EGU21-7154
https://doi.org/10.5194/egusphere-egu21-7154
EGU General Assembly 2021

Global scale hydrological modelling at 100 m, 1 h resolution, in Python

Kor de Jong[1,2], Marc van Kreveld[2], Debabrata Panja[2], Oliver Schmitz[1], and Derek Karssenberg[1]
[1]Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands
[2]Department of Information and Computing Sciences, Faculty of Science, Utrecht University, Utrecht, The Netherlands

Data availability at global scale is increasing exponentially. Although considerable challenges remain regarding the identification of model structure and parameters of continental scale hydrological models, we will soon reach the situation that global scale models could be defined at very high resolutions close to 100 m or less. One of the key challenges is how to make simulations of these ultra-high resolution models tractable ([1]).

Our research contributes by the development of a model building framework that is specifically designed to distribute calculations over multiple cluster nodes. This framework enables domain experts like hydrologists to develop their own large scale models, using a scripting language like Python, without the need to acquire the skills to develop low-level computer code for parallel and distributed computing.

We present the design and implementation of this software framework and illustrate its use with a prototype 100 m, 1 h continental scale hydrological model. Our modelling framework ensures that any model built with it is parallelized. This is made possible by providing the model builder with a set of building blocks of models, which are coded in such a manner that parallelization of calculations occurs within and across these building blocks, for any combination of building blocks. There is thus full flexibility on the side of the modeller, without losing performance.

This breakthrough is made possible by applying a novel approach to the implementation of the model building framework, called asynchronous many-tasks, provided by the HPX C++ software library ([3]). The code in the model building framework expresses spatial operations as large collections of interdependent tasks that can be executed efficiently on individual laptops as well as computer clusters ([2]). Our framework currently includes the most essential operations for building large scale hydrological models, including those for simulating transport of material through a flow direction network. By combining these operations, we rebuilt an existing 100 m, 1 h resolution model, thus far used for simulations of small catchments, requiring limited coding as we only had to replace the computational back end of the existing model. Runs at continental scale on a computer cluster show acceptable strong and weak scaling providing a strong indication that global simulations at this resolution will soon be possible, technically speaking.

Future work will focus on extending the set of modelling operations and adding scalable I/O, after

which existing models that are currently limited in their ability to use the computational resources available to them can be ported to this new environment.

More information about our modelling framework is at https://lue.computationalgeography.org.

References

[1] M. Bierkens. Global hydrology 2015: State, trends, and directions. Water Resources Research, 51(7):4923–4947, 2015.
[2] K. de Jong, et al. An environmental modelling framework based on asynchronous many-tasks: scalability and usability. Submitted.
[3] H. Kaiser, et al. HPX - The C++ standard library for parallelism and concurrency. Journal of Open Source Software, 5(53):2352, 2020.