Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles" by Fuxing Wang et al.

Anonymous Referee #1

Received and published: 27 February 2018

The paper describes a number of experiments assimilating GRDC runoff data into the ORCHIDEE land surface model across the Iberian peninsula. The assimilation adjusts the simulated runoff at sub-catchment scale where GRDC observations are available through a 'optimization parameter', effectively rescaling the simulated runoff towards the observations. The discharge bias is substantially reduced by adjusting this 'optimization parameter' and neighbouring sub-catchments are corrected by extrapolating the parameter to these.

The paper is clear and it is well written. The study is likely to be very relevant for future studies possibly extending and improving on the presented concept.

I propose a minor revision.

Some general remarks: The validation is performed on the basis of using GRDC both as an 'observation' and an independent 'validation' dataset? This should be discussed very critically. I am not an expert in the field of continental runoff and possibly there is no other independent data source to have a better independent validation. In general this is however quite uncommon in assimilation studies, i.e. satellite observations might be assimilated to improve soil moisture and the results would be validated against independent in-situ measurements.

For probably this reason the authors compare the corrected evaporation against GLEAM. As stated, GLEAM uses a different precipitation, the entire comparison therefore is challenging. Did the authors consider using the same precipitation as input for their experiments? It should be quite simple. Also, corrected evapotranspiration values could be compared to Fluxnet in-situ measurements. This should be either included or a strong case should be made why this was not done. The motivation of exactly / only using GLEAM should also be well presented. There are a number of alternative evapotranspiration products.

The correction factor x is applied to each sub-catchment for runoff. It was not quite clear to me how the evapotanspiration was then corrected, presumably at a grid cell level? This duality between correcting at catchment scale but the model essentially being a distributed one computing the water balance at each grid cell should be made clearer. The model runoff is corrected as it was a lumped conceptual land surface model but the relationship between this and the land surface heterogeneity is not clear to me. Also, is equifinality a serious issue? I suppose a number of optimized x can result in the same or very similar runoff downstream? Can this be mitigated by also looking at the correct seasonality of the generated runoff?

The proposed method is supposed to be superior to more simple water-balance methods? Can this be somehow quantified?

Despite the in general high-level language there are a number of inaccuracies (for instance missing articles).

Specific:

L155: . . . for different parameters . . ., parameters includes also variables, such as soil moisture, runoff etc.?

L172: Again, I'm getting confused with parameter and variable, I suppose parameter is x, but the actual runoff is a variable? Please take care with this throughout the text.

L173: The background error B is vital in DA, why was it chosen like this? More detail needed.

L218: Is WFDEI not being updated? Please recheck.

L237: Does each HTU have it's own location within a grid cell? Or is it more 'conceptual'. Might be helpful to clarify this in the model description. I'm assuming that they have a fixed location within each grid cell.

L266: What is meant by one optimization parameter? In my understanding the algorithm only perturbs x to find the optimum fit between the runoff simulations and observations? The river routing parameters are perturbed? Or does it depend on the number of upstream catchments with a separate x?. Not quite clear to me.

L274: . . . value '1' and a 'pre-estimated error': 'and' should be 'or'?

L282: the cost function is lower? The value of the cost function? Section Experiments design needs to be a bit clearer.

L288: factor m corresponds to number of GRDC stations?

L304: The river routing model runs at each grid cell? The distributed nature of the river routing model is not quite clear.

L322: higher than a factor of 1.5?

C3

L359: "Summary" seems misnamed for the amount of text following

L369: They most certainly do. . .

L375: → can allow, remove 'of'

L377: → patterns, some inaccuracies in this area

L383: Is it also connected to topographic or other land surface features which might be not well presented by the forcing data or the model itself? Just wondering.

L475: GREAM → GLEAM

L479: references, also maybe mention more global attempts to create gridded runoff data? (can be in the introduction).

L507: Throughout the paper most errors are attributed to the lack of human influences. For sure other factors also play a large role?

Figure3 top: With the logarithmic scaling the lines mostly seem pretty horizontal. Is there a clearly visible gradient when using a different scale? Maybe add this as a window. Missing unit for J?

C4