**Author Response to RC#2**

Thank you for the positive evaluation of our article. We appreciate the feedback that will contribute to improving the manuscript.

We would like to apology for the missing references. The error emerged when we specified the HESS format, and un-intentionally deleted many references from the reference list. The main author should nonetheless have detected this flaw prior to posting.

Replies and corrections are done as follows: the Author response (AR) is marked with red text, while the author's suggestions to corrections (AC) are marked with blue text. All Referee comments are kept in a black; we use page and line number when needed to specify the appropriate location.

## Interactive comment on "Streamflow forecast sensitivity to air temperature forecast calibration for 139 Norwegian catchments"

By Trine J. Hegdahl et al.

**Anonymous Referee #2**
Received and published: 15 November 2018

This manuscript presents analyses of the sensitivity of streamflow forecasts to air temperature forecast calibration. The manuscript is well written, well structured, and I only have a few minor comments to the presentation, most of them just edits.

I find the description of validation scores and evaluation scores in 3.2 somewhat short. The section could give a better description of the rank histograms, and what is actually meant by the different shapes. And what is meant by slope and convexity being "negatively oriented"? Something seems wrong with the last sentence.

AR: Thank you for bringing this to our attention. We will provide some more details in the description of the rank histograms. By "negatively oriented", we mean that lower values (of slope and convexity) are better (i.e. more reliable forecasts). We will revise the sentence to better explain the meaning of "negatively oriented", and rephrase the last sentence to make it clear.
AC: We will apply the following changes:

- "For reliable ensemble forecasts, the rank-histogram will be uniform (horizontal). A bias in the ensemble forecast is recognized as a slope in the rank-histogram, where a negative slope indicates over-estimation by the forecasts (and vice versa). A U-shape indicates that the ensemble forecast is under-dispersed whereas a convex shape indicates over-dispersion (Hamill, 2001).
- Negatively oriented: new "… and convexity are negatively oriented, (i.e. lower values are better), and with an optimum value of zero …."

P2L5 three main componentS?
AR: Thank you, will be corrected.
AC: 'component' replaced by 'components'

P2L14 Langsrud et al, 1998a and 1998b are missing from reference list. What kind of statistical uncertainty models? (One line, to understand better what is different from the ensemble forecast)
AR: The references will be added and the text revised explaining the uncertainty model referred to.

**Author Response to RC#2**

AC: We suggest writing: “the uncertainty model accounts for the strong autocorrelation in forecast errors and estimates an uncertainty band around the deterministic temperature, precipitation and streamflow forecasts.”

P4 2.1 Is Gjuvaa in the region South or East? Bulken is in the West region?
AR: We agree that the current manuscript is somewhat unclear on this issue and in the revised manuscript, we suggest adding in parenthesis to which region each catchment belongs.
AC: E.g.: “Bulken (W), Gjuvaa (E)”.
We will change the following sentence: “Gjuvaa (E) is non-glaciered and located inland.”

P5L20 PEST can be generic tools for parameter estimation or a particular software, what is it here?
AR: We use the PEST software to estimate parameters. We will specify this in the revised manuscript.
AC: We suggest to write: “… we used the operational model setup which has been calibrated using the PEST software to establish model parameters (Doherty, 2015)”

P6 2.2.4 / 3.1.1 Is the forecast from ECMWF point forecast (centre of the grid cell) or averages for the entire grid cell?
AR: The ECMWF forecasts should be considered as average values within the grid box, see Owens (2018, fig 3.2.1) for details.
AC: We will add this to the ECMWF description: “The ECMWF grid temperature, **which represents the average temperature for the grid cell**, was interpolated from a horizontal resolution of 0.25 ….”

P7L23-24 “In this study, the ensemble range (…) visually assessed the sharpness.” Something seems wrong, rephrase.
AR: Thank you. We will rephrase this sentence. We consider modifying this paragraph according to suggestions by RC#3, evaluate plot of empirical sharpness distribution.
AC: Suggestion: “In this study, the sharpness was visually assessed by looking at the ensemble range (i.e. the interval spanned by the lowest and highest forecasted values)”

P9L12-13 since “reliability has improved and some sharpness is maintained”. This could be better explained.
AR: We will modify this part including evaluating plots of the empirical distribution of sharpness, ref. RC#3, and information above.
AC: We will revise this paragraph in the manuscript

P6L24 I guess it should be “atmospheric lapse rate”?
AR: You are quite right; will be corrected.
AC: to “atmospheric”

P8L17 remove s from catchments.
P9L9 remove comma after convexity
P10L7 they performs – remove S.
AR: Thank you.
AC: We will correct as suggested.

P11L17 Rather than just sensitive, I think QM is unable to correctly map forecasts outside the observation range.

**Author Response to RC#2**

AR: We will rephrase to enhance the problems of QM mapping outside the observational range. It is important to note that all statistical methods will have problems outside the observational range. (ref RC#3, and discussion in RC#4)
AC: We suggest writing: "Quantile mapping (as most statistical calibration methods) is sensitive to forecasts outside the range of calibration values and period (Lafon et al. 2013), this can be an explanation for too high correction in the highest $T_{ens}$ quantile."
In addition, we add a sentence p7, l19 to clarify the use of quantile mapping: "The same coefficients based on the first 24h mapped, are applied to all lead times and ensemble members individually. For forecasts outside the observation range, a 1:1 extrapolation is used. I.e. if a forecast is 2°C higher than the highest mapped percentile, then the calibrated forecast is 2°C higher than the same percentile for the reference."

P12L2 temperatureS are?
AR: Thank you; will be corrected.
AC: Changed

P14 L29 "elevation correction dependency on lapse rate" – is this correct?
AR: We will rewrite to make this phrasing clearer (ref. Detailed comment no 2 by R#1),
AC: e.g. "… an elevation correction depending on lapse rate"

P16L17 No publisher?
AR: Thank you; will be corrected.
AC: "Engeland, K., Renard, B., Steinsland, I., and Kolberg, S.: Evaluation of statistical models for forecast errors from the HBV model. Journal of Hydrology, 384(1), 142-155, 2010."

Fig1 caption: Most of the catchments on the left are too small to be visible?
AR: We agree. The western catchments are small and thus difficult to distinguish on the map. We will revise the figure accordingly and further suggest adding a note on the fact that catchments on the western coast are small in the figure legend.
AC: We add the following to the caption:" Please note that many catchments are relatively small, for location confer the right map."

Reference
Owens, R G, Hewson, T D: ECMWF Forecast User Guide. Reading: ECMWF. doi: 10.21957/m1cs7h, 2018.