Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-532-RC1, 2020

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "Daily ensemble river discharge reforecasts and real-time forecasts from the operational Global Flood Awareness System" *by* Shaun Harrigan et al.

Anonymous Referee #1

Received and published: 22 November 2020

Summary

This paper describes the most recent version of the GloFAS ensemble streamflow forecasting system. While there are no major advanced in methods used to generate forecasts, GloFAS is a system of international significance, and highly relevant to readers of HESS. The manuscript is well structured, admirably clear and succinct, and was a pleasure to read. Figures are well presented, and while references are sparse (especially in the introduction), as this paper is essentially focused on presenting an operational system this is ok. As the authors note, a major development is the availability of GloFAS forecast outputs in near-real time, and this is well-explained and documented.

Printer-friendly version

Discussion paper

I therefore believe the study ultimately deserves publication. I nonetheless had two major issues with this paper, listed below. I therefore recommend the paper be revised before it can be published.

Major comments

1) There appeared to me to be an error in the calculation of CRPSS with respect to (wrt) persistence - see specific comments below. If this is not due to an error, I would like the authors to explain what to me were counterintuitive results.

2) The authors earmark the assessment of reliability to future work. I do not think this is good enough, given 1) that reliability is a key attribute - in my view at least as important as skill - of ensemble forecasts and 2) their statement in the introduction that "not also having direct access to the raw data precludes the use in further downstream applications (e.g. impact modelling, multi-model forecast systems, production of value-added products for specific sectors such as river transport and hydropower industries, and advancement in techniques requiring large-scale datasets such as machine learning)." This statement implies that the authors expect the outputs in the ways specified - i.e. as direct inputs to impact assessment models of some kind or other. In my experience such models very often require reliable ensembles wrt to observations (or at least unbiased ensembles) as inputs. As GloFAS does not treat hydrological uncertainty, it is highly likely that ensembles are overconfident, particularly at short lead times (e.g. Bennett et al. 2014). I think this is information that users of these outputs, and therefore readers of this paper, would want to know. I therefore would like to see the authors present an assessment of reliability as well as skill, and the ramifications of this assessment discussed. Given the forecasts are likely to be treated as continuous variables in impact models, I suggest using probability integral transforms (PIT, e.g. Gneiting and Katzfuss 2014) to assess reliability (noting the need to generate 'pseudo'-PIT values in cases where streamflow observations can equal zero). If the authors prefer, PIT values can then be summarised with either the alpha-index (Renard et al. 2010) or the beta-score (Keller et al. 2011) (whichever is more suitable) for presentation in plots

similar to Figure 5 or 6.

Specific comments

L88-97 Please provide the model time step at some point in this paragraph.

L125 "https://www.globalfloods.eu/" the hyperlink associated with this text 1) differs from the text and 2) returns a 404 error.

L250 Figure 5. To me, there's something very counterintuitive (and perhaps erroneous?) about the persistence skill plot. The accuracy of persistence (the benchmark, and the demoninator in eq 1) is often very high at short lead times and then declines with lead time - often rapidly. In my experience, this decline is usually much faster than the decline in the accuracy of forecasts. So I would expect CRPSS wrt to persistence to be very low - perhaps even close to 0 - at very short lead times, and then to rise with lead time. But Fig 5 shows the opposite of these trends - i.e. CRPSS wrt persistence starts high and falls with lead time. I can't see how this can occur without a calculation error - though perhaps I've missed something? Even if this is not due to an error, these results at least requires some discussion/explanation. CRPSS calculated wrt to climatology looks sensible to me, which makes the persistence results even more puzzling.

L271 Figure 6 As with Fig 5, I would expect skill wrt to persistence to rise with lead time, not to fall.

L343-345 "Future work should assess other aspects of forecast quality such as reliability (Robertson et al., 2013), value (Cloke et al., 2017) or performance during extreme events (Bischiniotis et al., 2019)." Not suggesting any change here, but the authors may also like to consider calculating the skill/reliability of accumulated volume forecasts (e.g. accumulated 30-day streamflows), as this may well be of interest to reservoir operators and others. The ability to simply sum streamflows of individual ensemble members over various lead times is a major benefit of ensemble streamflow forecasting systems such as this one (as opposed to probabilistic forecasts generated at discrete lead times).

Typos/grammar/style

L79 "descripted" - 'decribed'?

L140-141 "see for https://confluence.ecmwf.int/display/COPSRV/01.+GloFAS+operational+system a description" should be "see https://confluence.ecmwf.int/display/COPSRV/01.+GloFAS+operational+system for a description

L152-155 "Twice per week ... as real time (Vitart 2014)." Suggest breaking this long sentence in two at the comma.

References

Bennett JC, Robertson DE, Shrestha DL, Wang QJ, Enever D, Hapuarachchi P, Tuteja NK. 2014. A system for continuous hydrological ensemble forecasting (SCHEF) to lead times of 9 days. Journal of Hydrology 519: 2832-2846. DOI: 10.1016/j.jhydrol.2014.08.010.

Gneiting T, Katzfuss M. 2014. Probabilistic forecasting. Annual Review of Statistics and Its Application 1: 125-151. DOI: 10.1146/annurev-statistics-062713-085831.

Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW. 2010. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resources Research 46: W05521. DOI: 10.1029/2009wr008328.

Keller JD, Hense A. 2011. A new non-Gaussian evaluation method for ensemble forecasts based on analysis rank histograms. Meteorologische Zeitschrift 20: 107-117. DOI: 10.1127/0941-2948/2011/0217.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-532, 2020.