Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-206-RC3, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Evaluating the strength of the land–atmosphere moisture feedback in earth system models using satellite observation" *by* P. A. Levine et al.

Anonymous Referee #3

Received and published: 23 June 2016

In this study, the authors aim to assess the feedbacks between terrestrial water storage (TWS) and the atmosphere in remote sensing observations (using GRACE data for TWS), and they aim to use the estimate they obtain to evaluate climate model simulations from the CESM large-ensemble and from (a subset of) CMIP5 models. They use correlations at the interannual time-scale, with lead-lags of a few months between TWS and atmospheric variables, to assess both how the atmosphere (during months of TWS decrease) influences the minimum annual TWS (the forcing response), and how TWS (at time of maximum annual TWS) feeds back on the atmosphere in the following months (the feedback response). Their main result include some characterization of these relationships in model and observations, which show that the CESM model, and

Printer-friendly version

Discussion paper

most CMIP5 models, appear to overestimate both forcing and feedback relationships. Some discussion is then proposed on the possible reasons for that.

Although I like the idea of trying to use GRACE observations to derive a large scale "land-atmosphere moisture coupling" metrics to compare to models, I see a number of issues with the study as it stands, that I think warrant some major revision.

1) My first main comment has to do with the metric used. First, the brevity of the time period of analysis is obviously an issue. The metrics are essentially interannual correlations over 13 years, i.e., correlations over 13 data points. That's very short. I am not sure I have seen many studies looking at interannual variability over 13 years only, in particular in terms of land-atmosphere studies. Accordingly, results for the feedback metrics appear very noisy. First, I believe a discussion of field significance is warranted here: patches of apparently significant values may still be random in that context (e.g., see Livezey and Chen (1983)). Second, note that recent research underscores the need for long-record datasets to establish land-atmosphere coupling, that coupling metrics require more data than single-variable simple statistics (e.g., mean and variance) to be robustly estimated, and finally that, unlike single-variable statistics, coupling metrics are actually degraded by observational uncertainty (Findell et al. 2015). The latter point, in particular, is in my view a much likelier explanation for the weaker correlations found here in observations – between uncertain observation datasets that are independent of each other - compared to correlations computed with model outputs, which are by definition perfectly consistent with each other. The authors touch on the issue of observational uncertainty by computing correlations with the ERA reanalysis, but I don't think enough is made of that. So, given the brevity and uncertainty of observations, even without consideration of any other issues (but, see below), I am really uncomfortable with the approach proposed by this paper, which is to consider the observational estimate as a benchmark for model evaluation. Personally, I think an approach where observations and model results are used together to try to infer the 'real' coupling would make more sense here. But, this would lead to a very

different paper. Overall, if the authors are going to go on with their approach, I would recommend much more caution in how things are presented, including in the title of the study and the conclusions.

2) Another issue with the metrics involves the definition of the (feedback) metric. The way it is defined, it is looking at the impact of TWS at the end or peak of the rainy season on climate in the subsequent months. The authors indicate as much, and say they want to consider, in the Tropics, the impact of late-rainy season TWS on dry-season climate. I see two issues with that. First, in my view, while that may useful in the deep Tropics where the dry season is short, this approach is problematic in monsoon regions, or regions of the Tropics that have a well defined rainy season (i.e., outside of the deep Tropics): basically, after the rainy season, there is not much rain to look at any more. For instance, over the Sahel, what the authors are computing is the impact of September TWS on precipitation over September-May. But it doesn't rain much over that time period. In my view, it would be much more interesting to look at the impact of end-of-dry season TWS on the subsequent rainy season to see if, in these regions, available land moisture feeds back on precipitation during the rainy season. Second, in the same example over West Africa, whatever rainfall there is over Sept-May is actually probably the end of the monsoon, Sept-Nov. Because TWS in September is likely to be influenced by precip in September, and Sept. precip is likely to represent large part of the 'response' variable, the causation is muddied a little bit: a clearer temporal offset would be needed in such a case. But more importantly, even precip in the months following September (Oct, Nov) is likely to be correlated with precip in the previous months – for instance, a year with a strong monsoon that has more rain in Jun-Sept may well tend to also have more rain in Sept-Nov. Because September TWS will largely reflect JJAS rainfall, the TWS-based metric will then show a strong feedback - but the inferred causation would be a misinterpretation. This brings me to a more general point: the authors do not discuss how autocorrelation, here at the seasonal time scale, of climate variables, may impact their estimate of land-atmosphere coupling. This is a major issue affecting all empirical studies of land-atmosphere cou-

