

## ***Interactive comment on “Do land parameters matter in large-scale terrestrial water dynamics? – Toward new paradigms in modelling strategies” by L. Gudmundsson and S. I. Seneviratne***

**L. Gudmundsson and S. I. Seneviratne**

lukas.gudmundsson@env.ethz.ch

Received and published: 4 February 2014

Reviewer #1 raises several interesting concerns regarding the scope and the structure of the manuscript. In the following we provide point by point answers to these issues. For the sake of clarity we first repeat the reviewer's comments (*in italic*) and then provide our response.

**Comment 1:** *The title of this piece was very encouraging as the evaluation of the relative contribution to uncertainty of land surface parameterization versus model structure is certainly a hot topic in hydrologic science at the moment, but I was sorely disap-*

C7708

*pointed by this manuscript which does not live up to the excellent reputation of the authors. Certainly the article does NOT move us towards new paradigms in modelling strategies and the assertions that physical process development in land surface/large scale models is not necessary is unfounded based on the provided discussion (and most probably at all).*

**Reply 1:** We are pleased that Reviewer #1 finds the scope of the article interesting and regret that the submitted manuscript was not formulated clearly enough to communicate the content convincingly. We want, however, to emphasize that we never suggested that “*physical process development ... is not necessary*”. We instead focus on a specific area that has received high attention in a number of publications (i.e. the importance of local parameters for hydrological processes) and assess whether it is relevant for large-scale hydrological dynamics ( $\approx 50$  km, monthly). Our results show that for applications focusing on this scale (which encompass e.g. climate model applications), the representation of these small-scale parameters is actually less relevant than often assumed. This is an essentially new result.

However, we recognize that the previous title may have wrongly conveyed that we were proposing new paradigms for hydrological modeling, while our aim is rather to emphasize the need for such a development. In order to better define the scope of the article, we have decided to remove the second part of the title and now replaced it as follows: “Do land parameters matter in large-scale monthly terrestrial water dynamics?”.

**Comment 2:** *The manuscript is not helped in that it is not clearly formulated and difficult to follow in parts. The language and style is quite obtuse and the structure is very strange and complicated. Why all the appendices? Where is the coherent story?*

**Reply 2:** We acknowledge that parts of the manuscript had to be improved and have modified the revised version accordingly.

**Comment 3:** *The most important point however is that nothing has been evaluated on temporal or spatial scales that actually matter.*

C7709

**Reply 3:** We fundamentally disagree with the reviewer on this point, which is an essential part of our argumentation. The scope of this article is the efficient modelling of terrestrial water dynamics on large spatiotemporal scales ( $\approx 50\text{km}$ , monthly). Phenomena occurring on (sub-)continental and monthly scales are actually of great importance for both practical and theoretical aspects of earth science, in particular in the context of climate change projections. Typical climate models, which are used for a range of hydrological applications (e.g. assessments of projected changes in large-scale drought or flood occurrence) are generally run at coarse resolution (ca.  $2^\circ$ , i.e.  $200\text{ km}$ ) and their output is evaluated at monthly or yearly time scales. In this context, our results highlight that there is most likely little need to develop highly complex land surface schemes to assess such changes in water availability. This is an essential new finding, since land surface model development in the context of climate modeling is still aiming at increasingly high complexity, which as we show, is not essential for some of the key applications these models are used for. We now clarify this point in the revised manuscript.

Nonetheless, we agree that for some other applications (e.g. high-resolution extremes), a detailed (and high resolution) representation of the underlying processes is necessary. Possibly this point was not clear enough from our previous version of the manuscript. For this reason, we also add more text on this aspect in the revised article.

**Comment 4:** *I do not understand how if you use spatio-temporal scales greater than where runoff dynamics are thought to be influenced by local varying land properties (fig2) you can say anything about tools that consider locally varying land properties. Surely it is totally obvious that the high spatial resolution processes will not have an effect.*

**Reply 4:** This is actually again exactly the point of the article (see Reply 3 above). In Sections 2 & 3 we formally introduce the hypothesis that the influence of land parameters on terrestrial water dynamics diminishes if phenomena occurring on scales  $\gg 10\text{ km}$  and  $\gg 7\text{ days}$  are considered. If this holds true this implies that models

C7710

targeting phenomena on these scales can rely on parsimonious formulations. Similarly as highlighted in our Reply 3, we do now provide more text in the revised manuscript to better clarify the scope of the article, as well as the rationale for considering large spatial and temporal scales in hydrological dynamics.

