

EGU21-7138

<https://doi.org/10.5194/egusphere-egu21-7138>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Value of multi-source dataset for hydrological catchment modeling

Carolina Natel de Moura^{1,2}, João Marcos Carvalho¹, and Jan Seibert²

¹Federal University of Parana, Department of Hydraulics and Sanitation, Curitiba, Brazil

²University of Zurich, Department of Geography, Zurich, Switzerland

Global meteorological and hydrological datasets have become increasingly available in the past few decades, marked by an increase in the number of large datasets, often including hundreds of catchments. These data sets bring two main advantages: the ability to perform hydrological modeling over a large number of catchments located in different hydroclimate characteristics, - which leads us to more robust hypothesis testing, and the ability to address the uncertainties related to the hydrological model input data. However, the full added value to hydrological modeling is not yet fully understood. The main questions surrounding the use of multi-source and large-scale datasets are related to how much value these datasets add to the performance of hydrological models. How different are these datasets, how accurate are they, and whether their use results in similar or rather different hydrological simulations? Other questions are how can we better combine them for improved predictions, and what is the average uncertainty of the input datasets in hydrological modeling? We aimed here to investigate better those issues using Brazilian catchments as study cases. The Brazilian hydrometeorological network has several issues to overcome, such as an undistributed spatial network resulting in data-scarce areas, a large amount of missing data, and the lack of standardized and transparent quality analysis. In this study, we used a national hydrometeorological dataset (CAMELS-BR) along with other several global forecast and reanalysis meteorological datasets, such as the CFSv2 and ECMWF, for the streamflow prediction using the data-driven model Long-Short Term Memory (LSTM). Initial results indicate that calibrating a recurrent neural network is clearly depending on the data source. Moreover, the tested global meteorological products are found to be suitable for hydrological modeling. The combination of different data sources in the hydrological model seems to be beneficial, especially in those areas where ground-level gauge stations are scarce.