and how rainfall is partitioned between evaporation and runoff and therefore form an integral part of global climate models. LSMs have been extensively evaluated for simulated water, energy and carbon fluxes, typically at seasonal to inter-annual time scales (Abramowitz et al., 2007; Best et al., 2015; Blyth et al., 2011; Dirmeyer, 2011; Zhou et al., 2012). LSMs have been found to perform reasonably well under well-watered conditions (e.g. Best et al., 2015) but are less able to capture fluxes during water-stressed conditions (Li et al., 2012). Given the long history of systematic model evaluation (Best et al., 2015; Dirmeyer, 2011; Henderson-Sellers et al., 1995), it is remarkable that very few studies have systematically evaluated the ability of LSMs to simulate hydrological processes during precipitation or soil moisture deficits, although Powell et al. (2013) and Prudhomme et al. (2011) are some notable exceptions.~~

~~The ability to simulate a drying landscape is one of the pre-requisites for reliable projections of drought by a LSM and therefore by a climate model.~~ Droughts are expected to increase in frequency and intensity (Allen et al., 2010; Trenberth et al.,

2014) in some regions due to the effects of climate change (Collins et al., 2013). This would have profound implications for affected regions and their socio-economic systems. ~~Unfortunately, there are large uncertainties in the evolution of historical (Dai, 2012) and future~~Land surface models (LSMs) are a key tool for understanding the evolution of historical droughts and predicting future water scarcity when coupled to global climate models ~~(Prudhomme et al., 2014) droughts. This is associated, in part, to differences in projected climate, particularly regional scale rainfall, or sparse observational records in the case of historical drought trends (Dai, 2011). However, there are also major weaknesses in the capability of the LSMs used in climate models to capture drought events. For example, Prudhomme et al. (2011) showed large differences between three global hydrological models in simulating historical hydrological droughts (i.e. streamflow deficits) in European catchments. The participating Joint UK Land Environment Simulator LSM (JULES; Best et al., 2011) was shown to overestimate both the duration and severity of droughts. Similarly, the projected occurrence of future droughts~~(Dai, 2012; Prudhomme et al., 2014). LSMs have been extensively evaluated for simulated water, energy and carbon fluxes, typically at seasonal to inter-annual time scales (Abramowitz et al., 2007; Best et al., 2015; Blyth et al., 2011; Dirmeyer, 2011; Zhou et al., 2012), and been found to perform reasonably well under well-watered conditions (e.g Best et al., 2015). However, recent studies have indicated that the ability of current LSMs to simulate these fluxes during water-stressed conditions is limited (De Kauwe et al., 2015b; Powell et al., 2013). LSMs have been shown to poorly characterise the magnitude, duration and frequency of droughts when evaluated against site- and catchment-scale observations of latent heat (Q_E) and streamflow (De Kauwe et al., 2015b; Li et al., 2012; Powell et al., 2013; Prudhomme et al., 2011; Tallaksen and Stahl, 2014). Similarly, LSM projections of future drought occurrence has been shown to be highly model dependent, with greater uncertainty in future projections arising from differences between ~~hydrological models~~LSMs than from the climate model projections used to force them (Prudhomme et al., 2014). Further, changes in soil moisture in the future are also linked with changes in the probabilities and intensities of other extremes including heatwaves (Seneviratne et al., 2010). Clearly, a better understanding of limitations in LSMs under more extreme conditions, and ultimately improved performance by these models, is necessary for improving the future projections of drought and other land-surface influenced extremes by climate models.

We investigate the performance of the Australian Community Atmosphere-Biosphere Land Exchange (CABLE) in simulating observed declines in ~~Q_E during rainfall deficits. We use Q_E because it is the variable that links the land surface energy, water and carbon budgets. It is also one of the variables supplied by the land surface to the atmosphere and is therefore important to a climate model. We do not use soil moisture as evaluating soil moisture from LSMs directly~~latent heat during rainfall deficits.~~ is problematic (Koster et al., 2009) due to different soil structures assumptions, storage capacity and timescales inherent in how LSMs represent this variable. There are also very few examples of full soil moisture profile observations with co-located meteorological forcing. By using Q_E, we can identify when this flux begins to become limited by soil moisture and then evaluate CABLE against these "dry-down" periods. This forms an important first step towards establishing the capability of CABLE to simulate drought events. This is particularly important because~~ CABLE is the LSM used within the Australian Community Climate and Earth Systems Simulator (ACCESS; Bi et al., 2013), a global climate model which has participated in the 5th assessment report of the International Panel on Climate Change (IPCC, 2013) and is used for numerical weather prediction research in Australia (Puri et al., 2013). ~~If CABLE, or other LSMs, over- or underestimate the magnitude of Q_E at the onset of rainfall deficits, or drop too quickly or slowly during periods of low rainfall, they are also likely to fail to capture the magnitude and intensity of droughts.~~

~~We examine CABLE in the context of two key areas of uncertainty: how processes are parameterised and how associated parameters are selected. In terms of model parameterisations, we examine the hydrology and stomatal conductance modules. These have recently been revised~~In common with other LSMs (Prudhomme et al., 2011; Tallaksen and Stahl, 2014), CABLE has been found to poorly simulate the evolution of droughts, systematically underestimating site-scale Q_E during seasonal-scale droughts (De Kauwe et al., 2015b; Li et al., 2012). There could be many reasons for this systematic error. For example, unrealistic representation of plant drought responses has been identified as a major limitation in LSM simulations of drought (Egea et al., 2011; Powell et al., 2013), including CABLE (De Kauwe et al., 2015b). Recent studies have revised vegetation drought responses in CABLE but not fully resolved existing model biases (De Kauwe et al., 2015b; Li et al., 2012). In this paper, we examine CABLE in the context of another major area of uncertainty: how hydrological processes are parameterised and how associated parameters are selected. A better representation of hydrological processes, particularly soil moisture, has been identified as necessary for improving LSM simulations of drought (Tallaksen and Stahl, 2014) but this has not been widely explored. The parameterisations of soil hydrology and stomatal conductance (g_s) have recently been revised in CABLE and shown to improve seasonal to annual scale simulations of Q_E in CABLE ~~(Decker, in review; De Kauwe et al., 2015a). In terms of parameters, we quantify the uncertainty arising from soil texture and leaf area index (LAI) inputs. CABLE-simulated~~ ~~Q_E~~ ~~has been shown to be sensitive to these parameters but they remain uncertain at both site and large scales (Kala et al., 2014; Zhang et al., 2013). While other parameters, including other vegetation characteristics such as rooting depth (Li et al., 2012), are potentially important, soil texture and leaf area index can be constrained from readily available global-scale datasets widely used in large-scale LSM applications~~(De Kauwe et al., 2015a; Decker, 2015). We explore if changes to these hydrological processes can also improve simulations of Q_E during dry periods and to guide development of more realistic drought mechanisms in LSMs.

~~We note it is possible to measure or calibrate some key parameters at site scales (Gupta et al., 1999; Leplastrier et al., 2002; Wang et al., 2001) to enable a LSM to better match an observed time series. This is not our goal; we are not trying to demonstrate how well CABLE can reproduce observed~~ ~~Q_E~~ ~~*per se*. Rather, we are examining where the model appears to fail for lack of process-level representation of a land surface component in ways that should be applicable to larger-scale applications of the model. While using calibration would improve a model's metrics, calibration is likely to lead to parameter values that perform well at the calibration site, but are likely to be over-fitted for broad scale application. Rather than using calibration, we utilise observed datasets that are similar or identical to those commonly used to run CABLE (and other LSMs) at regional to global scales. We examine whether variations in the LAI and soil properties, informed by the observations, enable our LSM to accurately capture the observations of~~ ~~Q_E~~~~. Where the LSM cannot capture the observations, despite variations in LAI and soil parameters, points to systematic errors in the model's representation of physical processes.~~

~~We therefore evaluate CABLE's simulation of Q_E across 20 globally distributed flux tower sites, covering a range of climates and vegetation types, with a specific focus on identifying key processes controlling model performance during rainfall deficits. While our analysis is restricted to site scales, flux tower measurements are direct observations of latent heat (Wang and Dickinson, 2012) and offer a valuable reference for model performance.~~

We quantify the uncertainty arising from key model parameters: soil properties and leaf area index (LAI) inputs. CABLE-simulated Q_E has been shown to be sensitive to these parameters but they remain uncertain at both site and large scales (De Kauwe et al., 2011; Kala et al., 2014; Koster et al., 2009; Zhang et al., 2013). Quantifying the sensitivity of CABLE to LAI and soil properties is useful for separating parameter uncertainties from inadequate model parameterisations. Where the LSM cannot capture the observations, despite variations in LAI and soil parameters, systematic errors in the model's representation of physical processes are probable (assuming negligible errors in flux tower data used to drive and evaluate the model). While other parameters, including additional vegetation characteristics such as rooting depth (Li et al., 2012) are also potentially important, soil properties and leaf area index can be constrained from readily available global-scale datasets widely used in large-scale LSM applications.

We therefore explore CABLE performance at 20 flux tower sites distributed globally. We contrast model behaviour at annual to sub-seasonal scales to explore uncertainties in hydrological processes and parameters under conditions ranging from wet to dry. We concentrate on the ability of the model to capture the onset of drought in a drying phase as a pre-requisite for capturing the magnitude and intensity of droughts.

