

Interactive comment on “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” by E. Blyth et al.

Anonymous Referee #2

Received and published: 16 January 2011

The manuscript describes a suite of benchmarking tests suitable to reveal a land surface model's performance in simulating water and carbon fluxes concurrently, over multiple spatial and temporal scales.

The authors have chosen a synthetic and pragmatic approach to such a vast problem and their overall concept and approach is useful and immediately applicable as "first line of attack", also potentially useful in model inter-comparisons.

The review of existing literature is appropriate, so is the credit given to all the programmes and investigators that have provided access to the observational data that are the foundations such a multi-scale benchmarking suite.

C694

The abstract leads us to believe that the method will allow us to "understand the causes of model errors". The manuscript falls very short of this, in quite a few areas, however:

1) analysis of results is rushed and superficial, limited to a repetition of "too high / too low" instead of quantifying (e.g. an error of 20W/m²); without any attempt to point out what is a significant error, for instance in the light of known observational uncertainties, or in the light of requirements for climate prediction, and what is not. Here I suggest, for instance, that interannual variability is used as an indicator of significance of the errors that are presented; additionally, or alternatively, the literature review is used to point out which errors are significant.

2) the "metrics" provided in Table 2 are not properly introduced, nor justified, and not used in a standardised and integrated way to support analysis and discussion all throughout the manuscript. These metrics and their application should be the central point of the paper and end up featuring, instead, like an afterthought. I suggest inserting a new section, "Metrics" and dedicating substantial thought and effort to it.

3) while the simultaneous consideration of water and carbon fluxes is a strength of the proposed method, little effort is made to join up the information provided by the results at analysis time, for instance going into a much more mechanistic discussion of how errors in carbon assimilation are reflected in runoff and how quantitatively consistent the two results are. The two cycles seem quite disconnected in terms of the analysis and discussion. This, again, falls short of what had been promised in the abstract.

4) there is no information as to the fundamental characteristics of the model presented. Publications from the early 1990s are cited, then two model description publications "to be submitted". In lack of those, there is a need for a 1/2 to 1 page description of what this model is.

5) Overall, the errors presented seem very large, in particular for evaporative fluxes and for GPP. The fact that a model with such errors is used for climate prediction is alarming. The fact that no detailed, quantitative analysis is presented, nor any explanation given

C695

in the discussion as to why these errors are so large, is very disappointing.

6) The weight given to the discussion of other people's work, up to Section 3, is substantial. In comparison, the amount of work presented that was completed by the authors seems somewhat insubstantial. I suggest shortening and better integrating the paragraphs in Sections 1,2, which often sound repetitive (e.g. all the "importance of land surface" sentences) and dedicating much more space and effort to the in-depth discussion of the work completed by the authors for this specific study.

Specific points: page 1844, top, discussion of metrics. The discussion of "pass mark" is vague, difficult to follow, especially in the lack of rationale for the proposed metrics and provides no new insights. I suggest that either a rigorous discussion is provide, or the entire section is removed.

There are a few places in the text where carefully re-reading would reveal lack of prepositions, unwanted prepositions, else inconsistencies in stated information, e.g. the hourly/half-hourly FLUXNET measurements discussion at the start of Section 2.1

Interactive comment on Geosci. Model Dev. Discuss., 3, 1829, 2010.