



Bart van Osnabrugge, Jan Verkade

# Enabling hydrological (ensemble) verification as part of standard hydrological forecasting practice at operational centers around the world

## Introduction

Hydrological forecasts are routinely produced all over the world. When acting on forecast data, a mental model is made of the forecast quality, influencing the decision. Preferably this model is based on information provided through verification exercises.

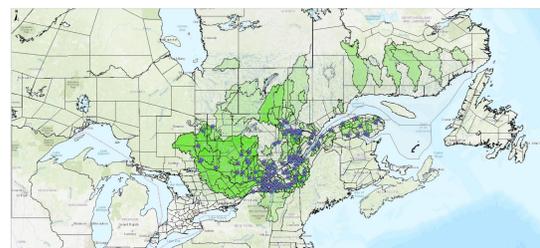
However, verification metrics are not very accessible. First, verification exercises can be time-consuming. Second, verification metrics are hard to interpret and difficult to translate into operational cases.

This raises the question how we can make verification metrics as information source on forecast quality more accessible.

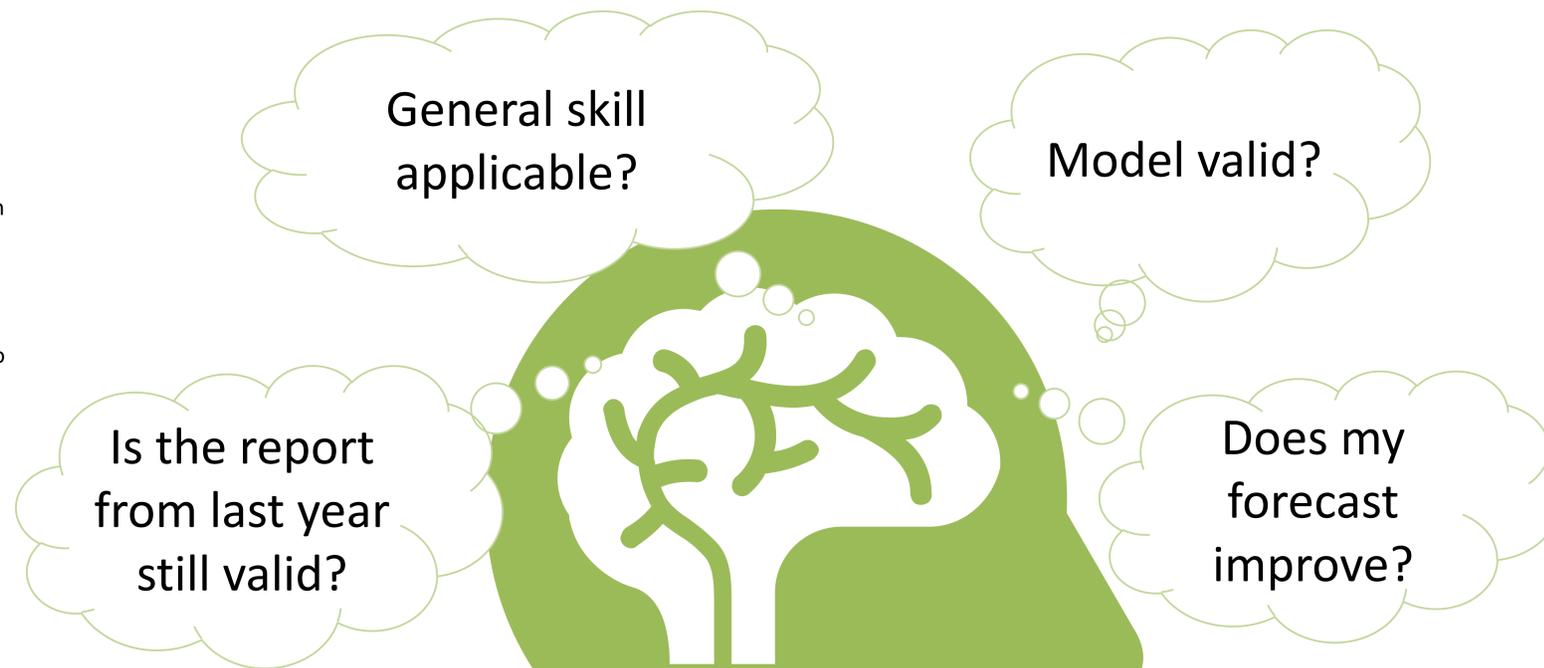
The answer explored here is to routinely calculate verification metrics by integrating the Ensemble Verification System (EVS) into a Delft-FEWS based forecasting system.

## Test system

The Système de Prévision Hydrologique (SPH), Quebec, Canada is based on Delft-FEWS and Deltares OpenArchive. For the verification it is expanded with the PI webservice and EVS.



Area covered by the SPH forecasting system

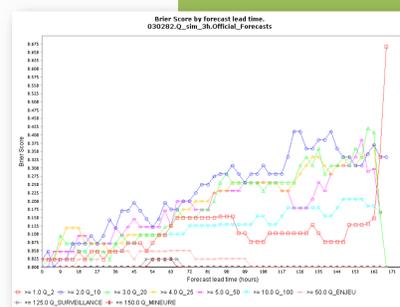


## Output

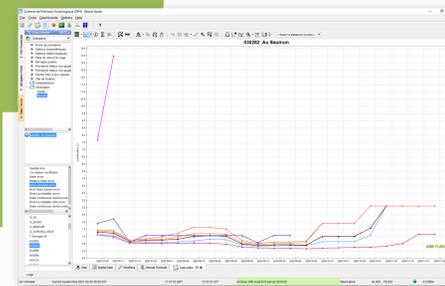
Time series of forecast metrics shown side-by-side with forecasts

HTML and PDF report with ensemble metrics

All verification metrics archived together with the data

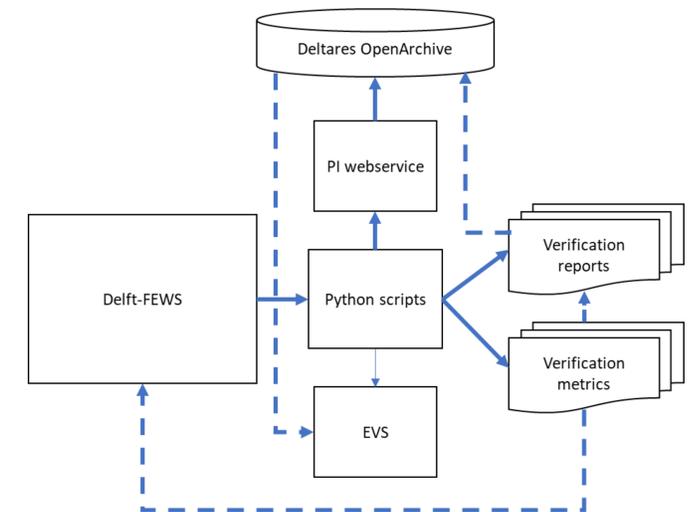


Ensemble Verification System output



Time series plot of metric evolution in time, shown in the same interface as the forecasts

## Workflow



High-level component description of the coupling between Delft-FEWS and the Ensemble Verification System (EVS)

## Lessons learned

- Automated verification is a step forward in forecast verification
- The multi-dimensional verification data can be imported and shown in forecasting software
- Offering the option to perform scheduled verification exercises brings the user questions regarding quality into focus:
  - How often to check for quality?
  - For which time periods?
  - Which metrics are useful in daily practice?
  - Does this metric answer the question I have?