Towards Open and FAIR Hydrological Modelling with eWaterCycle

Niels Drost¹, Jerom P.M. Aerts², Fakhereh Alidoost¹, Bouwe Andela¹, Jaro Camphuijsen¹, Nick van de Giesen², Rolf Hut², Eric Hutton³, Peter Kalverla¹, Gijs van den Oord¹, Inti Pelupessy¹, Stef Smeets, Stefan Verhoeven¹, and Ben van Werkhoven³

¹Netherlands eScience Center, Amsterdam, Netherlands (n.drost@esciencecenter.nl)
²Delft University of Technology, Faculty of Civil Engineering and Geosciences, Delft, Netherlands
³Community Surface Dynamics Modeling System, Boulder, CO, United States

The eWaterCycle platform (https://www.ewatercycle.org/) is a fully Open Source system designed explicitly to advance the state of Open and FAIR Hydrological modelling. While working with Hydrologists to create a fully Open and FAIR comparison study, we noticed that many ad-hoc tools and scripts are used to create input (forcing, parameters) for a hydrological model from the source datasets such as climate reanalysis and land-use data. To make this part of the modelling process better reproducible and more transparent we have created a common forcing input processing pipeline based on an existing climate model analysis tool: ESMValTool (https://www.esmvaltool.org/).

Using ESMValTool, the eWaterCycle platform can perform commonly required preprocessing steps such as cropping, re-gridding, and variable derivation in a standardized manner. If needed, it also allows for custom steps for a hydrological model. Our pre-processing pipeline directly supports commonly used datasets such as ERA-5, ERA-Interim, and CMIP climate model data, and creates ready-to-run forcing data for a number of Hydrological models.

Besides creating forcing data, the eWaterCycle platform allows scientists to run Hydrological models in a standardized way using Jupyter notebooks, wrapping the models inside a container environment, and interfacing to these using BMI, the Basic Model Interface (https://bmi.readthedocs.io/). The container environment (based on Docker) stores the entire software stack, including the operating system and libraries, in such a way that a model run can be reproduced using an identical software environment on any other computer.

The reproducible processing of forcing and a reproducible software environment are important steps towards our goal of fully reproducible, Open, and FAIR Hydrological modelling. Ultimately, we hope to make it possible to fully reproduce a hydrological model experiment from data preprocessing to analysis, using only a few clicks.