Geosci. Model Dev. Discuss., 4, C1067–C1074, 2011
www.geosci-model-dev-discuss.net/4/C1067/2011/

Geoscientific
Model Development
Discussions

Interactive comment on "The CSIRO Mk3L climate system model v1.0 coupled to the CABLE land surface scheme v1.4b: evaluation of the control climatology" *by* J. Mao et al.

J. Mao et al.

s.phipps@unsw.edu.au

Received and published: 15 November 2011

General Comment

This paper describes a comparison with respect to land of control simulations of the CSIRO Mk3L atmosphere model when the model is run with two different land surface schemes: K91 and CABLE. The CABLE scheme is a replacement for the K91 scheme that couples carbon fluxes to the energy and water budgets, among other changes. The control simulations are evaluated against a range of reference datasets (observation and model derived) that span different periods of the twentieth century.

The simulations are also compared with each other to assess the effect on the coupled land-atmosphere system of replacing K91 with the more sophisticated CABLE scheme. This is a useful study because it provides a relatively clean benchmark of basic model variables in two models that may be compared against subsequent CABLE developments or results from other climate models. The authors conclude correctly that the change in land surface scheme from K91 to CABLE produces broadly similar coupled land-atmosphere simulations while adding important functionality to the Mk3L model.

The manuscript is well written, is clear in its intent and the authors achieve an honest balance in reporting where the models do well and do poorly. The figures are clear and clearly explained. However, given that this is a benchmarking article rather than a model description one, the analysis is limited in places and excludes important land surface variables. I recommend publication subject to the authors addressing the following comments.

We thank the referee for his/her positive and constructive comments, which have enabled us to improve the manuscript considerably.

Specific Comments

(1) The sites selected for the offline comparisons are (with one exception) all in North America or Europe, with none in the tropics. This seems an odd selection given that the most interesting model biases and differences occur in Eurasia and tropical South America. I can't help thinking that Fig 1 plots for a site in, say, Amazonia would be useful context for the results in later sections. Could the authors comment on how these sites were selected?

While we certainly agree that the six selected flux tower sites are not evenly distributed globally, their distribution is representative of the availability of this type of data (see http://www.fluxdata.org:8080/SitePages/). The six sites cover four different vegetation types in the mid and high latitudes. While we also agree that tropical sites would be ideal in this situation, the vast majority of sites that are freely available (including those in the tropics) failed the quality control criteria we set for the use of flux tower data for model evaluation. These criteria included minimal gap-filling with synthetic data, especially in meteorological drivers, energy balance closure issues and restriction to whole year periods to allow model spin-up.

We have amended the manuscript text (Section 3) to reflect this:

Figure 1 shows the simulation by CABLE of monthly averaged latent heat flux, sensible heat flux and net ecosystem exchange for the six locations detailed in Table 1.

has been modified to

Figure 1 shows the simulation by CABLE of monthly averaged latent heat flux, sensible heat flux and net ecosystem exchange for the six locations detailed in Table 1. Sites were chosen based on the completeness and quality of their meteorological and flux measurements over whole year periods.

(2) The authors highlight (correctly) net radiation as being the most prominent difference between models in these simulations, particularly in Eurasia. They attribute this to increased insolation and albedo change (P1622, L15), but provide no evidence to support this. They also attribute the increase in Eurasian JJA sensible heat flux to the net radiation increase, but Figure 7 indicates that there is a coincident reduction in latent heat flux. This suggests a more complex and interesting picture involving a large scale reduction in soil moisture and possibly differences in snow melt timing, which in turn may contribute to the original insolation increase. The reader is also given no

indication about the source of the albedo changes: are they from different prescribed ancillary fields (e.g., soil and vegetation) or from prognostic differences (e.g., snow cover)? While this study does not need an analysis of boundary layer and convection changes it does need a more thorough and quantitative description of differences at the land surface.

We agree. We have added two new panels to Figure 7 showing the albedo changes, and we have added associated explanatory text to Section 5.2.

(3) Even without the above comments on the Eurasian energy budget I'd consider soil moisture and snow mass part of the basic set of variables that should be reported. I suspect that many of the main differences between K91 and CABLE simulations are affected by these variables, so it would be useful to see them. It's also rare to have the opportunity to compare so easily soil moisture between different land schemes (the authors note that the same soil layer configurations were used in both models).

The referee sensibly suggests that we add an analysis of soil moisture to the manuscript. We have not done this. Soil moisture is a very difficult quantity to interpret (see Koster and Milly, The Interplay between Transpiration and Runoff Formulations in Land Surface Schemes Used with Atmospheric Models, Journal of Climate, 10, 1578-1591, 1997). If the soil moisture from the two schemes appeared quantitatively similar, this might lead a reader to interpret this as meaning that this variable did not cause differences in the fluxes. Conversely, large differences might be interpreted to mean the soil moisture was driving the responses. However, soil moisture has to be interpreted within the context of its role in the supply of moisture for evaporation from the soil, and for transpiration from the vegetation. This is a full-scale analysis in itself and is beyond the scope of this paper. In the absence of this analysis we suggest that including soil moisture would more likely than not mislead the reader.