pling – see, for instance, Wei et al. (2008) and Orlowsky and Seneviratne (2010). The authors do cite the latter study, but, it seems, simply to say that if models underestimate SSTs influence on land climate, they will then appear to overestimate local L-A coupling. They somehow miss the point of that paper in how it should apply to to their own results. I just gave one example above (the Sahel) of how that might be the case. Other monsoon regions (e.g., India) might similarly be affected. Interannual variability in the coupled ocean-atmosphere (eg., ENSO) might also the source of confounding influence at the time scales investigated here. So, overall, I recommend these caveats be considered and discussed by the authors in their interpretation of their results. Personally, I would need to see some further analysis to be more convinced of the physical reality of the land-atmosphere feedbacks the authors claim to show (e.g,. some sensitivity test to the months and time lags considered, some investigation of atmospheric variability and persistence, etc.).

3) Another main comment has to do with the discussion section. The authors discuss why models might exhibit stronger feedback (and forcing) metrics than observations. As mentioned above, I think uncertainty in observations should be mentioned as a primary reason. The authors propose that ET may be consistently overestimated in climate models, and a large part of the discussion then consists in speculation as to why that may be the case. First, while I appreciate the effort to discuss things further and not just show results, I found this whole section a bit too speculative. IF the models overestimate ET, then IF stomatal conductance, IF convection, IF bare soil, etc... Can the authors actually point to any evidence that ET is consistently overestimated in climate models, in the first place (or at least in CESM)? Second, if soil water is too readily available in models, and ET is overestimated, wouldn't that actually mean that feedbacks should be underestimated in models? Indeed, ET would then be less water-limited and more energy-limited, with less potential for soil moisture-atmosphere feedbacks. Surface climate variability would then be influenced by the atmosphere to a greater extent. Along the same lines, the authors claim that their results, showing an overestimation of land-atmosphere feedback by models, are consistent with prior

studies, and have implications for projected warming (e.g., Cheruy et al., 2014). However, these previous studies, it seems to me, point to ET being underestimated in these models, and models getting to easily “locked” in a dry and warm surface mode. So, in effect, while the authors agree with prior studies that land-atmosphere feedbacks are overestimated in models, they provide opposite reasons for that (overestimated versus underestimated of ET). I would like to see the authors clarify that apparent contradiction.

4) Finally, the author interpret the relationship they find between the strength of the feedback and forcing metrics in CMIP5 models as showing that: “ the response limb of the feedback loop is important for understanding how conditions are set up for subsequent forcing via land–atmosphere coupling”. They claim that it highlights “the importance of the land surface response in priming the system for subsequent forcing on the atmosphere”. I am not convinced by this interpretation, which sounds a bit hand-wavy to me. I don’t see a strong physical reason why a model where, for instance, TWS responds strongly to precipitation, should have a strong feedback of TWS onto precipitation. Couldn’t the relationship on Figure 10 be due to intermodel differences in what TWS (or its estimate, here) encompasses in each model? For instance, different soil depths? A deeper soil would lead to weaker links between TWS and climate both in terms of response and feedback to the atmosphere. In any case, I found Figure 10 to be insufficiently explained and encourage the authors to discuss this further.

Here are some further comments along the text:

- p.2 line 4: “cloud radiative coupling”: please explain and clarify.

- P.2 line 24: actually, no: a surprising result of GLACE II was that predictive skill was not enhanced over the Great Plains “hot spot” from GLACE I, but rather to the North of it (see Koster et al. 2011). Consider rephrasing.

- P.3 line 5: the text should make it clear that GLACE-like metrics cannot be directly compared to observations, and that other more simple metrics, not strictly equivalent, have to be used, like SM-ET correlations, etc.

- P.3 line 20: Findell et al. 2011 is actually based on reanalysis data, not "modeling". Also, Findell et al. 2015 should be included in this discussion, to highlight the issue, discussed above, of data length requirements to estimate land-atmosphere coupling.

- P.5 line 15: is that version of the GRACE data downscaled in any way, and if so, how? I thought the basic GRACE data was at coarse resolution (e.g., 500km).

- P.5 line 17: "the TWS time series". I read that GRACE data are actually anomalies compared to the mean over 2004-2009. How is that accounted for in the computation of the metrics? Are the other variables centered on the same years? Does that affect results in any way? What about model outputs?

- P.6 line 32: see main comment above: I am not sure this is the most relevant time of year to investigate, and they are issues of rainfall autocorrelations.

- P.7 lines 12-15: that is, if the feedback is actually a positive moisture feedback. In other words, the authors adopt the a priori view that they are looking at a positive, moisture recycling feedback. This should be stated more explicitly, and perhaps earlier in the manuscript.

- P.8 line 18: what about AMIP simulations?