**Comment 5:** *Here there is only evidence that the CLPH can provide a simple rainfall-runoff transformation of the precipitation at gridscale and at low temporal and spatial resolution, i.e. a large scale overview which does not consider storage or routing. This is not surprising or new.*

**Reply 5:** We also disagree with the reviewer on this point. While the results may not be surprising to some readers, they are substantially novel compared to previous applications and analyses:

1. Usually rainfall-runoff transformations depend on location-specific parameters, which are generally assumed to be a function of land properties. This is contrasted by our application which has the same parameters at any location in space.
2. We consider (limited) storage effects explicitly. (See the time-lag operator  $\tau_n$  in Equations 9 & 10).
3. Our technique allows for complex interactions which for instance enables us to capture the contrasting effect of temperature on the seasonality of runoff in snow vs evapotranspiration dominated regions. This would e.g. not be possible with a pure unit hydrograph model. Further it is noteworthy that state-of-the-art LSMs have pronounced issues in capturing this pattern. (This is also documented in great detail in other publications; see Gudmundsson et al., 2012b).

We better highlight these points in the revised article.

C7711

**Comment 6:** *The comparison with observed has only been done against short residence times and there is no consideration of terrestrial water storage and longer residence times.*

**Reply 6:** This is not correct. We provide an extensive model validation (Figures 3, 4, B1 and Table C1) using three different variables ((i)runoff from small catchments, (ii) discharge from continental scale rivers, (iii) evapotranspiration) focusing on several features of terrestrial water dynamics:

1. The analysis of biases of runoff (BIAS in Figure B1) river discharge (Table C1) and evapotranspiration (Figure 4) provides insights on systematic differences in observed and modelled quantities at time scales of multiple years.
2. Correlations between observed and modelled mean annual cycles for runoff (climatology -  $R^2$  in Figure B1) provides insights to model performance on seasonal time scales.
3. The comparison of observed and modelled monthly anomalies gives insights to the models ability to capture runoff (anomaly -  $R^2$  in Figure B1) and river discharge ( $R_{\text{ano}}^2$  Table C1) at the highest temporal resolution considered.

Some of these results are presented as supplementary information in the Appendix, as they allow for a more detailed understanding of model performance, but are not essential for the main scope of the paper (introducing and testing the CLPH).

**Comment 7:** *Thus I find the conclusions are totally overstated. If CLPH-RFM can be used as a pragmatic estimator of continental scale terrestrial water dynamics, then so can the meteorological variables directly. Surely. What is the point?*

**Reply 7:** As already outlined in Equations 1, 5 & 6 and in Reply 5 we expect that terrestrial water dynamics result from complex transformation of atmospheric variables.

C7712

Consequently we do not expect that untransformed atmospheric forcing (e.g. raw precipitation) is an equally good estimator for the variables under investigation as the presented Random Forest Model (RFM). We do, however, show that this transformation is more general (and less dependent on land parameters) than usually assumed.

**Comment 8:** *Why is the RFM appropriate to use here? This seems like a horribly overparameterized methodology, perhaps even rivalling the more complex hydrologic models and land surface models in terms of free parameters. Why do you need this complexity? This is not obvious and seems at odds with the aim to free modellers from unneeded parameterization. Also there can be no physical basis after the RFM has been implemented and this makes tracing back processes impossible. Why is this a useful method?*

**Reply 8:** We do acknowledge that the background of the used techniques has to be introduced more carefully. Random Forests are one of the many methods from the rapidly growing field of statistical pattern recognition / machine learning (Hastie et al., 2009; Bishop, 2006). Random Forests are, like similar techniques (e.g. Support Vector Machines, Neural Networks, Gradient Boosting), built to identify complex relationships. This flexibility comes at the cost of losing the interpretability of the internal structure of the resulting models. In the context of the presented study, this flexibility is desirable as we want to make the assessment of the CLPH independently of the plethora of typical LSM formulations.

In principle we could also have calibrated global parameters of a LSM, however, the long recognised issue of over-parametrisation and equifinality (Beven, 2006) would have rendered the physical interpretation of the resulting parameters infeasible (as already indicated on p. 13199, l. 3-4). Therefore we find sacrificing the interpretability of the internal structure of Random Forests acceptable. Especially as our main result is based on the comparison of the performance of a RFM with (Equation 9) and without (Equation 10) land parameters (and not the comparison CLPH-RFM vs LSMs).