(4) Could the authors comment on why they chose to compare their simulations with the IPCC TAR (2001) models rather than the more recent IPCC AR4 (2007) models?

We compare our modelling results with the IPCC TAR because Mk3L has a resolution and configuration most comparable to the TAR models. In addition, the TAR included some key diagnostics against which we wished to evaluate at the zonal scale.

(5) P1625: Discussion comments appear to contradict each other stating the net radiation change is "probably an improvement in CABLE" (L11) but note that there is an "absence of independent observations" (L15). How do the authors deduce the former given the latter? In fact, given that this is a relatively short manuscript, there are a few too many vague performance statements (e.g., "likely related to... leaf area", "zonal estimates are likely reasonable", "simulations are good") These add little that the reader can't see for themselves in the plots or are just speculation.

We mean by "independent" that there are not additional observations that we can use other than those shown in Figure 5. The terminology used is rather general and we have modified the text a little to more clearly explain what we mean (Section 6).

(6) The manuscript would benefit from a table of "bottom line" values for global, annual mean fluxes (LE, H, NPP) along with some discussion of how these compare with other estimates (observation and/or model derived). Benchmarks are most useful when they are specific. While the maps are enlightening in the context of this paper, it is very difficult in practice to make useful comparisons with similar plots from other models.

We agree. We have added a new table (Table 3) which provides DJF, JJA and annual means, as well as the bias and root-mean-square error, for the two different versions

of the model and for the following variables: near-surface air temperature, precipitation, net surface radiation and net primary productivity. We have also added a new paragraph at the end of Section 5 which discusses the data presented in the table.

Technical Comments

P1614 S2: While the authors are right not reproduce the model description details of Phipps et al, a simple statement of the Mk3L grid resolution used (horizontal and vertical) would be useful orientation for the reader.

We have added this information to the description of the model in Section 2.

P1615 S2: The manuscript should include some brief description of differences in K91 and CABLE ancillary data where they are relevant to subsequent results and discussion. E.g., are these data derived from different sources for each model?

The main issue with ancillary data of relevance is the albedo. The impact of these changes is now reported in the paper in the seasonal and annual data in the new table (Table 3). Other background data are hard to encapsulate here since they have different effective meanings between the two schemes. For example, the water holding capacity cannot really be reported in isolation – the impact of this parameter is highly dependent on how the parameter is actually used within the code.

P1616 L15: Are offline simulations really forced with net radiation?

No – the off-line simulations were forced with incident solar and infrared radiation. We did not mean to imply this and we have now clarified this within the text (Section 3).

P1617 L25: "...there is clear skill in the upper tails...". This skill may be present, but it's difficult to tell from the plots alone. The percentage overlap metric refers to whole histograms, so it is presumably insensitive to large relative but small absolute differences in the tails? Would the skill be quantified better by some metric derived from upper percentiles?

We have removed this comment from the text. While it would be possible to quantify the skill in the tails of the probability distribution functions, we feel that it would be beyond the scope of this manuscript.

P1618: The authors should comment (here or in the results) on how a simulation using pre-industrial atmospheric CO2 but late twentieth century SSTs (which include 100+ years of warming) might affect their comparison with observed late twentieth century datasets.

We wished to ensure that the land surface model was integrated to equilibrium, particularly with regard to terrestrial carbon storage. We therefore felt that pre-industrial boundary conditions were more appropriate. In the absence of high-quality reconstructions of pre-industrial sea surface temperature, late 20th century values were used instead. While we acknowledge this inconsistency in the experimental design, we note that the increase in global-mean sea surface temperature between pre-industrial times and the late 20th century was only $\sim$0.5 K (Folland et al., 2001). Furthermore, the default configuration of the model (Mk3L-K91) has been extensively evaluated under the same boundary conditions and biases relative to 20th century observations were found to be very modest (for example, an RMS error in near-surface air temperature of 1.90 K; Phipps et al., 2011). We have added comments to this effect (Section 4.1).

P1620: The sections "Surface forcing fields" and "Surface fields" are structured oddly. The former heading is misleading because these are coupled land-atmosphere simu-

lations rather than offline simulations of K91 and CABLE, so these variables are not strictly forcing the land schemes. Also net radiation results are described in both sections, which makes the distinction of the two sections confusing. Perhaps it would be clearer to move the net radiation results entirely to the second section and rename the sections simply "Temperature and precipitation" and "Surface energy budget".

We agree that the headings of these sections were confusing. We have rectified this by renaming Section 5.1 from "Surface forcing fields" to "Coupled simulation of surface forcing fields", and by renaming Section 5.2 from "Surface fields" to "Coupled simulation of surface fields".

P1621 L17: Figure 8.3 rather than 7.3 of the TAR, no?

We have corrected this typographical error.

Figures: Global maps of land-only variables are better presented when they run [-180, 180] degrees longitude rather than [0, 360] degrees, as this emphasises land rather than the blank Pacific. This is, I admit, a matter of personal taste.

There is no fixed convention in this regard, and we have therefore chosen to leave the figures in their current form.

Interactive comment on Geosci. Model Dev. Discuss., 4, 1611, 2011.