2 Methods

2.1 Flux tower sites

~~We evaluate CABLE performance against eddy covariance measurements across 20 flux tower sites globally (Fig. 1 and Table S1). The~~The flux tower data were collated as part of the Protocol for the Analysis of Land Surface models (PALS; Abramowitz,

2012) Land sUrface Model Benchmarking Evaluation pRoject (PLUMBER; Best et al., 2015), originally obtained through the Fluxnet LaThuile Free Fair-Use subset (fluxnet.ornl.gov). The PLUMBER sites ~~were chosen to~~ represent a broad range of vegetation and climate types, ~~whilst~~and were also ~~maximising~~selected to maximise the length of measurement records (Best et al., 2015). Here, we focus on the results for six sites with a pronounced period of low precipitation~~,~~ (Fig. 1 and Table S1), each representing a different climate and vegetation type, but provide results for all study sites as Supplementary Information (Fig. S1-S4).

The 20 flux tower sites provide meteorological and flux measurements at 30-minute resolution. The observed meteorological data (precipitation, short- and long-wave radiation, surface air pressure, air temperature, specific humidity and wind speed) were used to drive CABLE simulations. Observed Q_E was then used to evaluate simulations because it is the variable that links the land surface energy, water and carbon budgets (Pitman, 2003). It is also one of the variables supplied by a LSM to the atmosphere and is therefore important to a climate model. We note that it would also be desirable to evaluate soil moisture outputs from LSMs. Ultimately this is problematic (Koster et al., 2009) ~~time steps~~as site measurements at depths which reflect the plants' root-zone access (i.e. deeper than the top few centimetres) are not readily available~~. The observed meteorological data (precipitation, short- and long-wave radiation, surface air pressure, air temperature, specific humidity and wind speed) were used to drive CABLE simulations. Observed~~ ~~Q_E~~ ~~was used to evaluate simulations.~~

2.2 Description of the CABLE LSM and model parameterisations

2.2.1 General description

The Community Atmosphere-Biosphere Land Exchange (CABLE) model is~~ a LSM~~ used to simulate energy, water and carbon fluxes and the partitioning of net radiation into latent and sensible heat fluxes. It can be employed offline with prescribed meteorology, as in this paper, or within the Australian Community Climate Earth System Simulator coupled climate model (ACCESS; Bi et al., 2013). It has been used widely in coupled (Cruz et al., 2010; Lorenz and Pitman, 2014) and offline (Haverd et al., 2013; Huang et al., 2015; Zhou et al., 2012) simulations and has been extensively evaluated against flux site (De Kauwe et al., 2015b; Li et al., 2012; Wang et al., 2011; Williams et al., 2009) and regional to global-scale observations (De Kauwe et al., 2015a; Decker, 2015). Previous model inter-comparisons have shown that simulated latent and sensible heat fluxes perform comparably to other LSMs (Best et al., 2015).

CABLE consists of sub-models for radiation, canopy, soil and ecosystem carbon. Canopy processes are represented with a two-leaf model, which calculates photosynthesis, stomatal conductance and leaf temperature separately for sunlit and shaded leaves ~~(Leuning, 1995; Wang and Leuning, 1998). The soil module simulates the transfer of heat and~~(Leuning, 1995; Wang and Leuning, 1998). The soil module simulates the transfer of water within the soil and snowpack following the Richards equation. CABLE has 11 plant functional types (PFT). A detailed description of model components can be found in Wang et al. (2011).

We ~~ran~~used CABLE version 2.0 (revision 2902; https://trac.nci.org.au/trac/cable/wiki) forced with site-specific meteorological data at 30-minute time steps. Site PFT was determined by matching site vegetation (fluxnet.ornl.gov) to CABLE PFTs. PFT parameters were taken from a standard look-up table provided with CABLE 2.0 and were not calibrated to match site characteristics. The model was run using two alternative hydrological modules, two stomatal conductance parameterisations, three soil types and three LAI time series. The new hydrological scheme implements a topographic slope parameter, which was varied between two values in additional simulations. This parameter controls the drainage rate and can in principle be constrained from high-resolution elevation data. We vary the slope parameter between 0 and 5 degrees, broadly coinciding with the observed range of 0-6 degrees at the flux sites as derived from the approximately 1 km spatial resolution Global 30-Arc Second Elevation (GTOPO30) elevation dataset (https://lta.cr.usgs.gov/GTOPO30).

CABLE was run with all parameterisation combinations, resulting in 18 simulations using the default and 36 using the new hydrological scheme. This enabled the quantification of individual parameter and/or parameterisation uncertainties on model simulations and accounts for interactions between different parameterisations. The individual model parameterisations varied in this study are detailed below.

2.2.2 Hydrological parameterisation

We use two different representations of hydrology. The two schemes are fully detailed in Decker (2015) but we will briefly describe the main differences here. The default soil hydrological scheme in CABLE simulates the exchange of water and heat based on six soil layers and up to three snow layers. The default parameterisation for soil moisture processes was developed by Kowalczyk et al. (1994) and later revised by Gordon et al. (2002) and is described in detail in Kowalczyk et al. (2006) and Wang et al. (2011). The default scheme only generates infiltration excess surface runoff when the top three soil layers are ≥95% saturated and otherwise lacks an explicit runoff generation scheme ~~(Decker, in review). It does not distinguish between saturated and un-saturated top soil fractions or simulate groundwater dynamics. The default version of CABLE tends to overestimate Q_E at annual to seasonal scales when used coupled with the ACCESS climate model (Lorenz et al., 2014), but was found to significantly underestimate Q_E during soil moisture deficits across six European flux tower sites (De Kauwe et al., 2015c). It is not known if this is the result of the hydrological parameterisation and we explore this here.~~ (Decker, 2015). It does not simulate saturated and un-saturated top soil fractions separately or consider groundwater aquifer storage. The default scheme solves the vertical redistribution of soil water using the 1-D Richards equation. The bottom boundary condition for the solution of the 1-D Richards equation is given as

$$q_{sub} = C_{drain}\theta_n \qquad (1)$$

where q_{sub} is the subsurface drainage (mm s^{-1}), θ_n the soil moisture content of the bottom soil layer (mm^3 mm^{-3}) and C_{drain} a tunable parameter (mm s^{-1}) (Decker, 2015). The scheme thus assumes a free draining lower boundary and water below the model domain (e.g. groundwater) cannot recharge the water content of the above soil column.

Soil evaporation (E_{soil}; W m^{-2}) is given by

$$E_{soil} = \beta_s \frac{L_v \rho_a (q_{sat}(T_{srf}) - q_a)}{r_g} \tag{2}$$

where L_v is the latent heat of vaporisation (J kg^{-1}), ρ_a the air density (kg m^{-3}), $q_{sat}(T_{srf})$ the saturated specific humidity at the surface temperature (kg kg^{-1}), q_a the specific humidity of air (kg kg^{-1}) and r_g (s m^{-1}) the aerodynamic resistance term. β_s is a dimensionless scalar (varying between 0 and 1) used to reduce E_{soil} when soil moisture is limiting and is given by the linear function

$$\beta_s = \frac{\theta_1 - 0.5\theta_w}{\theta_{fc} - 0.5\theta_w} \tag{3}$$

where θ_1 the soil moisture content of the first soil layer (m^3 m^{-3}), θ_w the wilting point (m^3 m^{-3}) and θ_{fc} the field capacity (m^3 m^{-3}).

Decker (2015) developed an improved representation of sub-surface hydrological processes similar to that implemented in the Community Land Model (Lawrence and Chase, 2007; Oleson et al., 2008). The new scheme explicitly simulates saturation- and infiltration-excess runoff generation and has a dynamic groundwater component with aquifer water storage. The scheme solves the vertical redistribution of soil water (θ) following the Modified Richards Equation (Zeng and Decker, 2009):

$$\frac{\partial \theta}{\partial t} = -\frac{\partial}{\partial z} K \frac{\partial}{\partial z} (\Psi - \Psi_E) - F_{soil} \tag{4}$$

where K (mm s^{-1}) is the hydraulic conductivity, Ψ (mm) the soil matric potential, Ψ_E (mm) the equilibrium soil matric potential, z is soil depth (mm) and F_{soil} (mm s^{-1}) is the sum of subsurface runoff and transpiration (Decker, 2015). An unconfined aquifer is located below the six soil layers and is presented with a simple water balance model:

$$\frac{dW_{aq}}{dt} = q_{re} - q_{aq,sub} \qquad (5)$$

where W_{aq} is the mass of water in the aquifer (mm), $q_{aq,sub}$ the subsurface runoff removed from aquifer (mm s^{-1}) and q_{re} the water flux between the aquifer and the bottom soil layer, given from Darcy's law as

$$q_{re} = K_{aq} \frac{(\Psi_{aq} - \Psi_n) - (\Psi_{E,aq} - \Psi_{E,n})}{z_{wtd} - z_n} \qquad (6)$$

where z_{wtd} is the water table depth (mm), K_{aq} is the hydraulic conductivity of the aquifer (mm s^{-1}), and z_n is the depth of the lowest model layer (mm). The bottom boundary condition is given as

$$q_{out} = 0 \qquad (7)$$

as the scheme assumes that the groundwater aquifer sits above an impermeable layer of rock (Decker, 2015). Subsurface runoff (q_{sub}) is calculated from

$$q_{sub} = \sin\frac{dz}{dl}\hat{q}_{sub} e^{-\frac{z_{wtd}}{f_p}} \qquad (8)$$

where dz/dl is the mean slope, $\hat{q}_{sub}$ the maximum rate of subsurface drainage for a fully saturated soil column (mm s^{-1}) and f_p is a tunable parameter (Decker, 2015). q_{sub} is removed from the bottom three soil layers (which account for 4.366 m of the total soil thickness of 4.6 m) by weighting the amount of water removed from each layer based on the mass of liquid water present in each layer (Decker, 2015).

