Corrigendum

to

Mitigating the impacts of climate non-stationarity on seasonal streamflow predictability in the US Southwest

(doi: 10.1002/2017GL076043)

Flavio Lehner[1], Andrew W. Wood[1], Dagmar Llewellyn[2], Douglas B. Blatchford[3], Angus G. Goodbody[4], Florian Pappenberger[5]

[1] Research Applications Laboratory, National Center for Atmospheric Research, Boulder, USA

[2] Bureau of Reclamation, Albuquerque Area Office, Albuquerque, USA

[3] Bureau of Reclamation, Lower Colorado Regional Office, Boulder City, USA

[4] National Water and Climate Center, Natural Resources Conservation Service, Portland, USA

[5]Forecast Department, European Centre for Medium-Range Weather Forecasts, Reading, UK

Corresponding author: Flavio Lehner (flehner@ucar.edu)

Published on Flavio Lehner's website in October 2018

An abbreviated version is published on the *Geophysical Research Letters* website

Abstract

In the original paper (Lehner et al. 2017) an approach was introduced to improve the skill of seasonal streamflow forecasts in the U.S. Southwest by adding climate model-based temperature forecasts to an operational statistical streamflow forecast model ('temperature-aided forecast' versus 'baseline forecast'). The 'baseline forecast' in the original paper was not a perfect replica of the current operational forecast model used by the Natural Resources Conservation Service (NRCS) due to lack of standardization of the default predictors (snow water equivalent and accumulated precipitation) before use in the principle component regression model used to predict seasonal streamflow. As a result, the skill of the 'baseline forecast' was less than the current operational forecast. Consequently, the relative skill improvement from implementing the 'temperature-aided forecast' was higher than what would be obtained by implementing the proposed approach in the operational system of NRCS. While there remains skill improvement in the corrected assessment, it is about half of what was reported originally.

1 Introduction

The Natural Resources Conservation Service (NRCS) streamflow forecast model bases on a principle component regression (PCR) using observed measurements of snow water equivalent (SWE) and accumulated precipitation (P) from snow telemetry monitoring (SNOTEL) stations as predictors and seasonal streamflow volume (e.g., April-July flow total) as predictand. In Lehner et al. (2017), we mimicked the NRCS model according to Garen (1992), termed ‘baseline forecast’ model. Then, skillful seasonal temperature forecasts from seasonal climate prediction models for the headwater region of the Colorado River and Rio Grande were added as a predictor to the ‘baseline forecast’ model. This ‘temperature-aided forecast’ was shown to be more skillful than the ‘baseline forecast’. Specifically, metrics of correlation, relative root-mean-square error (rRMSE), Brier Skill Score for the lowest tercile (BSS<33rd percentile), and the Continuous Ranked Probability Skill Score (CRPSS) all showed robust improvements across 20 headwater gages and five forecast issue dates.

Since publication, it was discovered that the ‘baseline forecast’ was not a perfect replica of the operational forecast by NRCS. Specifically, while NRCS standardizes their predictors prior to use in the PCR, we had not. Standardization of predictors (that is, subtracting the mean and dividing by the standard deviation) prior to use in PCR typically leads to more evenly distributed PC loadings among predictors and better forecast skill. Indeed, after making this pre-processing step consistent between our ‘baseline forecast’ model and NRCS, the ‘baseline forecast’ became slightly more skillful. The skill of the ‘temperature-aided forecast’ remained roughly the same, which resulted in a reduction of the relative skill improvement reported originally.

2 New Results and Conclusions

Figure 1 shows a comparison of the originally reported skill improvement and the corrected skill improvement. The figure is modeled after Figure 1 in Lehner et al. (2017), summarizing the results across streamflow gages and forecast issue dates/lead times. The corrected skill improvement across the different skill metrics is approximately half of what it was originally. Interestingly, for all metrics except BSS<33rd percentile, the skill improvement actually becomes more robust than originally, with almost no streamflow gage and issue dates showing a decrease in skill. For BSS<33rd percentile, however, more forecasts see skill decreases than in the original assessment.

These new results are confirmed in a bootstrapping exercise that assesses the sampling uncertainty of the skill improvement (see Lehner et al. (2017) for details on the exercise): for correlation, rRMSE, and CRPSS, 99% of forecasts are improved, with 98-100% of those significantly. For BSS<33rd percentile, only 62% of all forecasts are improved, 95% of those significantly (Table 1).

Another point is the attribution of the skill improvement. In the corrected assessment, the skill improvement from using just the observed long-term linear temperature trend as a predictor (dotted black line in Figure 1) is of approximately equal magnitude as when the forecasted interannual temperature variability is used as a predictor (solid black line in Figure 1). Thus, it cannot be excluded anymore that the skill improvement originates solely from the long-term warming trend.

Importantly, the main conclusion of Lehner et al. (2017) remain unchanged: adding seasonal temperature forecasts to current operational statistical seasonal streamflow forecasts in the U.S. Southwest improves their skill across the majority of metrics, gages, and issue dates.

References

Garen, D. C., 1992: Improved Techniques in Regression-Based Streamflow Volume Forecasting. *J. Water Resour. Plan. Manag.*, **118**, 654–670, doi:10.1061/(ASCE)0733-9496(1992)118:6(654).

