



Progress on a comprehensive Earth System model evaluation framework

Wouter J. M. Knoben¹, Vincent Vionnet² & Martyn P. Clark¹

¹University of Saskatchewan, Canada; ²Environmental Numerical Research Prediction, Environment and Climate Change Canada, Canada



Environment and
Climate Change Canada
Environnement et
Changement climatique Canada



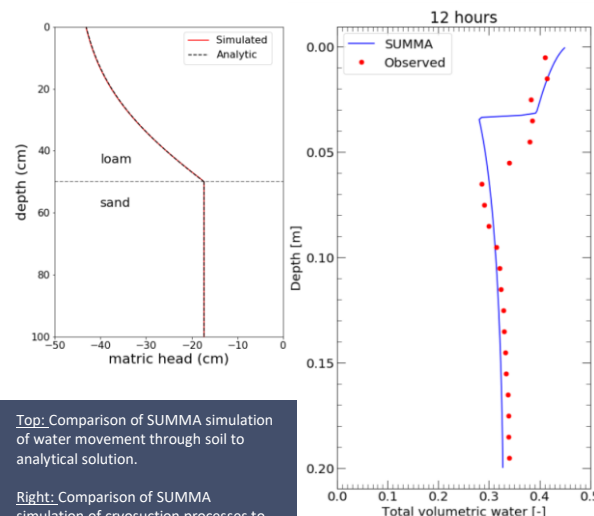
UNIVERSITY OF
SASKATCHEWAN

Thorough model evaluation is critical if a model is to provide simulations under changing conditions. Evaluation should include:

1. Confirmation that equations were implemented correctly;
2. Multi-scale and multi-variate comparison of simulations to observations;
3. Comparison of model performance against benchmarks.

Are the equations correctly implemented?

Laugh tests: compare isolated model components to known solutions^[1,2].



Top: Comparison of SUMMA simulation of water movement through soil to analytical solution.

Right: Comparison of SUMMA simulation of cryosuction processes to laboratory experiments.

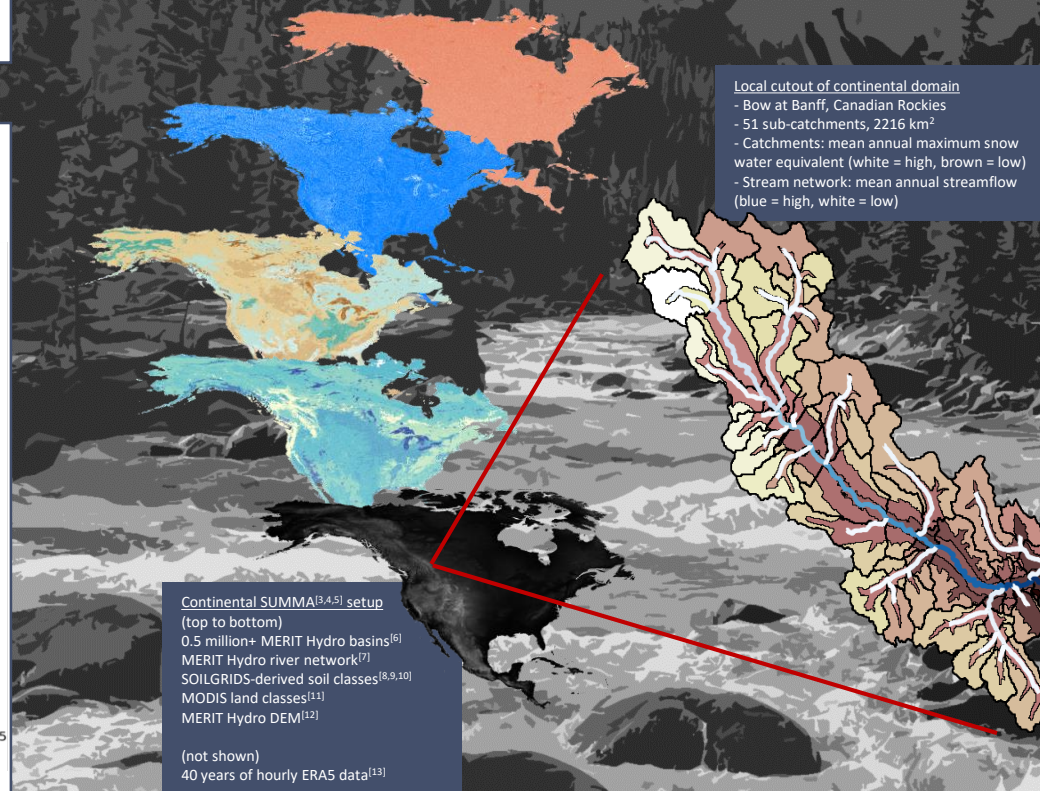
Progress on large-domain testbed for model evaluation

Goal 1: large-domain process-based model evaluation

Goal 2: local cutouts for data networks and experimental basins

Model setup effectively complete but some polishing needed.