Sub-grid scale heterogeneity in soil moisture is permitted and a modified soil evaporation formulation reflects this. At point scales the runoff generation from sub-grid heterogeneity in soil moisture is neglected as the saturated fraction of the grid cell is assumed to be equal to zero. Soil evaporation is given as (here assuming a saturated grid cell fraction (F_{sat}) of 0):

$$E_{soil} = F_{sat} E^*_{soil} + (1 - F_{sat}) \beta_s E^*_{soil} \qquad (9)$$

where E^*_{soil} is the soil evaporation prior to soil moisture limitation (mm s^{-1}). β_s is calculated using a non-linear function following Sakaguchi and Zeng (2009):

$$\beta_s = 0.25\left(1 - \cos\left(\pi \frac{\theta_{unsat}}{\theta_{fc}}\right)\right)^2 \qquad (10)$$

where θ_{unsat} is the first layer soil moisture content in the unsaturated portion (the entire soil layer in this study) (m^3 m^{-3}).

Both schemes calculate transpiration using the same method and, in common with many LSMs (Verhoef and Egea, 2014), limit gas exchange during low soil moisture using a dimensionless scalar (β) varying between 0 and 1:

$$\beta = \sum_{i=1}^{n} f_{root,i} \frac{\theta_i - \theta_w}{\theta_{fc} - \theta_w} \qquad (11)$$

where θ_i is the soil moisture content of soil layer *i* (m^3 m^{-3}) and $f_{root,i}$ the root mass fraction of soil layer *i*.

The default version of CABLE tends to overestimate Q_E at annual to seasonal scales when used coupled with the ACCESS climate model (Lorenz et al., 2014), but significantly underestimates Q_E during soil moisture deficits across six European flux tower sites (De Kauwe et al., 2015b) in uncoupled experiments. Decker (2015) showed that the new model reduced overestimations of Q_E by 50–70% compared to the default scheme and yielded an improved simulation of seasonal cycles when evaluated against observations from large river basins. The new scheme was also shown to better capture total water storage anomalies (an integral over depth of soil moisture changes) from the Gravity Recovery and Climate Experiment (GRACE; http://grace.jpl.nasa.gov) than the default scheme. It is not known if these improvements will also allow CABLE to better capture observed Q_E during dry-down and we explore this here.

2.2.3 Stomatal conductance parameterisation

We use two alternative parameterisations for stomatal conductance (g_s). The default CABLE currently implements an empirical g_s formulation following Leuning (1995):

$$g_s = g_0 + \frac{\alpha_1 \beta A}{(C_s - \Gamma)\left(1 + \frac{D}{D_0}\right)} \qquad (12)$$

where A is the net assimilation rate (μmol m^{-2} s^{-1}), Γ (μmol mol^{-1}) is the CO_2 compensation point of photosynthesis, C_s (μmol mol^{-1}) and D (kPa) are the CO_2 concentration and the vapour pressure deficit at the leaf surface, respectively. g_0 (mol m^{-2} s^{-1}), D_0 (kPa) and α_1 are fitted constants representing the residual stomatal conductance when A=0, the sensitivity of stomatal conductance to D and the sensitivity of stomatal conductance to assimilation, respectively. Although the g_s formulation following Leuning (1995) (or equivalent Ball-Berry model; Ball et al., 1987) are widely used in LSMs, the model parameters are empirical. Thus, we cannot attach any theoretical distinction as to how parameters vary across data sets or among species (De Kauwe et al., 2015a; Medlyn et al., 2011). Consequently, as is common with many LSMs (e.g. Community Land Model version 4.5 (Oleson et al., 2013) and the ORganizing Carbon and Hydrology in Dynamic EcosystEms model (Krinner et al., 2005)), the default scheme only varies stomatal conductance parameters between photosynthetic pathways (C_3 vs. C_4), rather than among PFTs.

~~At point scales the runoff generation from sub-grid heterogeneity in soil moisture is neglected as the saturated fraction of the grid cell is assumed to be equal to zero. This parameter controls the drainage rate and can in principle be constrained from high-resolution elevation data. We vary the slope parameter between 0 and 5 degrees, broadly coinciding with the observed range of 0-6 degrees at the flux sites as derived from the approximately 1 km spatial resolution Global 30-Arc Second Elevation (GTOPO30) elevation dataset (https://lta.cr.usgs.gov/GTOPO30).~~

~~Leuning (1995). The Leuning model and similar empirical schemes (e.g. Ball et al., 1987) are widely used in LSMs but due to the empirical nature of these models we cannot attach any theoretical distinction to parameters across data sets or among species (De Kauwe et al., 2015b; Medlyn et al., 2011). Consequently, as is common with many LSMs, the default scheme only varies stomatal conductance parameters between photosynthetic pathways (C_3 vs. C_4), rather than among different PFTs.~~

As an alternative, we also ran CABLE using the g_s model following Medlyn et al. (2011), a theoretical formulation based on the premise of optimal stomatal behaviour~~.~~ ~~In contrast to the default scheme, g_s parameters have been derived~~:

$$g_s = g_0 + 1.6\left(1 + \frac{g_1 \beta}{\sqrt{D}}\right)\frac{A}{C_s} \qquad (13)$$

where g_1 ($\mathrm{kPa}^{0.5}$) is a fitted parameter representing the sensitivity of conductance to the assimilation rate. Unlike the a_1 parameter in the Leuning model, g_1 has biological meaning, representing a plant's water use strategy. Values of g_1 were derived previously for each of CABLE's PFTs (De Kauwe et al., 2015a) based on a global synthesis of stomatal behaviour (Lin et al., 2015). Further details and associated parameter values can be found in De Kauwe et al. (2015a).

The Medlyn g_s model has been shown to improve existing CABLE biases, particularly overestimations of Q_E in evergreen needleleaf and C_4 biomes ~~(De Kauwe et al., 2015a). The Leuning and Medlyn g_s models, as implemented in CABLE, are fully detailed in De Kauwe et al. (2015a).~~ (De Kauwe et al., 2015a). We will explore if the re-parameterisation of g_s also improves the simulation of dry-down in CABLE.

2.2.4 Soil parameterisation

The soil parameters were derived from a dataset provided by CABLE developers (https://trac.nci.org.au/trac/cable/wiki; ~~Global Soil Data Task Group, 2000).~~Global Soil Data Task Group, 2000; Zobler, 1999). The dataset consists of nine soil classes; here the two classes with the highest sand and clay contents were used. The coarse sandy soil has a ~~87~~83% sand content and the fine clay soil a ~~66~~67% clay content~~; the~~. The soil classes have eight associated parameters for soil ~~water holding~~hydraulic and thermal capacities, fully detailed in Table S2. In addition, an arbitrary "medium" soil class was created with equal fractions of sand, silt and clay, with other soil parameters set as the median of the coarse sand and fine clay soil classes (Table S2). CABLE was run with these three alternative soil classes, fixing the soil parameters across all sites~~. The default hydrological scheme uses these three soil parameter sets directly, whereas the new scheme employs an empirical approach to calculate the parameters governing water holding and thermal capacities from sand, silt and clay fractions. The~~ to generate a range in soil parameter values. The default hydrological scheme uses all soil parameters directly, whereas the new scheme calculates the eight parameters governing hydraulic properties from sand, silt and clay fractions using the Clapp and Hornberger (1978) pedotransfer functions. The soil parameter values used by both schemes are detailed in Table S2.

2.2.5 Leaf Area Index

Leaf area index (LAI) plays an important role in the surface energy balance in CABLE~~ (Kala et al., 2014).~~, scaling sunlit and shaded leaf-level fluxes of photosynthesis, g_s and latent heat flux to the canopy. LAI was obtained from 8-daily gridded Moderate Resolution Imaging Spectroradiometer (MODIS) data at 1 km resolution (Yuan et al., 2011). The data were averaged to monthly time steps to smooth the time series and subsequently three alternative LAI time series were created for each site to take some account of uncertainties in LAI inputs. The first time series was constructed by extracting the grid cells that contained each site ("centre" time series). Two alternative time series were created using the minimum and maximum LAI value of the grid cell and its immediate neighbours ("minimum" and "maximum" time series, respectively). Time-varying LAI was used for years where the flux observations and MODIS data overlap (i.e. after 2000); a monthly climatology of common years was used otherwise. The minimum and maximum time series differ from the centre time series by 30% on average but the range varies between sites. The alternative LAI time series are ~~plotted~~shown in Fig. S5.

2.3 Analysis methods

We analyse ~~CABLE~~CABLE's performance across three time scales: the whole observational, annual and sub-annual periods. As the observational records are generally short for characterising hydrological extremes ~~(≤10 years,~~ (~5 years on average; Table S1), we have not adopted a formal statistical method for identifying periods of rainfall anomalies and thus do not refer to them as "droughts". We also note no one definition for droughts exists; instead, various indices have been employed based on, for example, precipitation, streamflow~~ and/or~~, soil moisture and measures of evaporative demand (Sheffield and Wood, 2011). In this study, the dry periods were defined based on precipitation as this allowed the use of available observations, but we note the simulated fluxes will also depend on other processes such as soil moisture availability. For the majority of sites (Howard Springs, Palang, and all supplementary sites), we selected the year with the lowest precipitation total as the one-year period, whilst for Amplero, Blodgett, Tumbarumba and UniMich, we selected a year when the default CABLE significantly underestimated latent heat fluxes during a rainfall deficit ("dry-down") period. The dry-down period generally coincides with the maximum and the following minimum observed latent heat flux during the one-year period but has been adjusted using expert judgment for some sites to best ~~capture~~demonstrate typical model behaviour (Fig. S6). Observed and simulated data were averaged to 14-day running means for all analyses.