Lehner, F., A. W. Wood, D. Llewellyn, D. B. Blatchford, A. G. Goodbody, and F. Pappenberger, 2017: Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. *Geophys. Res. Lett.*, **44**, 12,208-12,217, doi:10.1002/2017GL076043.

Figures

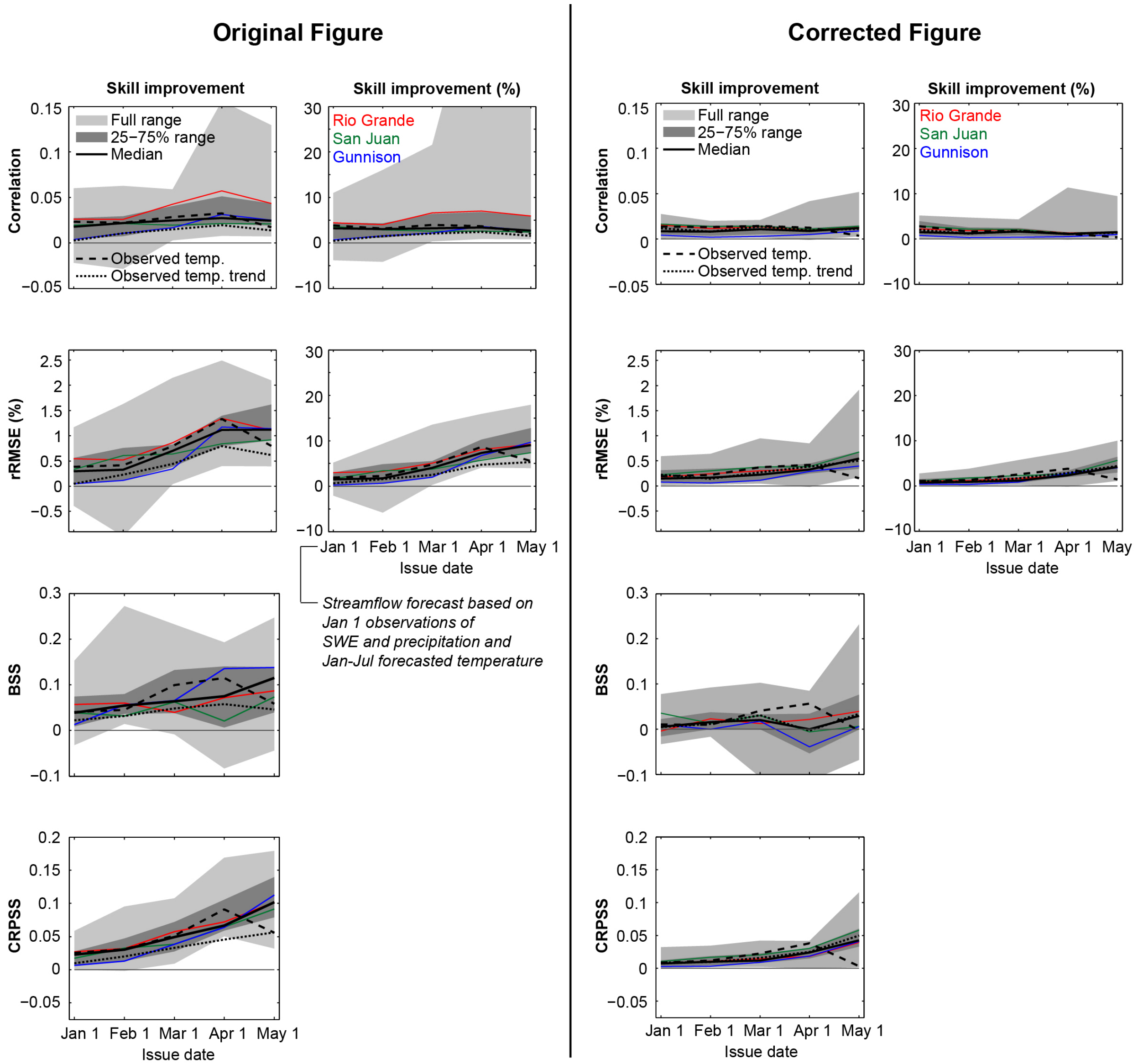

Figure 1: Comparison of the original and corrected version of Figure 3 from Lehner et al. (2017). Within each figure, the left column shows absolute skill improvement for all gages as a function of issue date, while the right column shows relative skill improvement for all gages as a function of issue date. Solid lines are the median across (black) all gages and (colors) the three basins. Dashed line is the median across all gages when using observed temperature, mimicking the hypothetical case where the future temperature is known at the time of forecast issue, and dotted line is the median when using only the linear trend of observed temperature.

Tables

Table 1: Forecast skill improvement, stratified by skill metric and forecast issue date. For each forecast issue date, the first column gives the fraction, in %, of gages (total of 20 gages) that show an improvement in the respective skill metric, while the second column gives the fraction, in %, of these improved forecasts that are significant at the 95% level. See Section 3.4 in Lehner et al. (2017) for details on the significance test.

	Jan 1		**Feb 1**		**Mar 1**		**Apr 1**		**May 1**		**Total**	
Correlation	100	100	100	100	100	100	95	100	100	100	99	100
rRMSE	100	100	100	100	100	100	95	95	100	100	99	99
BSS	55	100	65	100	70	93	45	89	75	93	62	95
CRPSS	100	100	100	95	100	100	95	95	100	100	99	98