Hydrol. Earth Syst. Sci. Discuss., 11, C6384–C6394, 2015
www.hydrol-earth-syst-sci-discuss.net/11/C6384/2015/

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "Towards observation based gridded runoff estimates for Europe" *by* L. Gudmundsson and S. I. Seneviratne

L. Gudmundsson and S. I. Seneviratne

lukas.gudmundsson@env.ethz.ch

Received and published: 5 February 2015

We would like to thank Reviewer # 2 for reporting his/her concerns. 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: *Like Balazs Fekete, I was greatly interested when I received this paper, and like him, I have been surprised by the content. As I far as I am concerned, I have been disappointed. The main problem is that the paper does not deliver what the title promises, namely gridded runoff estimates based on observations over Europe. The paper did not convince me either to achieve a significant step toward this goal.*

Reply 1: We regret that Referee #2 did not find find our manuscript convincing. We hope that the following reply to his/her concerns will help to clarify our approach and the points of discussion.

Comment 2: *From my point of view, an important reason stems from the ambiguity between runoff and streamflow in paper, where they are used as equivalents, although they are very different quantities, with different units (typically mm/s vs m3/s), and different dynamics, since streamflow results from the routing of runoff in watersheds, along hill slopes, under ground, and in the river network. What the authors are using as their reference for runoff is streamflow divided by the contributing area. This certainly leads to a runoff unit, but not to a runoff dynamic, as the operation does not deconvolute the effect of transfer times within the catchment. The paper would greatly benefit from clarifications about these concepts, and from some reviewing of the various attempts to estimate runoff based on streamflow observations, not overlooking the targeted timescales. In particular, I raise the attention of the authors to Gottschalk and Sauquet, who addressed these questions in quite many papers (e.g. Sauquet, E., Gottschalk, L., and Leblois, E., Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme, Hydrological Sciences Journal, 45, 799-815, 2000).*

Reply 2: While we agree that there are differences between streamflow and runoff, we disagree with the reviewer's evaluation that our approach is flawed:

1. The main difference between streamflow and runoff is that the former has been routed trough a chanel network. However, the associated processes operate on time scales that are much smaller than the resolution of the presented analysis. Hydrograph wave speeds are approximately 0.5 m sec^{-1} $= 1.8$ km h^{-1} $=$ 43.2 km day^{-1} (e.g. Wong and Laurenson, 1983), implying that at least daily resolution would be required to resolve these processes for 0.5° grid-cells. As our study operates on a monthly resolution on a 0.5° grid (≈ 50 km), we are confident that the effects of channel routing do not impair our results.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2. We thank Referee #2 for pointing us to the study of Sauquet et al. (2000). However, the setting of the study of Sauquet et al. (2000) is not comparable with that of the present manuscript as it is concerned with spatial interpolation of monthly runoff climatologies within the constraints of a high resolution channel network. In this context we are also surprised that Referee #2 did not point us to the more recent successor papers (e.g. Skøien et al., 2006; Skøien and Blöschl, 2007; Laaha et al., 2013), who formally refined the method of Sauquet et al. (2000).

3. While we agree that the above mentioned techniques (Skøien et al., 2006; Skøien and Blöschl, 2007; Laaha et al., 2013) have great potential for interpolating runoff at relatively small spatial scales, we are not aware of studies that applied this approach to estimate long (decadal) time series in large (continental scale) regions with a sparse observation network (see e.g. Blöschl et al. (2013) for a comprehensive synthesis report). Therefore it is difficult to judge upon the potential of these methods for estimating runoff at continental scales.

4. While the above mentioned approach (Skøien et al., 2006; Skøien and Blöschl, 2007; Laaha et al., 2013) is promising we decided in the present study to follow the approach of Jung et al. (2009, 2010, 2011) who used machine learning to estimate monthly evapotranspiration on the basis of gridded predictor variables. We are convinced that both approaches can yield reliable estimates, but a formal evaluation of their respective strengths and weaknesses lies beyond the scope of the presented paper.

Comment 3: *The above point is essential when comparing the "runoff estimates" to LSM products, which are runoff stricto sensu. Grid-cell runoff in a LSM is the spatial average over the grid-cell of point-scale runoff in mm=kg/m2, without any transfer function to the river network, so that it does not have the same dynamic as the selected streamflow data (even if the contributing catchments are smaller than the 0.5° $\times$ 0.5° LSM grid-cells). I believe this is a major reason for the large differences between "runoff* *estimates" and LSM runoff in the paper, so that these differences may not invalidate the LSMs, but rather the proposed "runoff estimates" as the appropriate data against which assessing LSMs.*

Reply 3:

1. The first author has discussed this issue with experts from the LSM community for previous assessments (Gudmundsson et al., 2012a,b). Note, e.g. that the majority of the co-authors of Gudmundsson et al. (2012a) are experienced LSM developers who all agreed with the approach to use grid-cell averages of streamflow from small catchments as observational runoff estimates.