C7713

**Comment 9:** *The whole point of the additional complexity in land surface models and large scale hydrological models is that they might want to do something more complicated than just predict gridscale runoff, for example, (i) act as an Earth System Model and feedback to the atmosphere or biosphere (water, energy, and carbon balance/cycle together), (ii) interact with changes in landuse or other environmental change, (iii) undertake higher spatio-temporal forecasting where the hydrograph dynamics, variations in soil column water, groundwater variability (sub-monthly and routed down the river) are essential to capture.*

**Reply 9:** We fully agree with Reviewer #1 that the scope of LSMs goes far beyond approximating terrestrial water dynamics at large spatiotemporal scales and do not put this objective into question. We note, however, the following points:

1. The complexity of these models needs to be justified by providing estimates of key variables that compare favourably with observations. To date, runoff and river discharge are among the best monitored variables of the terrestrial environment and thus we consider it essential that LSMs do capture at least the main features of their large-scale dynamics. However our results do not suggest a good performance of these models at this scale and thus question some of their underlying assumptions.
2. We obviously agree that land surface models are essential to capture a number of feedback processes when used in coupled mode, and that some of these feedbacks occur at smaller spatial and temporal scales than those considered here. Nonetheless, many relevant processes which involve such feedbacks actually happen at those large scales (e.g. droughts), and additionally, the Earth System Models themselves are rarely run at resolution of less than 200km.
3. Finally, although we agree that several hydrological applications require a high spatial and temporal resolution, it is important to remember that this does not

C7714

apply to a large number of other important applications (e.g. climate model predictions of changes in water availability)

**Comment 10:** *“However, the mismatch between their temporal and spatial resolution raises the question of whether this can be successful” - what do you really mean by temporal and spatial resolution here. What has one got to do with the other? What is the physical basis for this comment? Hydrologic physics is full of non-linear transformations which makes these generalised relationships between time and space unhelpful. What if the models were applied at higher spatio-temporal resolution (as many LSMs/GHMs are undertaking right now)? I agree that even higher spatial resolution does not solve all the problems of Beven & Cloke 2012, but it doesn't follow automatically that we should remove all land parameterization completely.*

**Reply 10:** Any model simplifies the world by dividing it into discrete space (e.g. grid cells) and time (e.g. time steps) units. The spacing of these units determines which phenomena can explicitly be represented by the respective model (cf. Nyquist frequency). It is true that we have not communicated this clearly enough in the first version of the article and we have added further text focusing on this topic in the manuscript.

For illustration, let us consider the hydrograph of a small catchment occurring after intense precipitation. Let us also assume that the typical temporal resolution of LSMs (hours) is sufficient to capture the time evolution of this hydrograph. The precise shape of this hydrograph is also dependent on e.g. the geomorphology of the small catchment, which cannot be represented by a 50 km × 50 km grid cell. Consequently there is a mismatch between the temporal resolution of the LSM, requiring detailed information on land properties and its spatial resolution which limits the amount of details.

**Comment 11** *The evidence provided distinguishing the bias in forcing compared to the land parameterizations (table C1) is not convincing. What happens if you do bias correct the precipitation and other variables per model instead of the bulk WATCH correction as might be more typically done in climate impact studies? What about the*

C7715

*additional uncertainty from the bias correction itself?*

**Reply 11:** We think that this is a misunderstanding, the results presented in Table C1 and Figure B1 do not report on any bias correction, but show the bias, i.e. the mean difference between observed and modelled variables.

The WATCH Forcing Data (WFD) used to force both the LSMs and the statistical models do actually provide bias corrected estimates of atmospheric variables (see Weedon et al. (2011) for details). As the uncertainty of this bias correction is identical in every case this should not impair our results.

**Comment 12** *What happens if you let the evaporation evolve as in a land-atmosphere coupled model instead of forcing it with constant forcing? Typically you would expect the model performances to change and yet these things have not really even been mentioned let alone carried out.*

**Reply 12** Reviewer #1 refers to the difference between offline simulations, where LSMs are forced with atmospheric variables and coupled simulations where LSMs are embedded into an atmospheric model. An important difference between these two setups is that offline simulations can be forced with historical estimates of atmospheric variables, which is not the case for coupled simulations. As our analysis relies on historic estimates of atmospheric forcing, the experiment suggested by the reviewer is not feasible. Nonetheless, we do not expect that coupled simulations would necessarily lead to different results, as long as one is focusing on the performance of the model(s) at large spatiotemporal scales.

