Hydrol. Earth Syst. Sci. Discuss., 10, C7720–C7729, 2014
www.hydrol-earth-syst-sci-discuss.net/10/C7720/2014/

Hydrology and Earth System Sciences
Discussions
Open Access

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 #2 raises a number of questions regarding the employed methods and the structure the submitted 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: *This paper addresses an important problem, namely the relative contribution of climate drivers and model parameters to the uncertainty of large scale hydrological or land surface models. The authors compare the skill of CLPH, a model that predicts discharge based on atmospheric variables alone to the skill of a set of land surface and hydrological models that have a spatially variable representation of the land surface.*

Reply 1: We thank the reviewer to recognizing the importance of the addressed problem. We note, however, that we do not only compare the skill of the statistical CLPH model (Equation 9) to that of Land Surface Models (LSMs) but also to that of a corresponding statistical model including land parameters (Equation 10). This latter point is important as we explicitly assess the influence of selected land parameters (topography and soil texture). Although the list of land parameters is not exhaustive, this is a strong indication that the influence of these land parameters on large-scale terrestrial water dynamics cannot be distinguished from other factors.

Comment 2: *CLPH is based on gridded fields of 8 atmospheric variables and has 11 parameters that are found using the Random Forest Approach.*

Reply 2: There is a misunderstanding regarding the number of parameters. Random Forests have by construction a number of (not interpretable) parameters that is much larger than the number of input variables. The data driven models (Equation 9 & 10) accounts for n + 1 time lags of the eight input variables.

Comment 3: *The main conclusion is that land parameters are less important and that discharge can be skillfully estimated using atmospheric parameters alone.*

Reply 3: We also conclude that the influence of selected land parameters on large-scale terrestrial water dynamics could not be detected (see also Reply 1). We have rephrased the corresponding text to better highlight this point.

Comment 4: *From this the authors conclude that substantial progress in modeling and forecasting can be achieved but do not offer any details on how exactly this could be achieved. If the goal is to simulate mean monthly discharge, then this could be achieved using atmospheric variables alone (or time series models of discharge). The* *reason we are adding complexity to models is to study processes and interactions between them. This is particularly true for the impact of humans in the water cycle (not discussed at all by the authors). If the goal is to simulate mean monthly discharge, then this could be achieved using atmospheric variables alone (or time series models of discharge). The reason we are adding complexity to models is to study processes and interactions between them. This is particularly true for the impact of humans in the water cycle (not discussed at all by the authors).*

Reply 4: The scope of the study is not to provide an exhaustive theory on terrestrial water dynamics at large spatiotemporal scales. We do, rather, (1) document empirically to which degree terrestrial water dynamics can be captured using only a limited set of assumptions. Further, (2) we explicitly test the influence of selected land parameters on terrestrial water dynamics at the scale of implementation (see also Reply 1). For us, this empirical evidence is a motivation to further develop the theory of terrestrial water systems at large scales. However, as indicated also in our answer to Reviewer 1, some misunderstanding on the scope of our study may stem from our subtitle. Hence we have decided to simplify the title as follow: “Do land parameters matter in large-scale monthly terrestrial water dynamics?”

Regarding the last two sentence of this comment, we fully agree that studying the complex interactions between different processes is an important and interesting subject within earth system sciences. However, a prerequisite for the application of models implementing these processes is that they capture the relevant observable variables. To date, runoff and river discharge are among the best monitored variables of terrestrial systems. Unfortunately several state-of-the-art LSMs have issues in capturing key features such as seasonality (Gudmundsson et al., 2012b).

Regarding the last point, the thorough understanding of the natural processes governing terrestrial water dynamics on large scales is a prerequisite for understanding human impacts on the water cycle. Therefore we do not focus on human impacts, but rather base our investigation on near natural catchments (see Stahl et al. (2010) for

catchment selection criteria).

Comment 5: *The main conclusion of the paper is neither new nor surprising but has been discussed for a long time in the large scale hydrological modeling community (e.g. Fekete et al. 2004, Biemans et al., 2009).*

Reply 5: We strongly disagree with the reviewer on this point. Both studies mentioned by Reviewer #2 (Fekete et al., 2004; Biemans et al., 2009) are focusing on the validation of global scale precipitation products on climatological time scales (long term means) and the effect of their uncertainty on runoff/discharge climatologies at continental to global scales.