~~We follow PALS for calculating model metrics. We use the normalised mean error (NME) to evaluate general model performance:~~

$$\cancel{NME = \frac{\sum_{i=1}^{n}\left|M_i - O_i\right|}{\sum_{i=1}^{n}\left|\bar{O} - O_i\right|} \qquad (1)}$$

~~where M represents the model values and O the observations. NME accounts for mean model biases and the temporal coincidence and magnitude of variability, but does not distinguish between them (Best et al., 2015). An NME of 0.0 represents perfect agreement and a value of 1.0 represents model performance equal to that expected from a constant value equal to the mean of all observations.~~

~~We examine mean bias error (MBE) to estimate absolute biases in CABLE simulations; it is simply the difference between the mean modelled and observed values:~~

~~$MBE = \overline{M} - \overline{O}$ (2)~~

We follow Abramowitz et al. (2007) and the PALS protocol for calculating model metrics. We use the normalised mean error (NME) to evaluate general model performance:

$$NME = \frac{\sum_{i=1}^{n} |M_i - O_i|}{\sum_{i=1}^{n} |\overline{O} - O_i|} \qquad (14)$$

where *M* represents the model values and *O* the observations. NME accounts for mean model biases and the temporal coincidence and magnitude of variability, but does not distinguish between them (Best et al., 2015). An NME of 0.0 represents perfect agreement and a value of 1.0 represents model performance equal to that expected from a constant value equal to the mean of all observations.

We examine mean bias error (MBE) to estimate absolute biases in CABLE simulations; it is simply the difference between the mean modelled and observed values:

$$MBE = \overline{M} - \overline{O} \qquad (15)$$

3 Results

3.1 Whole time period

We first evaluated ~~CABLE~~ Q_E simulated ~~Q_E~~by CABLE during the whole data period available for each flux site (ranging from 2 to 7 years for selected sites; Table S1). CABLE, using the default hydrological parameterisation, captures the general features, such as the timing and magnitude of seasonal cycles, in observed Q_E across the different sites (Fig. 2, left column). CABLE including the new hydrological scheme also captures these general features (Fig. 2, right column). Quantifying the performance of these two versions of CABLE over the full length of record does not indicate that there is a significant difference between the versions in either NME (Fig. 3) or MBE (Fig. 4). The average NME for all sites and parameter choices was 0.90 for the old scheme and 0.75 for the new scheme, and the average MBE was −1 and 6 W m^{-2}, respectively. The NME metric is <1.0 for the majority of sites using the new scheme, regardless of the choice of g_s, LAI or soil parameterisation. We note that the magnitude of Q_E for the evergreen broadleaf sites (Palang and Tumbarumba and supplementary site Espirra) is poorly captured (Fig. 4). Overall, both hydrological parameterisations tend to overestimate peak Q_E (Fig. 2); this tendency for excessive evapotranspiration has also been demonstrated in global applications of CABLE in both offline (De Kauwe et al., 2015a) and coupled (Lorenz et al., 2014) simulations. Furthermore, both schemes systematically overestimate Q_E in spring, particularly at cooler temperate sites such as UniMich (Fig. 2; also see deciduous broadleaf and needleleaf supplementary sites (Fig. S1)), and over-predict the short-term variability in Q_E (see e.g. Amplero in Fig. 2). ~~This is likely due to overly rapid drying of top soil layers, which strongly control Q_E in CABLE (De Kauwe et al., 2015c). This is particularly evident during warmer summer months when fluxes are more strongly moisture-limited (see e.g. Tumbarumba and UniMich in Fig. 2).~~ Despite clear biases in simulated fluxes, the MBE metric approaches zero at most sites when evaluated at inter-annual time scales. While encouraging, this is ~~likely~~ due to compensating errors, such that early season overestimations in Q_E are counteracted by underestimations during the dry-down periods~~.~~ (see e.g. Blodgett and Tumbarumba in Figure 5). This is particularly evident with the default hydrology scheme. We therefore focus the remaining analyses on shorter time periods where compensating biases are less likely to hide weaknesses in the model performance.

3.2 Annual and dry-down period

CABLE-simulated Q_E ~~were~~was then evaluated during annual and seasonal dry-down periods to explore model performance during rainfall deficits. The default scheme demonstrates a range of major biases (Fig. 5). The model dries down too quickly at Amplero, Blodgett, Palang, Tumbarumba and UniMich sites. At these sites, and at Howard Springs, Q_E drops too low and drops to that minimum too early in the year. At several sites, including Blodgett, Tumbarumba and UniMich, CABLE systematically overestimates Q_E in spring. These characteristics of CABLE are not dependent on the choice of LAI~~, g_s,~~ or soil ~~parameters~~inputs or g_s parameterisations; the range in Q_E fails to overlap the observations irrespective of how these properties are varied. This suggests parameterisation error as distinct from parameter choices as the cause of the model weaknesses.

The new hydrological scheme demonstrates clear improvements at Amplero, Howard Springs and Palang. At Blodgett, Tumbarumba and UniMich, the observations are within the uncertainty due to the choice of g_s, LAI or soil parameters in the second half of the year, but the excessive Q_E during spring and early summer remains a problem. While there are obviously remaining errors, the new hydrological scheme clearly improves the simulation of Q_E over the annual cycles ~~shown in~~ (Fig. 5~~.~~). Assessing the overall performance at annual time scales also highlights clear improvements with the new hydrology. Fig. 3 shows that for NME, the new hydrology scheme in CABLE performs as well as, or better than the default at every site, with an average NME across all sites of 0.68 compared to 0.90 for the default scheme. This is true also of MBE (Fig. 4) for all sites except Tumbarumba.

Assessing the performance of the two schemes over the dry-down period using NME is shown in Fig. 5. Using the default hydrology leads to worse performance on this shorter time scale at Amplero, Blodgett, Palang and to a lesser degree at Howard Springs and Tumbarumba compared to annual and inter-annual scales. In contrast, CABLE with the new hydrology performs similarly~~ well~~ to the longer (≥ 1 year) time scales at Blodgett, Palang, Howard Springs, Tumbarumba and UniMich and only marginally poorer at Amplero. Comparing NME over this dry down period shows that the new scheme strongly outperforms the default parameterisation (Fig. 3; the average NME is 0.68 and 1.27 for new and default schemes, respectively). A similar conclusion is reached using MBE (on average −4 and −22 W m^{-2} for the new and default schemes, respectively). In short, the new hydrology does not dramatically improve the performance of CABLE on the long term (i.e. inter-annual scales) (Fig. 2) due to compensating biases in the default CABLE. These include overestimated spring and early summer Q_E, and consequently, at least in part, underestimated Q_E during the dry-down. Once we focus on shorter, sub-annual timescales that lack these compensating biases, CABLE with the new hydrology strongly outperforms the default version in the simulation of Q_E.

3.3 Impact of varying LAI, g_s and soil parameters

We now explore the individual contributions from soil parameters, g_s and LAI to uncertainties in simulated Q_E. Fig. 6 and Fig. 7 show the uncertainty in model simulations due to soil parameters, g_s and LAI using the default and new hydrological schemes, respectively. Both hydrological schemes are sensitive to soil parameters during the dry-down period but show smaller variations due to soil during other parts of the year (see Amplero, Blodgett, Howard Springs and Palang in Fig. 6 and Fig. 7). This transition from low to high sensitivity occurs as soil moisture stores begin to deplete and Q_E becomes increasingly limited by moisture supply~~. The new hydrological scheme uses a narrower range of parameter values for water holding capacity and conductivity (Table S2) and thus results in a smaller range of uncertainty due to soil parameters.~~ (Figure S8). Both schemes show a similar sensitivity to g_s and LAI variations, which is generally smaller compared to soil variations, although the new scheme is more sensitive to g_s at Blodgett, Howard Springs and Palang, and to LAI in Amplero and Palang during dry-down.

While the new hydrological parameterisation systematically improved model performance across most sites (Fig. 3 and 4), the effect of LAI, g_s and soil parameters on the mean magnitude of simulated fluxes is highly site-specific during the annual and dry-down periods (Fig. 8). In agreement with (De Kauwe et al., 2015a), the choice of g_s scheme generally has a larger effect in needleleaf (Blodgett) and C$_4$ grass (Howard Springs) sites. Some sites, such as Howard Springs, are sensitive to multiple parameters, whilst others such as UniMich only respond minimally to parameter perturbations (Fig. 8). Whilst there is no *a priori* expectation that this should be the case, it highlights the importance of investigating model uncertainties and performance across multiple sites to capture full range of model sensitivities to parameter perturbations.

The results have so far assessed CABLE ~~when incorporating~~with the new hydrology using a 0° slope parameter ~~because~~as this enables a direct comparison with the default hydrology. The slope parameter, which can be derived from high-resolution elevation data, is scale dependent and was introduced by Decker (2015) to capture large-scale hydrological processes that are affected by landscape geometry. The slope parameter affects the rate of subsurface drainage and represents a key difference between the new and default schemes. With the exception of the UniMich site, Fig. 8 and 9 show that the model is highly sensitive to choice of the slope parameter across all sites, particularly during the dry-down period. The slope appears more critical for simulation of Q_E than the other parameterisations investigated here and has a strong effect on the magnitude of fluxes primarily during the dry-down (see e.g. Howard Springs and Palang in Fig. 9). Whilst this highlights the need to carefully set the slope parameter, it is unclear how well it can be constrained at the site scale. The surface slope derived from elevation data may not reflect large-scale features, such as subsurface geology, which can affect drainage rates and thus water availability for Q_E in highly site-specific ways.

