Utrecht University

Faculty of Geosciences
Department of Physical Geography

Introducing DL-GLOBWB

A deep-learning surrogate of a process-based global hydrological model

Bram Droppers, Marc F.P. Bierkens and Niko Wanders

b.droppers@uu.nl @DroppersBram

Why a deep-learning surrogate of a process-based global hydrological model?

Due to their process-based approach, global hydrological models have become an essential tools to support global change research and environmental assessments, as these models allow for predictions under changing climatic conditions and changes in water management. However:

1. The need for higher spatial resolution and larger ensemble scenario predictions is nearing the computational limitations of process-based models
2. Process-based models often perform poorly in regions where limited observational data for calibration is available

A deep-learning surrogate can help mitigate these challenges. Deep-learning models are orders of magnitude faster than their process-based counterparts. Moreover, although deep-learning models trained on observations are limited to the observed variable in the historical context, the surrogate will learn all water-balance components and can be trained on non-historical process-based model outputs.

Here we present a deep-learning surrogate framework for process-based hydrological models, applied to the PCR-GLOBWB (PCRaster Global Water Balance) hydrological model[1,2].

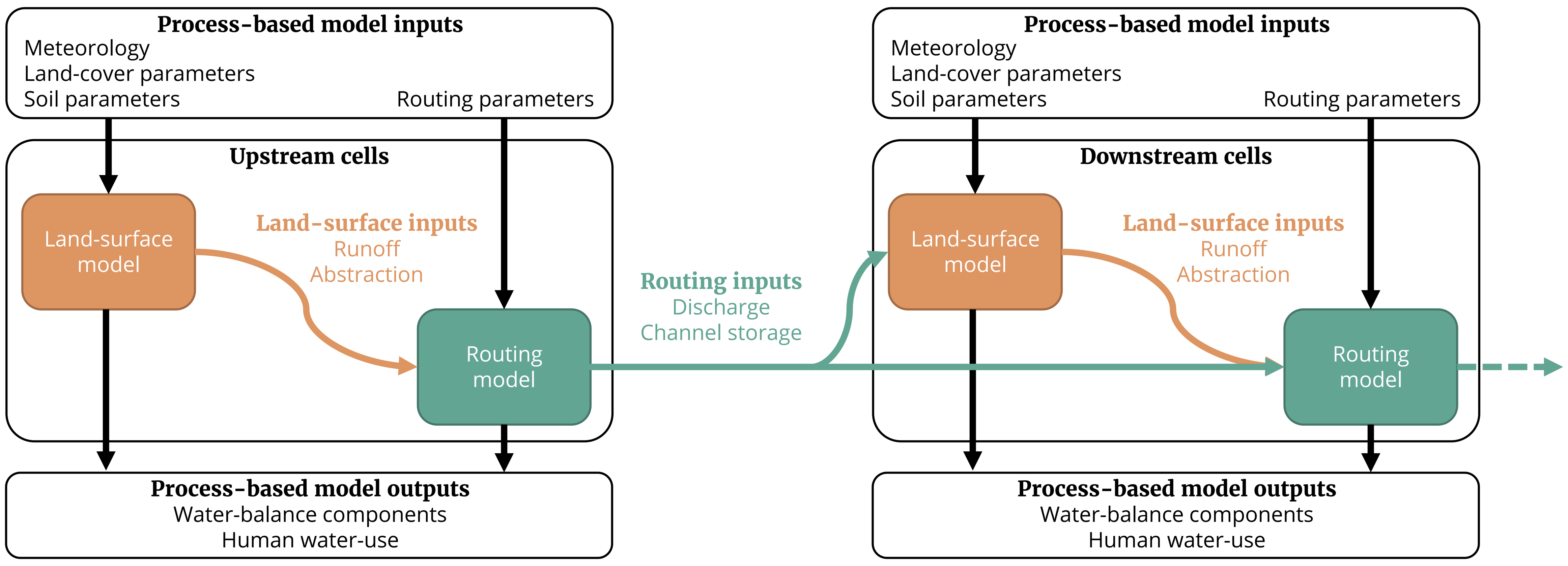
Surrogate design

Our deep-learning surrogate consists of two separate models:

- land-surface model*
- routing model*

*Rectified linear unit (ReLu) activated Long-Short Term Memory (LSTM) layer, enclosed by two linear layers.