2. See also our comments above.

Comment 4: *Another problem is the representativeness of the selected streamflow stations/catchments to constrain gridded runoff estimates over Europe by extrapolation. The streamflow measuring stations mostly come from mountainous areas, Germany, and UK. Non mountainous Mediterranean climate is almost absent, except a couple of points in Spain. The performances are not differentiated based on hydro-climatic regime, and Fig. 8, the only one where it might be possible to assess the extrapolation power of the RFM in ungauged areas (I mean here areas with different hydrologic regimes than the ones of the selected streamflow stations, i.e. areas on which the RFM has not been trained), is almost useless: scatter plots where the color points corresponding to the different large-scale river basins cover each other, colors that are difficult to distinguish, log scale hiding the differences, no average summary by basin. All this casts doubts on the main conclusions: (i) validity of a unique RFM for extrapolating "runoff" over Europe, (ii) negligible effect of land properties (they might have a larger effect in water stressed areas, in which the RFM is not trained).*

Reply 4:

HESSD

11, C6384–C6394, 2015

1. We agree that it would be beneficial to include more catchments for model training, especially in the Mediterranean region. Unfortunately neither the main data base of the GRDC, nor the European Water Archive currently hold runoff time series that extend to 2000 and fulfil the quality requirements of Stahl et al. (2010).

2. It is indeed interesting to assess whether model performance is dependent on climate conditions. To approach this we have conducted a correlation analysis, relating the spatial patterns in model performance to the mean climatic conditions (annual means of runoff, precipitation and temperature). Overall the results (Figure A1) indicate that there is little influence of mean climate on model performance (all correlations being $|r| < 0.5$). Nevertheless Figure A1 suggests that there is some dependence of the relative bias on mean annual runoff. In addition Figure A1 suggests a possible link between mean temperature and Anomaly-R^2. As these results help to clarify both strengths and weaknesses of the empirical model we suggest to include Figure A1 in the revised manuscript; alongside with an appropriate discussion.

3. It is not true that we do not provide a detailed basin-scale validation. Unlike suggested by Referee #2 Figure 8 is not the only information provided. Figure 9 shows the value of six performance metrics for each of the considered river basins. If needed we could include the numerical values underlying Figure 9 in the appendix or in an supplement.

Comment 5: *Other comments regarding the methods: (1) Why not test the effect of land use and vegetation in the model setup? (2) R2 is not the correlation coefficient but the determination coefficient, with a major difference, since R2 cannot be negative. (3) I'm not sure the cross-validation procedure really leads to an independent evaluation, since different sub-periods in one stream flow record are not independent (we would used paired statistics to compare them). Random sub-samples of different stations may also show some dependence given the hydroclimatic similarities raised above.*

Reply 5:

1. The main aim of this study was to suggest a framework which can be used to estimate gridded runoff at ungauged locations. In the presented manuscript we already tested the framework extensively and accounted for differences in atmospheric forcing and selected land parameters. We agree that assessing the influence of more land parameters on runoff dynamics within this framework would be of great interest. This, however would be a major independent research effort which goes beyond the scope of this study. (Ideally one should test the effect of all parameter maps that go into typical LSMs).
2. We thank the reviewer for noting the unprecise terminology used for R^2 (although, of course, R^2 is equal to the square of the correlation).
3. It is true that other sampling schemes may reveal strengths and weaknesses of the presented approach more clearly, especially for large climatic gradients. Nevertheless, we would like to emphasise that we already present an extensive validation focusing on (1) prediction at ungauged locations (CV in space); (2) prediction of ungauged times (CV in time); (3) prediction of continental river discharge [see also the comment above] and (4) prediction of Evapotranspiration climatologies.

Comment 6: *Results: (1) The sections on model selection and validation are very short: no discussion of the parameters values (time lags); Fig 4 is a very condensed summary of the three models' performances: couldn't we compare the pdf of the local RMSEs for instance? Fig. 7 suggests poorer performances in mountainous areas, which is not commented. (2) Still regarding validation: why not comparing to an independent estimate, such as provided by Fekete et al. (2002)?*

Reply 6:

HESSD

11, C6384–C6394, 2015

Interactive Comment

1. An inherent disadvantage of machine learning techniques such as Random Forests is that they often produce parameter sets that are not interpretable. It is therefore common practice that these techniques are evaluated in terms of their predictive performance (which we do) and not in terms of parameter values. (See e.g. the excellent text book of Hastie et al. (2009)).

2. It is not true that Figure 4 is the only information on the performance of the three considered models. The box plots in Figures 6 and 7 (orange boxes) show the error distribution of the three models for both cross validation experiments and all considered performance metrics. In addition Table 1 provides median grid-cell performance of the Random forest model with full atmospheric forcing.

