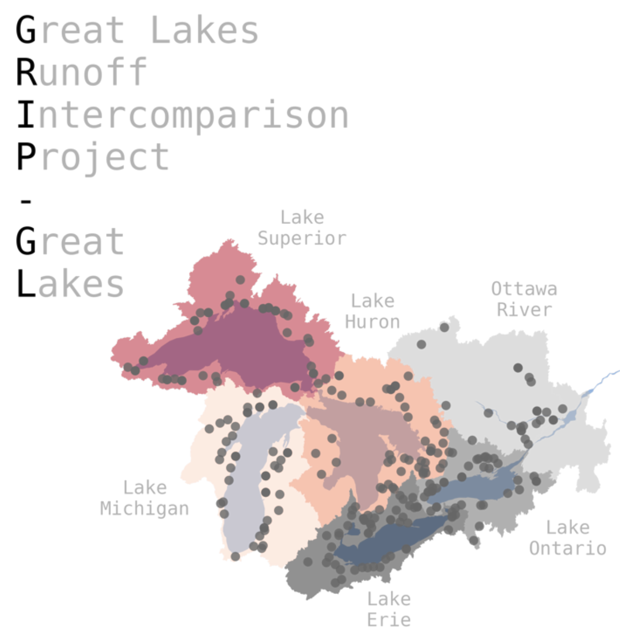
The Great Lakes Runoff Intercomparison Project Phase 4: The Great Lakes (GRIP-GL)

README created on June 21, 2022 by Juliane Mai and reviewed by Hongren Shen and Bryan A. Tolson.

Abstract: Model intercomparison studies are carried out to test and compare the simulated outputs of various model setups over the same study domain. The Great Lakes region is such a domain of high public interest as it not only resembles a challenging region to model with its trans-boundary location, strong lake effects, and regions of strong human impact but is also one of the most densely populated areas in the United States and Canada. This study brought together a wide range of researchers setting up their models of choice in a highly standardized experimental setup using the same geophysical datasets, forcings, common routing product, and locations of performance evaluation across the 1 million square kilometer study domain. The study comprises 13 models covering a wide range of model types from Machine Learning based, basin-wise, subbasin-based, and gridded models that are either locally or globally calibrated or calibrated for one of each of six predefined regions of the watershed. Unlike most hydrologically focused model intercomparisons, this study not only compares models regarding their capability to simulated streamflow (Q) but also evaluates the quality of simulated actual evapotranspiration (AET), surface soil moisture (SSM), and snow water equivalent (SWE). The latter three outputs are compared against gridded reference datasets. The comparisons are performed in two ways: either by aggregating model outputs and the reference to basin-level or by regridding all model outputs to the reference grid and comparing the model simulations at each grid-cell.



Citation (Journal Publication):
Mai, J., Shen, H., Tolson, B. A., Gaborit, É., Arsenault, R., Craig, J. R., Fortin, V., Fry, L. M., Gauch, M., Klotz, D., Kratzert, F., O'Brien, N., Princz, D. G., Rasiya Koya, S., Roy, T., Seglenieks, F., Shrestha, N. K., Temgoua, A. G. T., Vionnet, V., and Waddell, J. W. (2022).
The Great Lakes Runoff Intercomparison Project Phase 4: The Great Lakes (GRIP-GL)
Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-2022-113

Citation (Data Publication):
Mai, J., Shen, H., Tolson, B. A., Gaborit, É., Arsenault, R., Craig, J. R., Fortin, V., Fry, L. M., Gauch, M., Klotz, D., Kratzert, F., O'Brien, N., Princz, D. G., Rasiya Koya, S., Roy, T., Seglenieks, F., Shrestha, N. K., Temgoua, A. G. T., Vionnet, V., and Waddell, J. W. (2022).
The Great Lakes Runoff Intercomparison Project Phase 4: The Great Lakes (GRIP-GL)
FRDR. https://doi.org/10.20383/103.0598

Datasets: There are two categories of GRIP-GL datasets available. ***Part A*** includes all the data that original study participants were provided with in order to train/calibrate their model. ***Part B*** are request-only data associated with performing blind model validation. Users can access Part A at any time but, as described below, users can only be granted access to Part B of the dataset after they have accessed Part A.

Part A: Data shared for model building and calibration

These data were used for setting up models (Sec. A1 - Domain, Sec. A2 – Geophysical Datasets, Sec. A3 - Routing), to run/force the models (Sec. A4 – Forcing data), to calibrate the models regarding streamflow (Sec. A5 – Streamflow observations), and to evaluate the models (Sec. A6 – Auxiliary reference datasets). Model outputs for streamflow (Sec. A7) and additional variables (Sec. A8) are provided as well as scripts to generate the figures used in the journal publication (Mai et al., 2022).

Some data are held back:

- Sec. A3 – Routing product:
  Only contains Raven setups for 141 calibration locations; only calibration locations are labelled in routing shapefiles
- Sec. A4 – Forcing data:
  Only contains forcing data for calibration period 2000-2010 but not for validation period 2011-2017

- Sec. A5 - Streamflow observations
  Only contains observations for 141 calibration stations but not for the 71 validation stations; note that all observation files contain all observations including the validation period.