Our study goes well beyond these studies as we investigate the spatiotemporal dynamics of runoff at a much higher spatial and temporal resolution. The most important differences to the above mentioned studies are:

1. Model assessment at monthly resolution (including seasonal and inter annual variability) instead of (climatological) longterm mean monthly runoff.
2. Model assessment using observations at the grid-cell scale. This is at much higher resolution than the continental river basin scale considered in Biemans et al. (2009).

Consequently our investigation is a substantial progress compared to the studies mentioned by Reviewer #2. Nevertheless, both Fekete et al. (2004) and Biemans et al. (2009) are valuable for understanding the results of our analysis, as they document the massive uncertainties associated with large-scale estimates of atmospheric forcing. This in turn may be one of the reasons why we did not find an increase in model skill after explicitly including land properties in the statistical model (see also Reply 1). We have added text on this point in the revised version of the manuscript.

Comment 6: *Furthermore, the method used to arrive at the conclusion is not clearly* *described, partly confusing and a lot of important assumptions are buried in the appendices (see specific comments below). Overall, the paper Is not well written (too much theoretical background in the introduction that is not relevant of the paper, unclear description of the methods and no conclusion that helps addressing the uncertainties in large scale models).*

Reply 6: We acknowledge that some aspects of the methodology were not described in sufficient detail and we therefore have clarified the text in the revised version of the manuscript. The level of detail in the first sections of the article did receive variable feedback from the reviewers. This suggests that it is relevant for some parts of the target audience.

Specific comments/questions:

Comment 7: *Do all WATCH models really use all climate variables ? If not, are you comparing apples and oranges when some models use temperature and precipitation alone and others (including the CLPH) use the whole set ?*

Reply 7: In theory it would be optimal if all considered models would take the same atmospheric forcing variables into account. However, this is not the case for the WATCH ensemble. This reflects some parts of the uncertainty in model formulation, and the fact that different modelling groups have contrasting priorities with respect to the inclusion/exclusion of certain processes. Consequently we do not compare *“apples and oranges”* as suggested by the reviewer.

We did also assess the effect of reducing the forcing variables to precipitation and temperature (Appendix B1.3). The results (Appendix B2.3, Figure B1) show that this has a clear, but small influence on the statistical model. These differences, however do not impair our conclusions. (As this analysis is only a stability check, we decided to present it in the Appendix).

Comment 8: *Furthermore, all the LSM model will have parameters that control the water balance balance. These will not be calibrated whereas the CLPH-RF model simulates the water balance with 11 time lag parameters that are found through an optimization process (“training on ∼400 catchments in Europe). Is this really a fair comparison ?*

Reply 8: It is true that LSMs have many parameters and the current modelling practice attempts to identify these parameters from mapped land properties, such as soil types. Consequently LSMs undergo no, or only very limited, calibration. (In the considered model ensemble WaterGAP (Hunger and Döll, 2008) is an exception to this practice). In contrast to LSMs the data driven models in this study depend heavily on calibration with implications for model assessment and validation, as noted by Reviewer #2.

An important feature of any statistical modelling is therefore the procedure used for model testing. Here we use cross-validation (see page 13203, lines 9 - 10) to quantify the skill of the statistical models at locations that where not used for calibration. Although cross validation is to date not widely used in hydrology, it is a standard practice with well understood properties in advanced statistical modelling (see e.g. the textbooks of Hastie et al. (2009); Bishop (2006) for details). Cross validation is one of the most rigorous approaches for testing statistical models. Therefore we argue that the comparison between the statistical models and the LSMs is made as fair as possible.

As the reviewer’s comment indicates that we did not explain cross validation in sufficient detail we will expand the relevant section in the manuscript.

Comment 9: *What criteria were used to select the 400 catchments ?*

Reply 9: The considered catchments are taken from a previous study (Stahl et al., 2010), which also documents the selection criteria. (See also Reply 4).

Comment 10: *What is the area distribution of those catchments ?*

Reply 10: The size distribution of the considered catchments can be found in Gudmundsson et al. (2012a), Figure 2. The median catchment size is 258 km^2, which is below the average size of a grid-cell ($\approx 2500\text{km}^2$). We provide more details on this point in the revised version.