3. It is true that Figure 7 indicates somewhat lower values of Anomaly-R^2 in Scandinavia and mountainous area. This is consistent with the result shown in Figure A1, which show that Anomaly-R^2 is positively correlated with mean annual temperature. We suggest to discuss this alongside the interpretation of the correlation analysis as suggested above.

4. We have assessed the performance of the data product suggested by Mr. Fekete. In general its performance is in line with the WATCH LSMs. (See reply to Mr. Feketes comments for more details).

Comment 7: *Eventually, I find abusive the conclusion that the selected RFM is capable, performs reasonably, compare well with observations, is an ideal candidate for model evaluation (all these expressions come from the submitted paper), and I don't recommend the publication of this paper in its present form.*

Reply 7: We hope that our replies to the reviewer's concerns did help to resolve the main issues highlighted in his/her review.

References

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H., eds.: Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales, Cambridge University Press, 2013.

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, New York, 2 edn., 2009.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001–2013, doi:10.5194/bg-6-2001-2009, 2009.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951 – 954, doi:10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res., 116, G00J07, doi:10.1029/2010JG001566, 2011.

Laaha, G., Skøien, J. O., Nobilis, F., and Bløschl, G.: Spatial Prediction of Stream Temperatures

Using Top-Kriging with an External Drift, Environmental Modeling & Assessment, 18, 671–683, doi:10.1007/s10666-013-9373-3, 2013.

Sauquet, E., Krasovskaia, I., and Leblois, E.: Mapping mean monthly runoff pattern using EOF analysis, Hydrology and Earth System Sciences, 4, 79–93, doi:10.5194/hess-4-79-2000, 2000.

Skøien, J. O. and Blöschl, G.: Spatiotemporal topological kriging of runoff time series, Water Resources Research, 43, W09 419, doi:10.1029/2006WR005760, 2007.

Skøien, J. O., Merz, R., and Blöschl, G.: Top-kriging - geostatistics on stream networks, Hydrology and Earth System Sciences, 10, 277–287, doi:10.5194/hess-10-277-2006, 2006.

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.

Wong, T. H. F. and Laurenson, E. M.: Wave speed-discharge relations in natural channels, Water Resources Research, 19, 701–706, doi:10.1029/WR019i003p00701, 1983.

HESSD

11, C6384–C6394, 2015

Figure Captions:

Fig. A1. Correlation between grid-cell level performance metrics of the Random forest model with full atmospheric forcing and mean climatic conditions (Q: mean annual runoff; P: mean annual precipitation; T: mean annual temperature). Horizontal lines at $r = \pm 0.25$ and $r = \pm 0.5$ are included as a visualisation aid. Details on the performance metrics can be found in Section 3.3.3 of the discussion paper. Spatial patterns of the performance metrics are shown in Figure 6 and 8 in the discussion paper.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 12883, 2014.

HESSD
11, C6384–C6394, 2015

Interactive Comment

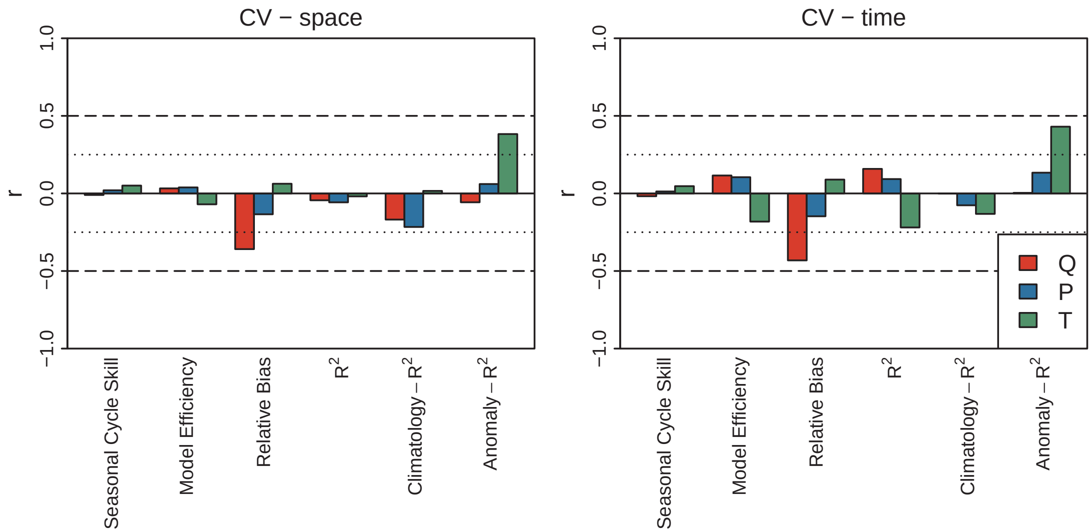

Fig. A1.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper