EGU24-8007, updated on 20 May 2024
https://doi.org/10.5194/egusphere-egu24-8007
EGU General Assembly 2024

Investigating the divide and measure nonconformity

Daniel Klotz[1], Martin Gauch[2], Frederik Kratzert[3], Grey Nearing[4], and Jakob Zscheischler[1]
[1]Helmholtz Centre for Environmental Research (UFZ), Leipzig, Germany
[2]Google Research, Zurich, Switzerland
[3]Google Research, Vienna, Austria
[4]Google Research, Mountain View, California, USA

This contribution presents a diagnostic approach to investigate unexpected side effects that can occur during the evaluation of rainfall--runoff models.

The diagnostic technique that we use is based on the idea that one can use gradient descent to modify the runoff observations/simulations to obtain warranted observations/simulations. Specifically, we show how to use this concept to manipulate any hydrograph (e.g., a copy of the observations) so that it approximates specific NSE values for individual parts of the data. In short, we follow the following recipe to generate the synthetic simulations: (1) copy the observations, (2) add noise, (3) clip the modified discharge to zero, and (4) optimise the obtained simulation values by using gradient descent until a desired NSE value is reached.

To show how this diagnostic technique can be used we demonstrate a behaviour of Nash--Sutcliffe Efficiency (NSE) that appears when evaluating a model over subsets of the data: If models perform poorly for certain situations, this lack of performance is not necessarily reflected in the NSE (of the overall data). This behaviour follows from the definition of NSE and is therefore 100% explainable. However, from our experience it can be unexpected for many modellers. Our results also show that subdividing the data and evaluating over the resulting partitions yields different information regarding model deficiencies than an overall evaluation. We call this phenomenon the Divide And Measure Nonconformity or DAMN.