Comment 11: *Are all catchments large enough to justify the application of a model with a resolution of 0.5 degrees ?*

Reply 11: We acknowledge that we did not describe the allocation of catchments to grid cells (see p. 13201, lines: 21 ff.) clearly enough in the previous version of the manuscript and have expanded the text accordingly:

The catchments considered are on average below the size of a 0.5° grid-box in mid latitudes. Therefore we did assign each catchment to the grid-cell containing the coordinates of the gauging station. Owing to the high station density in central Europe, many of the considered grid cells contained more than one catchment. In these cases, the average streamflow of the two catchments was considered. To reduce biases, stemming from possibly different catchment sizes, the area weighted average was used. We refer to this quantity as "observation-based runoff estimates". We interpret this quantity as the average amount of water draining from the land unit covered by the grid cell. Consequently this quantity can be used to analyse terrestrial water dynamics at spatiotemporal scales where phenomena governed by the water balance are considered.

Although it may seem counter intuitive, to compare streamflow from small catchments to runoff from much larger grid cells, we argue that this is a valid approach if the following points are considered:

1. As daily runoff dynamics depends on small scale catchment processes (e.g. channel routing) we do not expect modelled and observed runoff to match at a daily resolution. We do, however, expect that coarser resolution (e.g. monthly) phenomena are comparable at this spatial scale (see also discussion on the separation of scales).

2. The employed approach is common practice in the evaluation of large-scale hydrological models (see Gudmundsson et al. (2012a,b) and references therein)
3. The employed approach is comparable with approaches used in climate- and atmospheric sciences, where weather stations are often assigned to grid cells in a similar fashion.
4. The validation of the statistical model trained to estimate grid-cell scale runoff with respect to discharge from nine large rivers in Europe (Figure 3d-g, Table C1) further suggests that the approach is feasible.

Comment 12: *How are these basins impacted by human activities ?*

Reply 12: Only catchments with minimal human impacts have been selected. See Stahl et al. (2010) for details on catchment selection.

Comment 13: *What was used for comparison (p. 13201) ? Basin values ? Grid cell values ? This is very unclear. What is meant by "if more than one gauging station occurred in one catchment the area weighted average runoff rate was used".*

Reply 13: See Reply 11.

Comment 14: *Why not use the upstream area of the gauging station ?*

Reply 14: Unfortunately there is no standard, database with boundaries for all considered catchments available. As the catchments are on average smaller than the grid cells this approach seams to be feasible (see also Reply 11 & 13).

Comment 15: *Some of the appendices contain essential information for understanding the methods, I would suggest to merge this so that the manuscript is easier to read.*

Reply 15: Thank you for this suggestion, we have considered this in the revised version of the article.

Comment 16: *The discussions about scales (sec. 2) and the constant land parameter hypothesis is rather lengthy and not really relevant to the rest of the paper.*

Reply 16: We rely on the discussion of scales (Section 2) for the derivation of the CLPH (Section 3) which is the core of our study. (See also Reply 1 and Reply 6).

References

Biemans, H., Hutjes, R. W. A., Kabat, P., Strengers, B. J., Gerten, D., and Rost, S.: Effects of Precipitation Uncertainty on Discharge Calculations for Main River Basins, J. Hydrometeor, 10, 1011–1025, doi:10.1175/2008JHM1067.1, http://dx.doi.org/10.1175/2008JHM1067.1, 2009.

Bishop, C. M.: Pattern Recognition and Machine Learning, Information Science and Statistics, Springer, 2006.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O., and Willmott, C. J.: Uncertainties in Precipitation and Their Impacts on Runoff Estimates, J. Climate, 17, 294–304, doi:10.1175/1520-0442(2004)017<0294:UIPATI>2.0.CO;2, 2004.

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. Hydrometeor, 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.

Hunger, M. and Döll, P.: Value of river discharge data for global-scale hydrological modeling, Hydrology and Earth System Sciences, 12, 841–861, doi:10.5194/hess-12-841-2008, 2008.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow trends in Europe: evidence from a dataset

of near-natural catchments, Hydrology and Earth System Sciences, 14, 2367–2382, doi:10.5194/hess-14-2367-2010, 2010.

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