



## Is it bias or not?

Marie-Amélie Boucher

Canada (marie-amelie\_boucher@uqac.ca)

Bias is commonly referred as a systematic error between forecasts and observations. Bias can vary (slowly) over time as the climate evolves, or whenever there is a change in the forecasting system. For instance, meteorological forecasts, required as inputs for hydrological prediction, undergo improvements from time to time. Bias can also change seasonally and depends on the magnitude of the forecast variable. For instance, bias for streamflow predictions during spring melt is different than during summer low flows. In the context of hydrological ensemble forecasts, one can evaluate bias  $b$  over a defined period of time by computing the mean difference between the ensemble mean and the corresponding observation over  $N$  forecasts. Chances are that will never be zero. Then, is it possible that an ensemble forecasting system be mistakenly diagnosed as biased? Is there a systematic way to compute a threshold value for  $b$  above which the forecasts should be considered biased? To what extent is it important to include bias removal in a post-processing strategy? Similarly, to what extent is it detrimental to apply bias removal to unbiased forecasts? Such questions are addressed here through the use of both synthetic and real datasets. Bias correction methods are also coupled with kernel dressing methods for post-processing and the effect of bias correction (or lack of thereof) on the final post-processed forecasts is evaluated. This allows to end up with guidelines for a threshold value for bias  $b$ . In addition, bias could be defined differently than above. For instance, it could be evaluated distinctly for each rank in the ensemble. Both synthetic and real datasets are used to evaluate two bias correction strategies. The first strategy considers bias as the mean difference between the ensemble mean and the observation. The second strategy consists in sorting the ensemble members and computing the difference between each sorted member and the observation. Then there will be  $n$  values for bias instead of one, with  $n$  the number of ensemble members. The results show that in many cases computing and removing bias separately for each rank is more efficient than computing it using the ensemble mean.