



## Towards an enhanced objective function for hydrological model calibration to improve the performance along the whole flow duration curve

Stefania Grimaldi<sup>1</sup>, Davide Bavera<sup>2</sup>, **Andrea Ficchi**<sup>3</sup>, Francesca Moschini<sup>1,4</sup>, Andrea Toreti<sup>1</sup>, Alberto Pistocchi<sup>1</sup>, and Peter Salamon<sup>1</sup>

<sup>1</sup>European Commission, Joint Research Centre, Ispra, Italy

<sup>2</sup>Arcadia SIT, Milan, Italy

<sup>3</sup>Politecnico di Milano, Milan, Italy

<sup>4</sup>Rey Juan Carlos University, Madrid, Spain

The definition of the objective function for hydrological model calibration and goodness-of-fit metrics for validation are crucial in characterising model performance and driving model development. Ideally, model developers and users should identify the most suitable metric to characterize model performance based on their specific use cases and objectives. The benefits of this are evident; for example, the optimal objective function to calibrate a hydrological model only used for flood forecasting should be different from the optimal one to be used for a model focusing only on low flows. However, in practice, most hydrological models are used for multiple applications and a standard generalist function is adopted for their calibration. The two most widely used generalist functions are the Nash-Sutcliffe Efficiency (NSE) and the Kling-Gupta Efficiency (KGE), in its standard and modified versions. While the NSE is a simple normalization of the mean square error (MSE), the KGE overcomes some of the NSE limitations based on the decomposition of the MSE into three components, i.e., the error in the mean simulated streamflow, the relative variability, and the linear correlation between simulations and observations. Still, KGE presents some limitations, including a large sensitivity to outliers and an assumption of linearity and normality in the error distribution, which are impactful especially for the characterization of performances over low flows. Some alternative calibration functions have been proposed in the literature to overcome these limitations, but no calibration function overcomes the traditional two options (NSE and KGE) in improving the simulation along the whole range of the flow duration curve. Here we present the results of an extensive comparison of hydrological models calibrated and validated with multiple functions, including new variants and combinations of KGE and information-theory metrics, that can be suitable to characterize the performance over high, low, and regime flows. Two hydrological models (GR4J and Open Source LISFLOOD) were calibrated with several alternative calibration functions over more than 200 catchments in Europe with a varied range of hydroclimatic conditions. Both models were evaluated using multiple metrics, including use-case specific hydrological signatures focusing on flood characteristics, average regime and low flows. Based on our analysis, a new function is

proposed which combines the three KGE components with an additional component based on the Jensen-Shannon Divergence. The performance of the two models calibrated with the new function is shown to outperform the standard KGE and NSE over low flows with minimal change in performance over regimes and high flows. This study shows that more effort should be devoted to the choice of the optimal calibration function for hydrological model applications when aiming to improve specific aspects of model performance.