Water Supply Forecast Rodeo: Explainability and Communication Bonus Track

Report

Author: Matti Kämäräinen – *kamarain*

Abstract

One information-rich figure was found sufficient for presenting the previous and current model results for each target site and issue date. This figure also contains the historical observed quantiles of the naturalized flow to put the results into context.

The approach of presenting the explanations of the competition forecasts follows the bottom-up strategy. First, individual model features – the spatio-temporal principal components of each dataset – are explored with the SHAP analysis to find out which ones are contributing the most to the predictions of each issue date and target site. After that the identified features are visualized and narrated together with the other features and datasets.

Technical Approach

Forecast and uncertainty communication

Figure 1

The **MultiQuantileRegressor** is an ensemble model, and therefore the output from it is a distribution. In Figure 1 of the forecast summaries the model forecast means of target quantiles are shown with continuous colored lines, and the ensemble spread is indicated with shading for each predicted quantile. Thick dashed horizontal lines represent the observed 1981–2023 quantiles. Thin vertical lines indicate the actual issue dates, but forecasted values are shown also for dates between the issue dates. The latest forecasted values are shown in text boxes.

Explainability metrics and communication

Figure 2

The presentation of the explainability results follows the bottom-up approach. In this approach the contribution of features of the model – the PCs of each dataset – are first analyzed by using the SHAP values (Lundberg and Lee, 2017) to find out the most important ones in general for the site (pink bars in Figure 2) and for the current issue date (blue bars). The SHAP waterfall plot is also used to identify the sign of the contribution of each feature. These analyses determine the focus of the narrative analysis.

Figure 3

After finding the most important features via the SHAP analysis, Figure 3 is then plotted dynamically to put the top-six features into climatological context. This figure shows what is the typical variability range of the selected feature set on different days of the water year, and the time series of each feature for the current forecast year. PCs from observational datasets contain historical data, including information about differences and other autoregressive variables, which help the model understand the ongoing local processes as well as more distant processes from the teleconnection indices. These processes can be used to predict the future evolution of those data and the resulting streamflow. On the other hand, the ECMWF PCs contain more direct (but naturally uncertain) future information from the simulations extending to the future months.

Visualizing the PCs of the datasets related to streamflows was challenging. Distributions of these principal components are highly skewed (because the datasets from which they are calculated are also skewed), and because the PCs typically contain negative values, transforming the figure y-axis to logarithmic does not work. For this reason only the climatological median is presented in these cases along with the observed time series of the forecast year. Despite that, the linear y-scale makes the low flow values difficult to read if high values are also presented in the same figure.

Figure 4

The left part of the static Figure 4 offers always the same information: the spatial distribution of the most up-to-date PDSI data, the boundaries of the catchment, and the three key SNOTEL variables near the station locations they were observed: accumulated precipitation, snow water equivalent, and the 30-day running mean of temperature. Different marker symbols separate the different variables, and color coding describes the *climatological percentiles*. Even though *percents of median* used in the provided example reports is a handy metric to show the values for zero-bounded distributions such as streamflow, snowpack or precipitation amounts, the climatological percentiles are better for describing the temperature and they are also suitable for other quantities. The only major flaw is that percentiles can not be used to show values exceeding the observational range of the years 1981–2023. Here the climatological percentiles are defined separately for each day of year.

Also the right part of Figure 4 is static, presenting the climatological quantiles of the key datasets calculated as means over the catchment and nearby regions. All these datasets are either directly observational (SNOTEL temperatures, precipitation, and snow water equivalent) or derived from observations based on modeling (PDSI and the monthly naturalized flow).

Narrative analysis

The narrative analysis flows between the figures and is driven by the important anomalies in the observational and future-looking datasets. Spatial and temporal aspects are considered, and chained interactions between variables and processes in long and short time scales are explored.

Discussion

The principal components (PCs) used for fitting and predicting contain dense information about spatial and temporal variability and ongoing processes in the catchment domains. Usually the first component describes the mean over the recent past and the catchment, but understanding the physical meaning of the higher components is not always straightforward. Some of them describe the temporal variability from the smoothing, lagging and differencing while others are more strongly connected to the spatial variability. It is also worth noting that the higher components are very sensitive to small differences in the data used for fitting of the PC models:

for example, the sign of them can be different for different fitting samples, and for this reason it is necessary to always use the same PC models for transforming the data after fitting. Furthermore, the units do not correspond to the units of the original data after transformation. Therefore, the interpretation of the principal component features should always be done together with the non-transformed, raw predictor data (Figure 4).

Visualizing spatially the important side product of the principal component analysis, the static weights (so called empirical orthogonal functions in meteorology) could help interpret the PCs by showing the regions where the variance of each PC is coming from. However, as the PCs here also contain temporal information, that might affect the spatial analysis. Including these analyses and figures to the forecast summaries was not possible due to limited page numbers either. On the other hand, for practical purposes, the accurate physical meaning and interpretation of the PCs might not be needed: it is often sufficient to see just their impact on the prediction, i.e. the sign and magnitude from the SHAP analysis (Figure 2).

References

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems, 30*, 4765-4774. Retrieved from https://arxiv.org/abs/1705.07874