4 Discussion

4.1 Simulation of dry-down

We have shown that the default version of CABLE significantly underestimates Q_E during rainfall deficits~~.~~ (Fig. 5). We have also shown that it is unlikely that uncertainties in key model soil and vegetation (LAI) inputs~~, account for these biases.~~ account for these biases (Fig. 6). The observations used to drive and evaluate the model themselves include errors, notably lack of energy balance closure (Leuning et al., 2012). However, given the systematic and large seasonal biases in CABLE simulations, any errors in the flux tower measurements are unlikely to explain poor model performance during dry-down. Instead, our results point to deficiencies in the representation of hydrological processes in the default version of CABLE~~.~~ (Fig. 3 and

4). The default CABLE has been shown to perform similarly to other LSMs in Best et al. (2015) and indeed in other model evaluation studies (Abramowitz et al., 2008). Hence, it is likely that the errors of the kind identified here may be common among other models~~.~~, as model benchmarking rarely examine sub-annual behaviour. The poor simulation of dry-down periods is important: if LSMs in general struggle to simulate ~~the dry-down~~this period they will fail to correctly capture water fluxes when serious soil moisture deficits are established. A model that ~~dries~~ simulates dry-down too fast will enter drought early and will tend to simulate longer, deeper and more frequent droughts than a model that ~~dries down~~enters drought too slowly. We suggest that systematic evaluation of LSMs during dry down periods would lead to the identification of major limitations in some models that are hidden by compensating errors over longer timescales. Resolution of those problems has the potential to improve the simulation of drought in climate models.

We also showed that the effect of individual parameterisations was magnified during dry periods~~.~~ (Fig. 6 and 7). Whilst the new hydrological scheme did not ~~have~~present a significant ~~impact~~improvement on the annual and inter-annual timescales analysed here, it had an increasingly large positive impact on shorter time scales and in particular during the dry-down periods~~.~~ (Fig. 3,4 and 5). Similarly, the contribution of ~~LAI,~~soil (Fig. 6a and 7a), g_s ~~and soil~~(Fig. 6b and 7b) and LAI (Fig. 6c and 7c) parameterisations to model uncertainties was generally larger during the dry-down. ~~This highlights the value~~We will discuss each of ~~evaluating model parameterisations against both mean and (more) extreme states. It also~~ these points ~~to the challenge that, in these dry down periods that are critical to how a landscape develops towards drought, the skill needed to capture the relevant processes will be higher than during wet periods~~below.

4.1.1. Soil and LAI inputs

We evaluated the uncertainty in Q_E simulations arising from inputs of soil properties and LAI. These variables are generally obtained from gridded datasets in LSM simulations and remain uncertain at the site (and larger) scale (De Kauwe et al., 2011; Koster et al., 2009).

Soil parameters feature in many hydrological model components and our results show that the range in simulated Q_E due to the choice of soil parameters is largest during the dry-down for both hydrological schemes (Fig. 6 and 7). Parameters for wilting point (θ_w) and field capacity (θ_{fc}) are particularly important during drying conditions as they determine how evapotranspiration is reduced as soil moisture becomes limiting (following Eq. (3), (10) and (11)). The model is also sensitive to value of the matric potential at saturation (Table 2). The vertical diffusive flux of soil water between adjacent soil layers is proportional to the saturated matric potential. This control on the rate of vertical water movement alters the profile of vertical soil water during the dry-down impacting the water available for transpiration in a given soil layer. Using the default hydrological scheme, the observed Q_E could only be captured by varying the soil properties at Howard Springs. Elsewhere, the model underestimated observed Q_E during dry-down regardless of how the soil properties were varied (Fig. 6). This suggests that uncertainties in soil parameters cannot account for the poor simulation of dry-down by the default model.

Similarly, LAI, as it was varied here, could not explain the underestimation of Q_E. The range in LAI varied by site (30% on average) according to the remotely sensed data but was not lower at the drought sites. The model was generally not sensitive to changes in LAI during dry-down regardless of the choice of hydrological scheme (Fig. 6 and Fig. 7). This implies that the correct characterisation of canopy structure is probably not critical for the simulation of Q_E in CABLE during dry-down or that the scale of the errors in the simulations are too large to see any more subtle impact of these LAI variations. Nevertheless, we do note that leaf drop during drought events could lead to an increased or compensatory reflectance signal from deeper in the canopy profile, resulting in erroneous estimates of LAI from optically remote sensed products (cf. Amazon drought studies; Samanta et al., 2010).

4.1.2. Hydrological schemes

The new hydrological scheme was shown to improve CABLE simulations of Q_E during dry-down (Fig. 3 and 4). This results from higher soil moisture content simulated by the new scheme compared to the default model, particularly in the bottom soil layers (Fig. S7). This allows higher ET fluxes to be maintained during dry periods, mainly due to higher transpiration rates (Fig. S8).

The alternative hydrological schemes make different assumptions about subsurface drainage and how this is treated upon exiting the bottom soil column boundary. The default model assumes a free draining boundary for solving vertical water flow (Eq. (1)), such that the bottom soil layer essentially acts as a sink for the rest of the soil column as it can only remove water from the column. Conversely, the new scheme simulated an unconfined groundwater aquifer below the bottom soil layer that is assumed to sit on an impermeable layer of rock so that no water is lost from the aquifer through downward flow (Eq. (7)). Soil moisture content of above soil layers can then be replenished through recharge from the aquifer to maintain higher soil moisture during dry periods (given a water table depth near the bottom of the soil column; Eq. (6); Fig. S7). Zeng and Decker (2009) demonstrated that assuming a free draining lower boundary requires an unrealistically high precipitation rate to maintain a relatively wet soil moisture content that allows vegetation to transpire without encountering soil moisture stress. Using a hypothetical example, the authors estimate that a minimum precipitation rate of 17.2 mm/day is required to maintain non-water-stressed conditions (a value much higher than is observed in most environments), implying overly dry soil conditions in many cases. Our results therefore suggest that the replacement of a constant drainage assumption in the original model with a physically based, dynamic bottom boundary condition for the soil column is important for improving Q_E fluxes in CABLE under water stressed conditions.

Whilst it was not possible to evaluate these simulations against soil moisture data due to a lack of observations for soil depths used in CABLE, Decker (2015) showed that the new scheme could better capture total soil column water anomalies and evapotranspiration (two variables that strongly depend on the correct simulation of soil moisture content) in comparison to the default scheme at river basin scales. This gives us confidence that the higher soil moisture levels simulated by the new scheme are supported by some observations. This result should be evaluated in future work against locations where deep soil moisture measurements are made available, or efforts to obtain observed soil moisture coincident with tower measurements of the fluxes should be encouraged.

4.1.3. Stomatal conductance schemes

Our results showed that CABLE is generally not sensitive to the choice of g_s scheme during dry-down at most sites (Fig. 6, 7 and 8), with the exception of Howard Springs (a C_4 grass site) and Blodgett (an evergreen needleleaf site). This result is largely to be expected: during drought both schemes are limited in the same fashion, with β (Eq. (11)) reducing the slope that relates g_s to photosynthesis. The noted differences between schemes at the C_4 grass and evergreen needleleaf sites are consistent with results from De Kauwe et al. (2015a). At Howard Springs, De Kauwe et al. (2015a) found that the high g_0 value assumed in the Leuning model (0.04 mol m^{-2} leaf s$_{-1}$) accounted for the difference between schemes when g_s approached zero (for example during a drought). Differences between schemes at Blodgett stem, at least in part, from the use of a parameterisation of a conservative water use, found in evergreen needleleaf forests (Lin et al., 2015).

We note that the two stomatal schemes have different sensitivities to vapour pressure deficit (see De Kauwe et al. (2015a) for details). However, under current climatic conditions this assumption only results in a small difference between schemes, although this effect could be amplified in the future with expected increased in vapour pressure deficit in a warmer world.

4.2 Overestimation of soil evaporation

We identified systematic biases in the simulation of peak and spring Q_E, particularly at forested sites (e.g. Tumbarumba and Blodgett) (Fig. 2 and S7). The biases in the timing and magnitude of spring and peak fluxes not only have implications for the correct simulation of seasonal cycles, but can also affect the magnitude of dry-down simulated by the model. The excessive spring and early summer Q_E may reduce soil moisture levels prior to the dry-down, leading to the simulation of more severe reductions in Q_E during dry periods.~~).~~ Both hydrological schemes showed a tendency to significantly overestimate these fluxes. The reason for the overestimation of peak fluxes is not clear but is not resolved by the new hydrological scheme despite this parameterising many of the relevant processes differently. ~~The~~At many sites, the high

Q_E in spring is associated with excessive soil evaporation and is not linked to transpiration, which closely follow the observed seasonal cycle (~~Fig. S7~~see e.g. Bugac, Harvard, Howland and Hyytiälä in Fig. S9).

~~There are multiple potential causes and solutions to this excessive Q_E.~~There are a number of possible causes and solutions to this excessive soil evaporation. Insufficient drainage, and consequently overestimated surface soil moisture, and/or insufficient reduction of soil evaporation during soil drying may explain the excess spring Q_E. The default scheme uses a linear function to reduce soil evaporation when soil moisture is limiting following Eq. (3). This is replaced with a non-linear function presented in Eq. (10) in the new scheme. The non-linear function provides a much stronger limitation on soil evaporation as soil moisture declines but, based on these results, this approach is not sufficient for resolving the excessive soil evaporation.