Cells are processed sequentially from upstream to downstream so that (un)available water can propagate throughout the river basin.



Multi-resolution

A key requirement for our deep-learning surrogate is that the surrogate should be able to *predict across different spatial resolutions*. Therefore, we trained three models:

- A 30 arc-minute model
- A 5 arc-minute model
- A multi-resolution model*

*Combines half of the 30 arc-minute and half of the 5 arc-minute samples

Conclusions

- In general all models performed well for their target resolutions and captured the spatial and temporal patterns of the outputs.
- However, The single-resolution models perform poorly outside their target resolutions, indicating limited understanding of underlaying processes.
- Conversely, the multi-resolution model performs well on both resolutions and, surprisingly, often outperforms the single-resolution models on their target resolution.

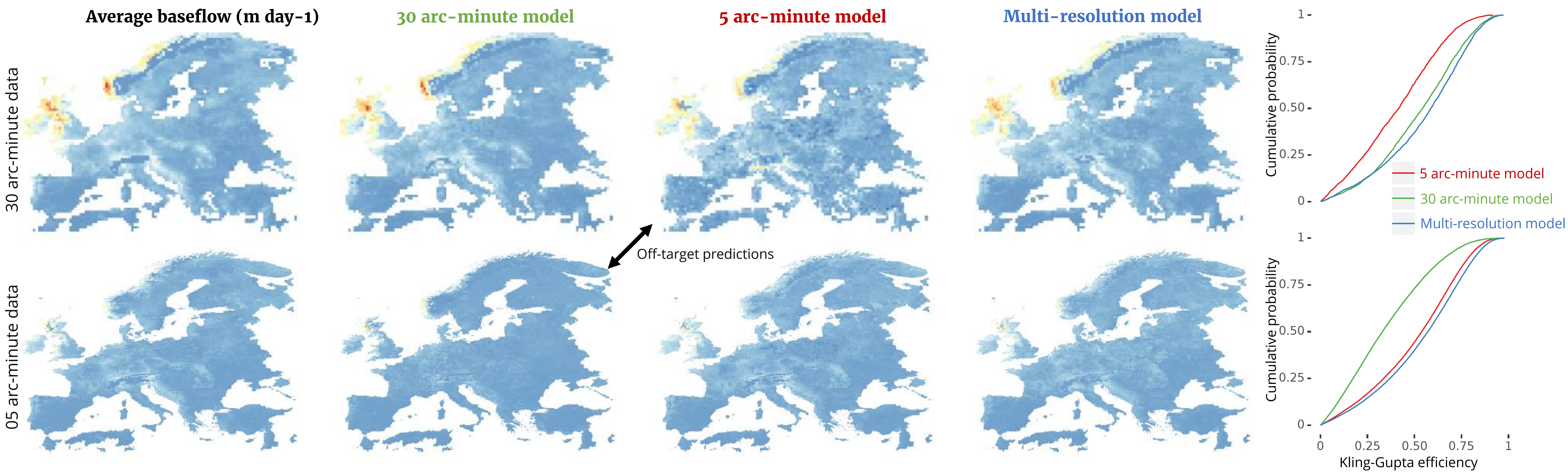
Outlook

The deep-learning surrogate helps us to make high-resolution, large-ensemble predictions computationally feasible, while, if trained appropriately, also allowing for climate-change and adaptation scenario analyses.

Moreover, the deep-learning surrogate, being a differentiable model, will allow for:

- parameter learning to improve model accuracy[3] (compared to observations)
- flux matching[4] (across spatial resolutions).

Multi-resolution training is crucial for model understanding and generalization.



References

1. Sutanudjaja, E. H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H., Drost, N., ... & Bierkens, M. F. (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geoscientific Model Development, 11(6), 2429-2453. doi:10.5194/gmd-11-2429-2018.
2. https://pcrglobwb.readthedocs.io/en/latest/
3. Tsai, W. P., Feng, D., Pan, M., Beck, H., Lawson, K., Yang, Y., ... & Shen, C. (2021). From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. Nature communications, 12(1), 5988. doi:10.1038/s41467-021-26107-z.
4. Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resources Research, 46(5). doi:10.1029/2008WR007327.

uu.nl/staff/BDroppers