- P.8 line 28: It's unclear to me why the authors restrict themselves to the GLACE-CMIP5 models. There is no further comparison in the manuscript, on a model-by-model basis, with results from that experiment. So why not use the whole CMIP5 ensemble?

- P.9 line 9: so what? What is made of that? What are the implications for the correlation-based metrics? This comment applies to the whole sub-section, actually, including the result about climate variability. If anything, higher variability in model outputs would point to lower correlations, if the covariance between TWS and climate is similar.

- P.9 line 11: aren't trends removed from this data? Please clarify.

- P.9 line 26: still, why would the covariance be positive?

- P.9 whole section 3.2: this whole subsection feels very descriptive. On the other hand, there is not much description of the processes themselves. This might feel obvious to the authors, but some further discussion of what the correlations mean physically, when describing the figures, may be welcome.

- P.10 line 13: the link with cloud cover and precipitation should be explicitly mentioned here.

- P. 11 lines 14-15: see main comment #4 above.

- P.11 line 23: "Discussion".

- P.11 line 28: as mentioned above, these "well understood mechanisms" are actually never really explained.

- P.12 lines 3-4: that's exaggerated. Feedback results on Figures 5-7 are very noisy, and even from a simply qualitative perspective, it is a stretch to say that they agree with results from GLACE 1. One could just as well point out all the regions on Figures 5-7 that do NOT show up in GLACE 1 and say results are completely different. Besides, I find it a bizarre impulse (or maybe, a testament to the strength of the GLACE 1 study) that every land-atmosphere study seemingly feels the need to point out some level of agreement with GLACE results, even when, as is the case here, the match is very weak at best, and more importantly, when different data (observations versus models), processes and spatio-temporal scales are considered. Consider removing that comparison.

- P.12 line 16: see main comment #4 above.

- P.12 line 26: the authors could still look at this in models results, though. In fact, showing the link between TWS and ET, for instance, would reinforce their results and the physical interpretation that they propose.

- P. 13, first paragraph: this is unclear. Do the author mean that that the models underestimate remote influences of SSTs, for instance, and thus appear to have too strong a coupling?

- P.13 lines 16-18: see main comment #3 above.

- P. 14 line 18: but here observations show positive coupling, too! Please clarify.

- P. 14 line 21: but reduced stomatal opening would be associated with reduced ET, too. Please clarify.

- P. 14 lines 18-30: See main comment #2. There is a fundamental issue with the manuscript here.

- P.15 line 3: see main comment #3.

- P.15 lines 8-9. not really: Seneviratne etal. (2013) show that long-term soil moisture change leads to more warming, differently across models in the GLACE-CMIP5. That, in and of itself, could be considered an estimate of (long-term) soil moisture-atmosphere coupling in these models; but, in any case, there is no comparison to estimates of present-day coupling.

- P.15 lines 11: No. Warmer air "holding" more water vapor and leading to more precipitation would lead to positive temperature-precipitation correlations – not negative.

- P.15 line 13: "determined": not really. What Berg et al. (2015) show is that because of land-atmosphere interactions, the interannual negative temperature-precipitation relationship that they identify in present-day climate holds on longer time scales, including in the case of climate change. This may be interpreted as suggesting, as the authors say here, that models with too strong a coupling will then overestimate future warming; however, it is not directly shown by that study. Consider rephrasing.

- P.16 line 10: see comment above on P.12.

- P.16 line 11: "regions of strong RESPONSE metrics", I believe.

- P.16 line 14: the implication is bit too implicit here. Consider being more explicit.

Figures

- Figure 1: nice figure that helps understand the study. The y-axis on a) refers to anomalies, I presume – see comment on GRACE values above.

- As noted above, Figure 3 and 4 are nice, but not much is made of them in the analysis.

- Figure 5-7: I suggest the authors modify the color legend here. More color shades is not always better. It is actually not easy to see differences in color shades on a continuous bi-color palette like here, and for the reader things essentially end up being two colors, one positive (green) and one negative (red). It would actually be easier to have fewer shades, more clearly separated, and with perhaps several different colors as well.

- Figure 8: I suggest showing the mean of the CESM distribution as well.

References cited in this review:

Findell, Kirsten L., P Gentine, B R Lintner, and B P Guillod, 2015: Data Length Requirements for Observational Estimates of Land-Atmosphere Coupling Strength. Journal of Hydrometeorology, 16(4).

Livezey R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. Mon. Wea. Rev., 111, 46–59.

Wei, J., R. E. Dickinson, and H. Chen, 2008: A negative soil mois- ture–precipitation relationship and its causes. J. Hydrometeor., 9, 1364–1376.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-206, 2016.