Haverd and Cuntz (2010) showed the inclusion of litter layer dynamics in an earlier version of CABLE improved the simulated timing of spring Q_E at Tumbarumba by suppressing soil evaporation but this was not implemented in the current study. Adding a litter layer may resolve excessive soil evaporation at ~~forested~~some sites by adding an additional resistance to evaporation, but ~~is unlikely to~~it is unclear that this approach would resolve errors ~~for other~~at all PFTs. ~~However, before we attempt to implement litter dynamics, we need to be sure that this addition is not simply masking a major deficiency elsewhere in the model. For example, errors~~Errors in the timing of spring green-up at deciduous sites in the LAI inputs (e.g. Fisher and Mustard, 2007) may ~~contribute to excessive spring evaporation, whereby a delayed green-up would allow excessive radiation to reach the ground surface in early spring, increasing soil evaporation rates. Insufficient drainage, and consequently overestimated surface soil moisture, and/or insufficient reduction of soil evaporation during soil drying may also explain the excess spring Q_E. Alternatively, high soil evaporation may result from the simulation of excessive within canopy turbulence (Raupach, 1989a, 1989b, 1994). The biases in the timing and magnitude of spring and peak fluxes not only have implications for the correct simulation of seasonal cycles, but may also affect the magnitude of dry-down simulated by the model.~~ also contribute to excessive spring evaporation, whereby a delayed green-up would allow excessive radiation to reach the ground surface in early spring, increasing soil evaporation rates. We encourage researchers to make use of the Best et al. (2015) experimental protocol to fully explore this problem. Using multi-LSM simulations should~~The excessive spring and early summer Q_E may reduce soil moisture levels prior to the dry-down, leading to the simulation of more severe reductions in Q_E during dry periods.~~

~~Our strategy to resolve the excessive spring soil evaporation is linked with the Best et al. (2015) experimental protocol. Using multi-LSM simulations we hope to~~ be able to identify where CABLE is anomalous, and ideally implement the model parameterisations used in other LSMs that do not simulate excessive spring soil evaporation.

4.3 Further model uncertainties

In this study, we explored and quantified model uncertainties due to LAI, g_s, hydrological and soil parameters, limiting our analysis to parameters that can be derived from observationally based global datasets (despite considerable uncertainties). ~~Other model processes, particularly vegetation response to drought, have been identified as critical for capturing drought processes and shown to improve CABLE performance during droughts but were not explored here. The simulation of the effects of soil moisture limitation on photosynthesis and stomatal conductance remains a key uncertainty for drought responses in LSMs (Zhou et al., 2013). Models rely on differing assumptions about the effects of water stress on photosynthesis and stomatal conductance (Egea et al., 2011; Keenan et al., 2009) but generally assume similar drought responses across different PFTs (including CABLE as employed here) (De Kauwe et al., 2015c; Zhou et al., 2013). De Kauwe et al. (2015b) evaluated CABLE against flux site observations during the 2003 European drought using an alternative drought model with experimentally derived drought sensitivities. The authors similarly showed significant underestimations of Q_E using the default CABLE but these were improved using different plant species sensitivities to drought and a dynamic weighting of water uptake across soil layers. Experimental data to inform the parameterisation of PFT-specific drought responses, however, remains limited (De Kauwe et al., 2015c), complicating the implementation of such responses into LSMs.~~Other model processes, particularly more realistic representations of vegetation drought responses, have been identified as critical for capturing drought processes and shown to improve CABLE performance during droughts but were not explicitly explored here. ~~Li et al. (2012) showed the underestimation of CABLE-simulated Q_E under water-stressed conditions could be improved by employing an alternative root water uptake scheme. The default root water uptake function in CABLE employed here (Wang et al., 2011) assumes a constant efficiency of water uptake per unit root length (Li et al., 2012). CABLE with the alternative scheme, combining a function allowing variable root-density distribution (Lai and Katul, 2000) with a hydraulic redistribution scheme (which allows roots to move water from wetter to drier soil layers), was shown to correctly capture the magnitude of seasonal droughts across three flux tower sites. The implementation of more realistic vegetation responses and adaptations to droughts should further refine the performance of the new hydrological scheme during dry-down periods.~~

The simulation of the effects of soil moisture limitation on photosynthesis and stomatal conductance remains a key uncertainty for drought responses in LSMs (Zhou et al., 2013). Models rely on differing assumptions about the effects of water stress on photosynthesis and stomatal conductance (Egea et al., 2011; Keenan et al., 2009) but generally assume similar drought responses across different PFTs (including CABLE as employed here) despite experimental evidence pointing to systematic differences in plant adaptations to drought (De Kauwe et al., 2015b; Zhou et al., 2013). In common with many other LSMs (Verhoef and Egea, 2014), CABLE limits gas exchange during low soil moisture using the dimensionless scalar β following Eq. (11). The function is strongly linked to soil properties (through wilting point and field capacity parameters) and does not directly consider vegetation characteristics beyond rooting depth (which varies little by PFT). De Kauwe et al. (2015b) evaluated CABLE against flux site observations during the 2003 European drought using an alternative drought model with experimentally derived drought sensitivities. They showed significant underestimations of Q_E using the default CABLE but these were improved using different vegetation sensitivities to drought (varying from low sensitivity in xeric environments to high in mesic environments in line with experimental evidence) and a dynamic weighting of water uptake across soil layers. Experimental data to inform the parameterisation of PFT-specific drought responses, however, remains limited (De Kauwe et al., 2015b), complicating the implementation of such responses into LSMs. Li et al. (2012) showed the underestimation of CABLE-simulated Q_E under water-stressed conditions could be improved by employing an alternative root water uptake scheme. The default root water uptake function in CABLE employed here (Wang et al., 2011) assumes a constant efficiency of water uptake per unit root length (Li et al., 2012). CABLE with the alternative scheme, combining a function allowing variable root-density distribution (Lai and Katul, 2000) with a hydraulic redistribution scheme (which allows roots to move water from wetter to drier soil layers), was shown to correctly capture the magnitude of seasonal~~Furthermore, in current simulations prescribed monthly MODIS LAI was used.~~-scale droughts across three flux tower sites. The implementation of more realistic vegetation responses and adaptations to droughts should further refine the performance of the new hydrological scheme during dry-down periods.

Furthermore, in the simulations described here prescribed monthly MODIS LAI was used. Whilst CABLE and many other LSMs are capable of simulating LAI dynamically, it is common practice, particularly in coupled online simulations, to rely on prescribed monthly climatology instead of time-varying LAI. This limits the realistic simulation of reductions in LAI during severe droughts and consequent feedbacks with radiative and evaporative processes such as interception losses. Canopy defoliation may, for example, decrease transpiration and interception but also increase radiation reaching the soil surface, potentially increasing soil evaporation in the presence of available moisture~~, whilst also decreasing albedo and total ground-reaching radiation.~~. As these feedbacks were not considered in this study, the rate of dry-down may have been overestimated at sites which experienced LAI reductions during rainfall deficits, but which may not have been captured in the MODIS LAI inputs. However, as only the magnitude of LAI was varied in this study, it is not possible to quantify the effects of temporal errors in LAI on simulated Q_E. ~~But as~~Since both hydrological models were forced with identical LAI, it is unlikely uncertainties in the prescribed LAI explain the excessive dry-down in the default hydrological scheme.

We have limited our analysis to short-term, seasonal-scale rainfall deficits. Multi-annual droughts, such as the Millennium drought in eastern Australia (van Dijk et al., 2013), are likely to exhibit different dynamics in terms of vegetation responses and consequent feedbacks with land surface fluxes, soil moisture states and albedo.

Prudhomme et al. (2011), for example, showed the JULES LSM to more successfully reproduce long-term hydrological droughts than short-term events in terms of duration and severity. Realistic representations of plant adaptations to drought and dynamically varying LAI are likely to be increasingly important for representing vegetation resilience and coupled land surface processes during long-term droughts. We therefore suggest future studies of LSM performance under water-stressed conditions should evaluate models against drought events at different temporal scales.

5 Conclusions

This study~~ has~~ evaluated the CABLE land surface model for seasonal-scale precipitation deficits using 20 flux tower sites distributed globally. We varied the soil hydrological and stomatal conductance parameterisations, and the inputs for LAI and soil properties. Our goal was to determine whether CABLE ~~can~~could capture dry-down associated with rainfall deficits as these components of the model are varied, or whether the model lacks the ~~mechanisms~~physical parameterisations to simulate this phenomenon.

On long time scales (annual and above), compensating biases mean that the two versions of CABLE ~~did not perform very differently. However, as our analysis focused on periods of rainfall deficit, a new hydrological parameterisation based on Decker (in review) clearly improved the capability of CABLE to simulate Q_E.~~performed similarly. However, as our analysis focused more on periods of rainfall deficit, a new hydrological parameterisation based on Decker (2015) clearly improved the capability of CABLE to simulate Q_E. However, neither version of CABLE, and no reasonable choice of soil parameter, LAI or stomatal conductance resolved systematic seasonal-scale biases in excessive spring soil evaporation. The reasons for these biases cannot be determined in isolation and we will next pursue these model limitations using the PLUMBER multi-model benchmarking framework (Best et al., 2015).

Our study highlights some opportunities for land modellers. First, our study again demonstrates the value in freely-available flux tower data for identifying systematic biases in LSMs. The value of these data extends well beyond their common use in evaluating means or seasonal cycles. Second, a major role for LSMs is to simulate feedbacks to the atmosphere associated with rainfall deficits. We have demonstrated that there is skill in CABLE in simulating these feedbacks as a landscape dries but clearly more work needs to be invested in capturing all the elements of a drying soil and its impacts on Q_E. While the parameterisation of hydrology has been explored over the years, we remind the community that there are on-going challenges in modelling soil moisture and links between soil moisture and evaporation that are not yet resolved. Third, we note that CABLE performs ~~comparatively~~reasonably relative to other LSMs (Abramowitz et al., 2007; Best et al., 2015) and yet when we interrogate the model's performance at timescales when compensating biases are limited, CABLE displays some concerning behaviour. It is inevitable that other LSMs, if examined using these periods of precipitation deficit, will also exhibit problems. Clearly, formally testing LSMs against more extreme conditions, and in the context of specific phenomenon (e.g. drought or heatwave) is a necessary step to build confidence in the projections from climate models that utilise LSMs.

