Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-584-RC1, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Performance of ensemble streamflow forecasts under varied hydrometeorological conditions" *by* H.-J. F. Benninga et al.

Anonymous Referee #1

Received and published: 16 December 2016

General comments

This manuscript presents an interesting analyse of the performance of hydrological ensemble predictions. The skills are screened according the regime (low and high streamflow) and the generating processes (snow melt, short rain, long rain floods etc.). This study further disentangles hydrological model errors and errors from meteorological forcing. The methodology is applied to a mountainous catchment. The combination of existing methodologies is pertinent and is worth being published in HESS.

However the reading is not easy and a major revision is necessary. Some information is redundant in the introduction, methodology and results sections and long lists of references are not always necessary. The focus should be made on the main contribution

Printer-friendly version

Discussion paper

of the paper i.e. the analysis of the skill for different hydro-meteorological conditions and skip or shorten secondary experiments. Some validation methodologies are described but their results are not shown. A balance should be found: either shorten the description or include those results. Some suggestions are given in the specific comments. The English should be improved.

The authors are using ensemble predictions from ECMWF from 2007 to 2013 with a training of the pre- and post-processing during two water years between 2011 and 2013. They associate the failure of the quantile mapping for post-processing method to the short time series of forecasts for training and to the inconsistency of the bias between the training and the validation period. They forget that the ensemble prediction system has undergone many changes during this period including spatial resolution changes. This is why retrospective forecasts are available since long and provide samples of 18 to 20 years back for post-processing purposes. Re-forecasts have been widely used and reported in the literature. These meteorological re-forecasts have also been used for the preparation of hydrological re-forecasts for the statistical post-processing of hydrological ensemble predictions.

Figure 5 to 9 are the core of the paper. They will gain value if the plots are associated with confidence intervals.

The use of the term “perfect forecast” is questioned because it is neither a forecast nor perfect and, would the future meteorological forcing be known, predictions with the model would include growing errors due to initial conditions as somehow shown in Table 5.

Specific comments

P1, L20-24 Should be rephrased e.g. too many occurrence of “improve”.

P3 L23-P4, L3 How do you correct measurement? Do you correct each station for the difference between the elevation of the station and the average of the elevation

Printer-friendly version

Discussion paper

in the area defined by the intersection of the Thiessen polygon corresponding to the station and the watershed? Then average the corrected values of the stations using their relative contribution to the catchment area as weights?

P5, L20-21 Equations would be appropriate here in order to define Y, NS and E_RV.

P5, L28 preceding the first forecast day.

P5, L32-P6, L2 I would suggest to skip this experiment or, if impossible to skip, tell already that it failed (according to P11, L3-5). This is to lighten the methodologies to keep in mind until the result section.

P6, L31-P7, L11 Some information (and references) is redundant with the sub-sections.

P7, L1 Three properties of probabilistic forecast quality . . .

P7, L8 “The histograms accompanying ...” the histograms of what?

P7, L20-21 “CRPS approaches the average value of the evaluated variable” What do you mean with “approaches”?

P7, L24-27 “and compares the forecasts with a relevant alternative forecast” somehow redundant with the beginning of the sentence.

P8, L1-2 “... argue that this” choice . . . these two lines should be rephrased. I would prefer a positive phrasing saying that the choice of another alternative forecast may result in a more robust estimation of forecast skill.

P8, L22-23 Either provide an equation for the “numerical indicator delta” if it adds to the understanding of the adopted methodology or skip any reference to delta.

P8, L30-31 “... contain a random element ...” explain how it works for the flatness coefficient.

P9, L3 “... for a certain event ...” It would be useful to define “event” and refer to sub-section 3.4 or Table 1.

P9, L24-28 Almost the same thing is repeated.

P9, L29 At a first reading, it was tempting to replace this ratio with a CRPSS of sim against meas but the purpose is different and since it is a major tool in this paper, this paragraph should be written with much care.

P10, L11-12 Are the rules given also by Merz and Blöschl or defined for this catchment based for instance on data from both simulation and observations during the training period?

P10, L16 Do you mean that the distribution of the generating processes shown in the figure is like we can expect for this region?

P10, L19, Table 3 What is the rule for precipitation deficit?

P11, L21 “preceding day” the day before the forecast issuing day.

P11, L28-29 “not shown in the paper ” therefore, going back to section 3.1.3, the methodology description should be simpler and not encumber with strategy numbers.

P10 L20 What do you mean by “reliable distribution”?

P12, L13 with more skill instead of “skilful”

P12, L16 “functional” what do you mean?

P12, L28 “... are in general less predictable by historical measurements ...” please re-phrase

P12, L32 “not shown” a figure is missing with the rank histograms for the low streamflow forecasts and for the high streamflow forecasts, two lead times. Apparently, for high flow, the rank histogram is not exactly U-shaped but skewed according to P13, L12-13.

P13 L10-13 Difficult to figure out . . . Please add a figure with the reliability diagrams and corresponding sharpness histograms for the low streamflow forecasts and for the high streamflow forecasts two lead times.

P13 L15-17 Note that good sharpness without reliability is useless.

P13, L18 reference already given, please re-phrase.

P14, L11-13 "... the below zero skill . . . do not result in positive skill ..." ...

P14, L23 What is the amount of this fake drizzle?

P14, L24-26 Re-phrase: "... meteorological forecasts accumulated in the forecasting system are better model inputs ..."

P15, L8 & Figure 10 I would skip this figure which highlights the weakness of drawing such a detailed profile with just a water-year data. The legend is missing for the thin plain lines.

P16, L8-10 Do you have evidence that such coincidence occurs and is the main explanation for the high ratio for short-rain floods?

P17, L13-15 "longer time series of forecasts", "longer forecasts datasets" see general comments; "more sophisticated" and first of all more robust.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-584, 2016.

Printer-friendly version

Discussion paper