Earth Syst. Sci. Data Discuss.,
doi:10.5194/essd-2015-38-AC2, 2016

Open Access Earth System Science Data Discussions

Interactive comment on "Observational gridded runoff estimates for Europe (E-RUN version 1.0)" *by* Lukas Gudmundsson and Sonia I. Seneviratne

Lukas Gudmundsson and Sonia I. Seneviratne

lukas.gudmundsson@env.ethz.ch

Received and published: 19 April 2016

We would like to thank G.V. Ayzel for the open and very positive evaluation of our manuscript. In the following we provide point by point answers to his comments. For the sake of clarity we first repeat the reviewer's comments (*in italic*) and then provide our response.

Comment 1:
As noted in the first referee comment "The beauty of first publishing the method and results... that the discussion and review comments of GS15 are publicly available" and it's totally true for this article too – following previous article history helps us to understand general concepts and, in addition, each referee can focus on the own key points of presented work without any intersections.

Printer-friendly version

Discussion paper

I am really impressed by the described database construction and provided description of the workflow – it is one another significant step to the modern science with shareable data, community models and reproducible research.

For this interesting and useful article and dataset I would recommend such minor revisions related to deeper analysis around a key questions I marked.

Reply 1:
We would like to thank G.V. Ayzel for this overall positive rating of the presented manuscript.

Comment 2:
General comments: Machine learning model cross-validation in space and time is a valuable tool for accessing performance fluctuations in respect with changing conditions, and for this work it was done well. Nevertheless there are two points I marked as unclear and required more detailed explanation.

1. The most dataset points were derived from the Random Forest model (RFM) implementation to E-OBS forcing data, but only a few points were used to model training and testing – and I guess that performance of RFM roughly decreases for eastern part of Europe (e.g. Russia) because of no observational data were account in the model. It is a theme of a great interest and I recommend describe it more clearly – it may have an impact of the decision to use provided dataset for Eastern Europe domain.

Reply 2:
Thank you for noting that the accurracy of our estimates may decrease with station coverage. We will discuss this as an possible limitation in the revised version of the article.

Note also that we managed to increase station coverage in eastern Europe by including station data that are only available in monthly resolution.

Comment 3:

2. This question is mainly related to the GS15 [1] article, but it is the main question for my own interest too. What is the reason of such significant differences between LSM- and RFM-generated monthly runoff? Our recent results (SWAP model runs for ISI-MIP basins) show monthly NSE about 0.8, and I think other models demonstrate the same efficiency. Why should we use E-RUN instead of the averaged LSMs outputs for large-scale case studies?

Reply 3:
The difference between the RFM estimates and the WATCH LSM ensemble was described and discussed in GS16. Note also that such a discussion is not within the scope of ESSD.

In addition we would like to emphasise that any estimation technique has its specific strengths and weaknesses. As a consequence we belief that it is important to develop "ensembles" of estimates that rely on different techniques (e.g. statistical tools vs process based models). This is already common practice in atmospheric sciences, where both statistical interpolation as well as model driven reanalysis are used to provide reconstructions of weather and climate.

Comment 4:
Specific comments:

Title. Actually it is combined (observational plus RFM-predicted) dataset;

Reply 4:
Following the suggestions of Reviewer #1 we will change the title to "Observation-based gridded runoff estimates for Europe (E-RUN version 1.0)"

Comment 5:
Abstract, P1, L1: "hydrological cycle variable" instead of "climate variable";

P1, L13: same comment as previous one;

Reply 5:

As already noted in our reply to Reviewer #1 runoff is listed as an “essential climate variable” (see e.g. Bojinski et al., 2014).

Comment 6:
P8, L20: average number is not useful metric, histogram of assigned stations per cell will be suitable;

Reply 6:
Thank you for this suggestion, we will consider to include such a figure in the revised manuscript. Note, however, that we try to balance the number of figures with the length of the overall text.

Comment 7:
P9, L 29: model efficiency is the same metric with Nash-Sutcliffe-Efficiency;

Reply 7:
See our reply to the corresponding comment of Reviewer #1.

Comment 8:
It is necessary to attach figures in high quality to the supplement section.

Reply 8:
We will consider to provide high-resolution versions of the figures in the supplement.

Comment 9:
Data availability assessment: Provided link to the dataset hosted by PANGAEA is not stable, there is the reason to make some reserve open data storage. Reading, testing and visualization was done by the NETCDF4 and Basemap libraries for python (ver.3.5) without any errors.

Reply 9:
We are glad to read that the data can be accessed without problems.

References

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M.: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy, Bull. Amer. Meteor. Soc., 95, 1431–1443, doi:10.1175/BAMS-D-13-00047.1, 2014.

Interactive comment on Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2015-38, 2016.