Acknowledgements

We acknowledge the support of the Australian Research Council Centre of Excellence for Climate System Science (CE110001028). M.D.K was supported by the Australian Research Council Linkage grant LP140100232. This work used eddy covariance data acquired by the FLUXNET community and in particular by the following networks: AmeriFlux (U.S. Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (DE-FG02-04ER63917 and DE-FG02-04ER63911)), AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada (supported by CFCAS, NSERC, BIOCAP, Environment Canada, and NRCan), GreenGrass, KoFlux, LBA, NECC, OzFlux, TCOS-Siberia, USCCC. We acknowledge the financial support to the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, Max Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and the database development and technical support from Berkeley Water Center, Lawrence

Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California - Berkeley, University of Virginia.

References

Abramowitz, G.: Towards a public, standardized, diagnostic benchmarking system for land surface models, Geosci. Model Dev., 5, 819–827, doi:10.5194/gmd-5-819-2012, 2012.

Abramowitz, G., Pitman, A., Gupta, H., Kowalczyk, E. and Wang, Y.: Systematic Bias in Land Surface Models, J. Hydrometeorol., 8, 989–1001, doi:10.1175/JHM628.1, 2007.

Abramowitz, G., Leuning, R., Clark, M. and Pitman, A.: Evaluating the performance of land surface Models, J. Clim., 21, 5468–5481, doi:10.1175/2008JCLI2378.1, 2008.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H. (Ted), Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J. H., Allard, G., Running, S. W., Semerci, A. and Cobb, N.: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, For. Ecol. Manage., 259, 660–684, doi:10.1016/j.foreco.2009.09.001, 2010.

Ball, T., Woodrow, I. and Berry, J.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in Progress in ~~photosynthesis research~~Photosynthesis Research, edited by J. Biggins, pp. 221–224, Martinus Nijhoff, Leiden, the Netherlands., 1987.

Best, M. J., ~~Pryor, M., Clark, D. B., Rooner, G. G., Essery, R. L. H., Menard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B. and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description-Part 1: energy and water fluxes, Geosci. Model Dev., 4, 677–699, 2011.~~

~~Best, M. J.,~~ Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. ~~a~~A., Dong, J., Ek, M., Guo, Z., Haverd, V., van den Hurk, B. J~~.~~., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. ~~a~~A., Stevens, L. and Vuichard, N.: The plumbing of land surface models: benchmarking model performance, J. Hydrometeorol., 16, 1425–~~1422~~1442, doi:10.1175/JHM-D-14-0158.1, 2015.

Bi, D., Dix, M., Marsland, S. J., O'Farrell, S., Rashid, H. a, Uotila, P., Hirst, A. C., Kowalczyk, E., Golebiewski, M., Sullivan, A., Yan, H., Hannah, N., Franklin, C., Sun, Z., Vohralik, P., Watterson, I., Zhou, X., Fiedler, R., Collier, M., Ma, Y., Noonan, J., Stevens, L., Uhe, P., Zhu, H., Griffies, S. M., Hill, R., Harris, C., Puri, K. and Farrell, S. O.: The ACCESS coupled model: description, control climate and evaluation, Aust. Meteorol. Oceanogr. J., 63, 41–64, 2013.

Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M. and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, Geosci. Model Dev., 4, 255–269, doi:10.5194/gmd-4-255-2011, 2011.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14, 601–604, 1978.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J. L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J. and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, pp. 1029–1136, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.

Cruz, F. T., Pitman, A. J., McGregor, J. L. and Evans, J. P.: Contrasting Regional Responses to Increasing Leaf-Level Atmospheric Carbon Dioxide over Australia, J. Hydrometeorol., 11, 296–314, doi:10.1175/2009jhm1175.1, 2010.

Dai, A.: ~~Drought under global warming: a review, Wiley Interdiscip. Rev. Clim. Chang., 2, 45–65, doi:10.1002/wcc.81, 2011.~~

~~Dai, A.:~~ Increasing drought under global warming in observations and models, Nat. Clim. Chang., 3, 52–58, doi:10.1038/nclimate1633, 2012.

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~~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, ~~2015b~~2015a.

De Kauwe, M. G., Zhou, S.-X., Medlyn, B. E., Pitman, ~~a~~A. J., Wang, Y~~.-~~. P., Duursma, R. ~~a~~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 ~~Discuss.,~~, 12, 12349–12393, doi:10.5194/bgd-12-12349-2015, ~~2015c~~2015b.

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

Dirmeyer, P. A.: A ~~History~~history and ~~Review~~review of the Global Soil Wetness Project (GSWP), J. Hydrometeorol., 12, 729–749, doi:10.1175/JHM-D-10-05010.1, 2011.

Egea, G., Verhoef, A. and Vidale, P. L.: Towards an improved and more flexible representation of water stress in coupled photosynthesis-stomatal conductance models, Agric. For. Meteorol., 151, 1370–1384, doi:10.1016/j.agrformet.2011.05.019, 2011.

Fisher, J. I. and Mustard, J. F.: Cross-scalar satellite phenology from ground, Landsat, and MODIS data, Remote Sens. Environ., 109, 261–273, doi:10.1016/j.rse.2007.01.004, 2007.

Global Soil Data Task Group: Global gridded surfaces of selected soil characteristics (IGBP-IDS), [online] Available from: http://www.daac.ornl.gov, 2000.

Gordon, H. B., Rotstayn, L. D., McGregor, J. L., Dix, J. L., Kowalczyk, E. A., O'Farrell, S. P., Waterman, L. J., Hirst, A. C., Wilson, S. G., Collier, M. A., Watterson, I. G. and Elliott, T. I.: The CSIRO Mk3 climate system model, Aspendale, Victoria, Australia. [online] Available from: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+CSIRO+Mk3+Climate+System+Model#0, 2002.

~~Gupta, H. V., Bastidas, L. a., Sorooshian, S., Shuttleworth, W. J. and Yang, Z. L.: Parameter estimation of a land surface scheme using multicriteria methods, J. Geophys. Res., 104, 19491–19503, doi:10.1029/1999JD900154, 1999.~~

Haverd, V. and Cuntz, M.: Soil-Litter-Iso: A one-dimensional model for coupled transport of heat, water and stable isotopes in soil with a litter layer and root extraction, J. Hydrol., 388, 438–455, doi:10.1016/j.jhydrol.2010.05.029, 2010.

Haverd, V., Raupach, M. R., Briggs, P. R., Canadell, J. G., Isaac, P., Pickett-Heaps, C., Roxburgh, S. H., van Gorsel, E., Viscarra Rossel, R. ~~a~~A. and Wang, Z.: Multiple observation types reduce uncertainty in Australia's terrestrial carbon and water cycles, Biogeosciences, 10, 2011–2040, doi:10.5194/bg-10-2011-2013, 2013.

~~Henderson-Sellers, A., Pitman, A., Love, P., Irannejad, P. and Chen, T.: The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3, Bull. Am. Meteorol. Soc., 76, 489–503, 1995.~~

Huang, M., Piao, S., Sun, Y., Ciais, P., Cheng, L., Mao, J., Poulter, B., Shi, X., Zeng, Z. and Wang, Y.: Change in terrestrial ecosystem water-use efficiency over the last three decades, Glob. Chang. Biol., 21, 2366~~–~~_2378, doi:10.1111/gcb.12873, 2015.

IPCC: Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.

Kala, J., Decker, M., Exbrayat, J.-F., Pitman, A. J., Carouge, C., Evans, J. P., Abramowitz, G. and Mocko, D.: Influence of ~~Leaf Area Index Prescriptions~~leaf area index prescriptions on ~~Simulations~~simulations of ~~Heat, Moisture~~heat, moisture, and ~~Carbon Fluxes~~carbon fluxes, J. Hydrometeorol., 15, 489–503, doi:10.1175/JHM-D-13-063.1, 2014.

Keenan, T., García, R., Friend, ~~a~~ A. D., Zaehle, S., Gracia, C. and Sabate, S.: Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO_2 and water fluxes through combined in situ measurements and ecosystem modelling, Biogeosciences ~~Discuss.,~~, 6, ~~2285–2329~~1423–1444, doi:10.5194/bgd-6-2285-2009, 2009.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. ~~a~~A., Mitchell, K. and Puma, M. J.: On the nature of soil moisture in land surface models, J. Clim., 22, 4322–4335, doi:10.1175/2009JCLI2832.1, 2009.

Kowalczyk, E. A., Garratt, J. R. and Krummel, P. B.: Implementation of a soil-canopy scheme into the CSIRO GCM: Regional aspects of the model response. ~~CSIRO Division of Atmospheric Research Technical Paper no. 32.,~~[online] Available from: http://www.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&recid=N9524192AH, 1994.

Kowalczyk, E. A., Wang, Y. P., Law, R. M., Davies, H. L., Mcgregor, J. L. and Abramowitz, G.: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, Aspendale, Victoria, Australia., 2006.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, 19, GB1015, doi:10.1029/2003GB002199, 2005.

Lai, C.-T. and Katul, G.: The dynamic role of root-water uptake in coupling potential to actual transpiration, Adv. Water Resour., 23, 427–439, doi:10.1016/S0309-1708(99)00023-8, 2000.

Lawrence, P. J. and Chase, T. N.: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), J. Geophys. Res. Biogeosciences, 112, G01023, doi:10.1029/2006JG000168, 2007.