**Comment 13** *In addition the observed (transformed) runoff has been used here as a benchmark directly and this is not usually a totally fair comparison as it is not in the 'model world'.*

**Reply 13** We would like to note that the comparison of observed runoff estimates to grid-cell runoff is a common practice which has been accepted by large parts of

C7716

the modelling community (see Gudmundsson et al., 2012a,b; Stahl et al., 2011, and references therein). Note also that similar comparisons are also done in atmospheric sciences where grid-cell scale precipitation and temperature are compared to weather station data. As considered small catchments actually cover substantial parts of the grid cells we are confident that the comparison between observed and modelled runoff is fair, at least at the considered temporal resolution (monthly).

**Comment 14** *I would have suggested that there should have been further comparisons with long-term model climatologies of runoff (routed discharge too) and/or reanalysis products which integrate observational data with the model structure.*

**Reply 14** We actually compare observed and modelled climatologies (see Reply 6). We also compare observed and modelled river discharge from large basins, where routing surely plays a role at a sub-monthly resolution (Figure 3d-g, Table C1)

**Comment 15** *Also the analysis with the reduced meteorological variable set seems rather tagged on at the end when really this forms quite an important part of the analysis as many models are only driven by precipitation and temp/evap etc. This again undermines the conclusions.*

**Reply 15** Albeit interesting, the comparison between the “full” and the “reduced” atmospheric forcing is not directly relevant for testing the CLPH. Therefore we have presented this as supporting information in the appendix.

**Comment 16** *“We note that issues common to all statistical applications can limit the interpretation of the presented results. Uncertainty in the used data, correlations between atmospheric forcing and land parameters, as well as an incomplete list of possible explanatory variables can influence the analysis”. Quite right - surely this should have been tested then - what is the sensitivity of your results to these? “these limitations do also imply that the effect of the considered land parameters on large scale features of terrestrial water dynamics may have a similar order of magnitude as the mentioned disturbing factors” - I don't understand how this can be justified as this has*

C7717

*not been evaluated.*

**Reply 16** Reviewer #1 highlights that our discussion of confounding factors was not clear enough and we will expand the related section along the following lines:

1. All considered data do unfortunately not have uncertainty estimates. Consequently a formal assessment of the influence of data uncertainty is not possible.
2. Although the correlation between land properties and atmospheric variables can be documented it is impossible to resolve the question which of both have the stronger influence phenomena we investigate. For example topography can influence precipitation patterns which should propagate to streamflow dynamics. Similarly topography is also expected to impact streamflow dynamics. If we observe differences in streamflow dynamics between flat and mountainous regions it is consequently impossible to decide whether these are related to changes in precipitation or related to differences in e.g. channel routing.
3. We acknowledge that the list of land properties we tested is incomplete. However, an exhaustive investigation of all available land parameters would lie beyond the scope of this study. We see, however, our study as a first step in this direction, which will hopefully also motivate other researchers to put the CLPH under scrutiny by confronting it with further land parameters.

**Comment 17** *So overall I find that although the manuscript has opened up a debate on complexity issues in land surface/global hydrologic modelling, which is a good thing, unfortunately the paper itself is not of a high scientific quality, the conclusions are not supported by the evidence presented and it is written in a way which is very complicated and difficult to understand.*

**Reply 17** We hope that our answers above could clarify the concerns of Reviewer #1.

C7718

## References

- Beven, K.: A manifesto for the equifinality thesis, *Journal of Hydrology*, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.
- Bishop, C. M.: *Pattern Recognition and Machine Learning*, Information Science and Statistics, Springer, 2006.
- Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F., and Koirala, S.: Comparing Large-Scale Hydrological Model Simulations to Observed Runoff Percentiles in Europe, *J. Hydrometeorol.*, 13, 604–620, doi:10.1175/JHM-D-11-083.1, 2012a.
- Gudmundsson, L., Wagener, T., Tallaksen, L. M., and Engeland, K.: Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe, *Water Resour. Res.*, 48, W11 504, doi:10.1029/2011WR010911, 2012b.
- Hastie, T., Tibshirani, R., and Friedman, J. H.: *The Elements of Statistical Learning – Data Mining, Inference, and Prediction*, Second Edition, Springer Series in Statistics, Springer, 2 edn., 2009.
- Stahl, K., Tallaksen, L. M., Gudmundsson, L., and Christensen, J. H.: Streamflow Data from Small Basins: A Challenging Test to High-Resolution Regional Climate Modeling, *Journal of Hydrometeorology*, 12, 900–912, doi:10.1175/2011JHM1356.1, <http://dx.doi.org/10.1175/2011JHM1356.1>, 2011.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, *Journal of Hydrometeorology*, 12, 823–848, doi:10.1175/2011JHM1369.1, 2011.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 13191, 2013.

C7719