See also: <https://meetingorganizer.copernicus.org/EGU21/EGU21-3091.html>



Local cutout of continental domain
- Bow at Banff, Canadian Rockies
- 51 sub-catchments, 2216 km²
- Catchments: mean annual maximum snow water equivalent (white = high, brown = low)
- Stream network: mean annual streamflow (blue = high, white = low)

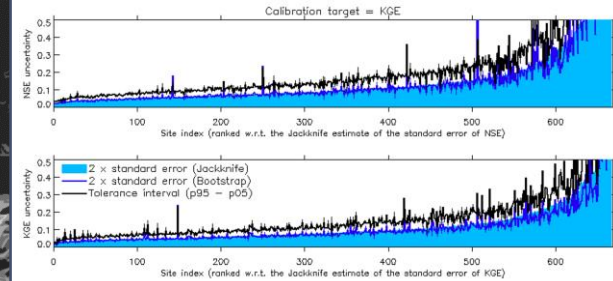
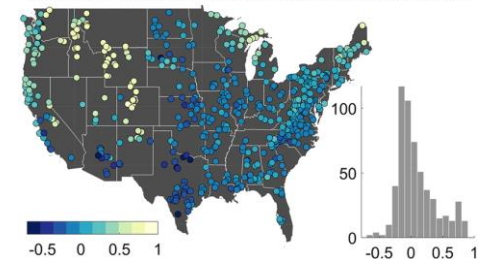
Continental SUMMA^[3,4,5] setup
(top to bottom)
0.5 million+ MERIT Hydro basins^[6]
MERIT Hydro river network^[7]
SOILGRIDS-derived soil classes^[8,9,10]
MODIS land classes^[4,11]
MERIT Hydro DEM^[12]
(not shown)
40 years of hourly ERA5 data^[13]

Benchmarking: on the use of efficiency metrics

Commonly used metrics such as NSE and KGE:

- Need context for interpretation (top^[14])
- Are highly uncertain (bottom^[15])

KGE scores obtained from streamflow seasonality - models must beat this lower benchmark



- [1] Clark et al., 2021. Journal of Hydrometeorology. <https://doi.org/10.1175/JHM-D-20-0175.1>
- [2] Laugh tests repository. <https://github.com/CH-Earth/laughTests>
- [3] Clark et al., 2015a. Water Resources Research <https://doi.org/10.1002/2015WR017198>
- [4] Clark et al., 2015b. Water Resources Research <https://doi.org/10.1002/2015WR017200>
- [5] Mizukami et al., 2016. Geoscientific Model Dev <https://doi.org/10.5194/gmd-9-2223-2016>
- [6] Based on Lin et al., 2019. Water Resources Res. <https://doi.org/10.1029/2019WR025287>
- [7] Lin et al., 2019. Water Resources Res. <https://doi.org/10.1029/2019WR025287>
- [8] Hengl et al., 2017. PLoS ONE. <https://doi.org/10.1371/journal.pone.0169748>
- [9] Benham et al., 2009. Tech note. <https://bit.ly/3uWQQ1l>
- [10] Knoben, 2021. HydroShare <https://doi.org/10.4211/hydroshare.1361509511e44adfb814f6950c6e742>
- [11] Friedl et al., 2019. Online. <https://doi.org/10.5067/MODIS/MCD12Q1.006>
- [12] Yamazaki et al., 2019. Water Resources Research. <https://doi.org/10.1029/2019WR024873>
- [13] Copernicus Climate Change Service, 2017. <https://cds.climate.copernicus.eu/cdsapp#!/home>
- [14] Knoben et al., 2020. Water Resources Research. <https://dx.doi.org/10.1029/2019WR025975>
- [15] Clark et al., (under review). Water Resources Research.

Background image: Keirnan Fowler