~~Leplastrier, M., Pitman, A., Gupta, H. and Xia, Y.: Exploring the relationship between complexity and performance in a land surface model using the multicriteria method, J. Geophys. Res., 107, 1–12, doi:10.1029/2001JD000931, 2002.~~

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C_3 plants, Plant. Cell Environ., 18, 339–355, 1995.

Leuning, R., van Gorsel, E., Massman, W. J. and Isaac, P. R.: Reflections on the surface energy imbalance problem, Agric. For. Meteorol., 156, 65–74, doi:10.1016/j.agrformet.2011.12.002, 2012.

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.

Lin, Y.-S., Medlyn, B. E., Duursma, R. ~~a~~A., Prentice, I. C., Wang, H., Baig, S., Eamus, D., de Dios, V. R., Mitchell, P., Ellsworth, D. S., de Beeck, M. O., Wallin, G., Uddling, J., Tarvainen, L., Linderson, M.-L., Cernusak, L. a., Nippert, J. B., Ocheltree, T. W., Tissue, D. T., Martin-StPaul, N. K., Rogers, A., Warren, J. M., De Angelis, P., Hikosaka, K., Han, Q., Onoda, Y., Gimeno, T. E., Barton, C. V. M., Bennie, J., Bonal, D., Bosc, A., Löw, M., Macinins-Ng, C., Rey, A., Rowland, L., Setterfield, S. ~~a~~A., Tausz-Posch, S., Zaragoza-Castells, J., Broadmeadow, M. S. J., Drake, J. E., Freeman, M., Ghannoum, O., Hutley, L. B., Kelly, J. W., Kikuzawa, K., Kolari, P., Koyama, K., Limousin, J.-M., Meir, P., Lola da Costa, A. C., Mikkelsen, T. N., Salinas, N., Sun, W. and Wingate, L.: Optimal stomatal behaviour around the world, Nat. Clim. Chang., 5, 459–464, doi:10.1038/nclimate2550, 2015.

Lorenz, R. and Pitman, A. J.: Effect of land-atmosphere coupling strength on impacts from Amazonian deforestation, Geophys. Res. Lett., 41, 5987~~-~~–5995, doi:10.1002/2014GL061017, 2014.

Lorenz, R., Pitman, A. J., Donat, M. G., Hirsch, A. L., Kala, J., Kowalczyk, E. A., Law, R. M. and Srbinovsky, J.: Representation of climate extreme indices in the ACCESS1.3b coupled atmosphere–land surface model, Geosci. Model Dev., 7, 545–567, doi:10.5194/gmd-7-545-2014, 2014.

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton,

C. V. M., Crous, K. Y., De Angelis, P., Freeman, M. and Wingate, L.: Reconciling the optimal and empirical approaches to modelling stomatal conductance, Glob. Chang. Biol., 17, 2134–2144, doi:10.1111/j.1365-2486.2010.02375.x, 2011.

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~~ 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.

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. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J.-F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J. and Yang, Z.-L.: Technical description of version 4.5 of the Community Land Model (CLM), Boulder, Colorado., 2013.

Pitman, A. J.: The evolution of, and revolution in, land surface schemes designed for climate models, Int. J. Climatol., 23, 479–510, doi:10.1002/joc.893, 2003.

Powell, T. L., Galbraith, D. R., Christoffersen, B. O., Harper, A., Imbuzeiro, H. M. a, Rowland, L., Almeida, S., Brando, P. M., da Costa, A. C. L., Costa, M. H., Levine, N. M., Malhi, Y., Saleska, S. R., Sotta, E., Williams, M., Meir, P. and Moorcroft, P. R.: Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought, New Phytol., 200, 350–365, doi:10.1111/nph.12390, 2013.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S. and Voss, F.: How Well Do Large-Scale Models Reproduce Regional Hydrological Extremes in Europe?, J. Hydrometeorol., 12, 1181–1204, doi:10.1175/2011JHM1387.1, 2011.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment., Proc. Natl. Acad. Sci. U. S. A., 111, 3262–7, doi:10.1073/pnas.1222473110, 2014.

Puri, K., Dietachmayer, G., Steinle, P., Dix, M., Rikus, L., Logan, L., Naughton, M., Tingwell, C., Xiao, Y., Barras, V., Bermous, I., Bowen, R., Deschamps, L., Franklin, C., Fraser, J., Glowacki, T., Harris, B., Lee, J., Le, T., Roff, G., Sulaiman, a, Sims, H., Sun, X., Sun, Z., Zhu, H., Chattopadhyay, M. and Engel, C.: Implementation of the initial ACCESS numerical weather prediction system, Aust. Meteorol. Oceanogr. J., 63, 265–284, 2013.

~~Raupach, M. R.: A practical Lagrangian method for relating scalar concentrations to source distributions in vegetation canopies, Q. J. R. Meteorol. Soc., 115, 609–632, doi:10.1002/qj.49711548710, 1989a.~~

~~Raupach, M. R.: Applying Lagrangian fluid mechanics to infer scalar source distributions from concentration profiles in plant canopies, Agric. For. Meteorol., 47, 85–108, doi:10.1016/0168-1923(89)90089-0, 1989b.~~

~~Raupach, M. R.: Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index, Boundary-Layer Meteorol., 71, 211–216, doi:10.1007/ BF00709229, 1994.~~

Sakaguchi, K. and Zeng, X.: Effects of soil wetness, plant litter, and under-canopy atmospheric stability on ground evaporation in the Community Land Model (CLM3.5), J. Geophys. Res. Atmos., 114, 1–14, doi:10.1029/2008JD010834, 2009.

Samanta, A., Ganguly, S., Hashimoto, H., Devadiga, S., Vermote, E., Knyazikhin, Y., Nemani, R. R. and Myneni, R. B.: Amazon forests did not green-up during the 2005 drought, Geophys. Res. Lett., 37, 1–5, doi:10.1029/2009GL042154, 2010.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B. and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Rev., 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.

Sheffield, J. and Wood, E.: Drought: past problems and future scenarios, Earthscan, London, United Kingdom., 2011.

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.

Trenberth, K. E., Dai, A., Schrier, G. Van Der, Jones, P. D., Barichivich, J., Briffa, K.

R. and Sheffield, J.: Global warming and changes in drought, Nat. Clim. Chang., 4, 17–22, doi:10.1038/NCLIMATE2067, 2014.

van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. ~~a~~A. M., Liu, Y. Y., Podger, G. M., Timbal, B. and Viney, N. R.: The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, Water Resour. Res., 49, 1–18, doi:10.1002/wrcr.20123, 2013.

~~Wang, K. and Dickinson, R. E.: A review of global terrestrial evapotranspiration: observation, modeling, climatology and climatic variability, Rev. Geophys., 50, RG2005, doi:10.1029/2011RG000373, 2012.~~

Verhoef, A. and Egea, G.: Modeling plant transpiration under limited soil water: Comparison of different plant and soil hydraulic parameterizations and preliminary implications for their use in land surface models, Agric. For. Meteorol., 191, 22–32, doi:10.1016/j.agrformet.2014.02.009, 2014.

Wang, Y. and Leuning, R.: A two-leaf model for canopy conductance , photosynthesis and partitioning of available energy I : Model description and comparison with a multi-layered model, Agric. For. Meteorol~~. ,~~., 91, 89~~-~~–111, 1998.

~~Wang, Y. P., Leuning, R., Cleugh, H. and Coppin, P.: Parameter estimation in surface exchange models using nonlinear inversion : how many parameters can we estimate and which measurements are most useful ?, Glob. Chang. Biol., 7, 495-510, doi:10.1046/j.1365-2486.2001.00434.x, 2001.~~

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E. and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, J. Geophys. Res., 116, G01034, doi:10.1029/2010JG001385, 2011.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M. and Wang, Y.-P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341–1359, doi:10.5194/bg-6-1341-2009, 2009.

Yuan, H., Dai, Y., Xiao, Z., Ji, D. and Shangguan, W.: Reprocessing the MODIS Leaf

Area Index products for land surface and climate modelling, Remote Sens. Environ., 115, 1171–1187, doi:10.1016/j.rse.2011.01.001, 2011.

Zeng, X. and Decker, M.: Improving the numerical solution of soil moisture–based Richards equation for land models with a deep or shallow water table, J. Hydrometeorol., 10, 308–319, doi:10.1175/2008JHM1011.1, 2009.

Zhang, H., Pak, B., Wang, Y. P., Zhou, X., Zhang, Y. and Zhang, L.: Evaluating ~~Surface Water Cycle Simulated~~surface water cycle simulated by the Australian Community Land Surface Model (CABLE) across ~~Different Spatial~~different spatial and ~~Temporal Domains~~temporal domains, J. Hydrometeorol., 14, 1119–1138, doi:10.1175/JHM-D-12-0123.1, 2013.

Zhou, S., Duursma, R. ~~a~~A., Medlyn, B. E., Kelly, J. W. G. and Prentice, I. C.: How should we model plant responses to drought? An analysis of stomatal and non-stomatal responses to water stress, Agric. For. Meteorol., 182-183, 204–214, doi:10.1016/j.agrformet.2013.05.009, 2013.

Zhou, X., Zhang, Y., Wang, Y., Zhang, H., Vaze, J., Zhang, L., Yang, Y. and Zhou, Y.: Benchmarking global land surface models against the observed mean annual runoff from 150 large basins, J. Hydrol., 470-471, 269–279, doi:10.1016/j.jhydrol.2012.09.002, 2012.

Zobler, L.: Global soil types, 1-degree grid, , doi:10.3334/ ORNLDAAC/418, 1999.