



Massive Parallelization of the Global Hydrological Model mHM

Maren Kaluza, Luis Samaniego, Stephan Thober, Robert Schweppe, Rohini Kumar, and Oldrich Rakovec
UFZ, CHS, Leipzig, Germany

Parameter estimation of a global-scale, high-resolution hydrological model requires a powerful supercomputer and an optimized parallelization algorithm. Improving the efficiency of such an implementation is essential to advance hydrological science and to minimize the uncertainty of the major hydrologic fluxes and storages at continental and global scales. Within the ESM project [1], the main transfer-function parameters of the mHM model will be estimated by jointly assimilating evapotranspiration (ET) from FLUXNET, the TWS anomaly from GRACE (NASA) and streamflow time series from 5500 GRDC gauges to achieve this goal.

For the parallelization of the objective functions, a hybrid MPI-OpenMP scheme is implemented. While the parallelization into equally sized subdomains for cell-wise computations of fluxes (e.g., ET, TWS) is trivial, cell-to-cell fluxes need to be computed for streamflow routing. For time series datasets, the advanced parallelization algorithm MPI parallelized Decomposition of Forest (MDF) will be used.

In this study, we go beyond the standard approach which decomposes the river into tributaries (e.g. the Pfaffenstetter System [2]). We apply a non-trivial graph algorithm to decompose each river-network into a tree data structure with nodes representing subbasin domains of almost equal size [3].

We analyze several aspects affecting the MDF parallelization: (1) the communication time between nodes; (2) buffering data before sending; (3) optimizing total node idle time and total run time; (4) memory imbalance between master processes and other processes.

We run the mHM model on the high-performance JUWELS supercomputer at Jülich Supercomputing Center (JSC) where the (routing) code efficiently scales up to ~180 nodes with 96 CPUs each. We discuss different parallelization aspects, including the effect of parameters onto the scaling of MDF and we show the benefits of MDF over a non-parallelized routing module.

[1] <https://www.esm-project.net/>

[2] <http://proceedings.esri.com/library/userconf/proc01/professional/papers/pap1008/p1008.htm>

[3] <https://meetingorganizer.copernicus.org/EGU2019/EGU2019-8129-1.pdf>