This is to guarantee that new models added later to this benchmark will be evaluated in a blind validation setup. The missing data will be provided upon request in exchange for the modelled streamflow time series at the calibration locations for the time period 2001-2010 (2000 used as warmup period). The data are specified in Part B:

- Sec. B3 – Routing product:
  Will contain Raven setups for validation locations and updated shapefiles where now also the validation locations are labelled
- Sec. B4 – Forcing data:
  Will contain forcings for entire period of time (2000-2017)
- Sec. B5 – Streamflow observations:
  Will contain streamflow observations at the 71 validation locations (all timesteps available per gauge)

A1. Study domain

These are shapefiles that specify the study domain. These files are **<u>only used for visualization</u>**. For shapefiles for individual basins see section "A3. Routing Product" (under "Shapefiles → 1_Lumped_version") below.

Data under "A1_Study domain"

- Shapefile of whole modeling domain (outer shape of entire Great Lakes basin
  ESRI shapefile: GRIP-GL_one_domain.zip
- Shapefile of seven subdomains (obsolete; one domain was discarded later)
  ESRI shapefile: GRIP-GL_seven_subdomains.zip
- Shapefile of six subdomains (six regions used in study)
  ESRI shapefile: GRIP-GL_six_subdomains.zip
- Shapefile of only lake area (shapes of Great Lake waterbodies)
  ESRI shapefile: GRIP-GL_lakes.zip

A2. Geophysical datasets

These datasets are static data. If you want to allow for a fair comparison with the published GRIP-GL suite of models, you **<u>must</u>** use these datasets. The datasets provided here have been cropped to the GRIP-GL domain. Several derived products (aggregates, regridded data, or alternative file formats) have been created for convenience of the participating modelers.

A2a. Digital Elevation Model (DEM)

The HydroSHEDS DEM has a 3'' resolution which corresponds to about 90 m at the equator. Since the upscaled HydroSHEDS DEMs with 15'' (500m) and 30'' (1km) are consistent with the best resolution dataset, the collaborators are allowed to pick the resolution most appropriate for their setup.
The data can be downloaded here: https://www.hydrosheds.org/hydrosheds-core-downloads
Flow direction, accumulation, and slope have been derived from the above-mentioned DEM.

Data under "A2_Geophysical_datasets/dem_HydroSHEDS-90m_GreatLakes"

- Conditioned SRTM DEM from HydroSHEDS, 3 second (90 m)
  ASCII raster format: rect_dem_GreatLakes.txt.zip
  NetCDF format: rect_dem_GreatLakes.nc
- Derived flow accumulation from DEM using ArcGIS
  ASCII raster format: rect_flow_accumulation_GreatLakes.txt.zip
  NetCDF format: rect_flow_accumulation_GreatLakes.nc
- Derived flow direction from DEM using ArcGIS
  ASCII raster format: rect_flow_direction_GreatLakes.txt.zip
  NetCDF file: rect_flow_direction_GreatLakes.nc
- Derived slope from DEM in TIF format
  TIF format: slope_HydroSHEDS-90m_GreatLakes.tif_v1.1.zip
- Aggregated slope to RDRS-v2 forcing grid
  tabulated TXT format: slope_HydroSHEDS-90m_GreatLakes_aggregated_v1.2.zip
  NetCDF format: slope_HydroSHEDS-90m_GreatLakes_aggregated_v1.2_RDRS-v2.nc

A2b. Soil database
As a common soil dataset for all models, we used the Global Soil Dataset for Earth System Models (GSDE) (30'' ≈ 1 km) containing 8 layers of soil to a depth of 2.3m.
The data can be downloaded here: http://globalchange.bnu.edu.cn/research/soilw.

Data under "A2_Geophysical_datasets/soilclass_GSDE_GreatLakes"

- Soil classes (1 km; average over 0-229.6cm soil depth = all layers)
  ASCII raster format: rect_GSDE_soil_class_v1.1.txt.zip
  NetCDF format: rect_GSDE_soil_class_v1.1.nc
- Soil classes (1 km; average over 0-30cm soil depth)
  ASCII raster format: rect_GSDE_soil_class_0-30cm.txt.zip
  NetCDF format: rect_GSDE_soil_class_0-30cm.nc
- Soil classes (1 km; average over 30-100cm soil depth)
  ASCII raster format: rect_GSDE_soil_class_30-100cm.txt.zip
  NetCDF format: rect_GSDE_soil_class_30-100cm.nc
- Soil textures (sand, silt, clay, bulk density; 1 km; averaged over all layers and per layer)
  ASCII format: rect_GSDE_soil_texture.zip
  NetCDF format: rect_GSDE_soil_texture.zip
- Soil classes and textures aggregated to RDRS-v2 forcing grid
  tabulated TXT format: soilclass_GSDE_GreatLakes_aggregated_v1.3.zip
  NetCDF format: soilclass_GSDE_GreatLakes_aggregated_v1.3_RDRS-v2.nc

A2c. Landcover dataset
As a common landcover dataset for all models we used the NALCMS product including 19 land cover classes for North America (30m, Landsat, 2010 from Mexico and Canada, 2011 for U.S.).
The data can be downloaded here:
http://www.cec.org/north-american-environmental-atlas/land-cover-2010-landsat-30m/
A documentation can be found here:
http://www.cec.org/wp-content/uploads/wpallimport/files/Atlas/Files/2010nalcms30m/nalcms_2010_30m_metadata.doc

Data under "A2_Geophysical_datasets/landcover_NALCMS_GreatLakes"

- Landcover classes (30m; original meter-grid)
  TIF format: landcover_NALCMS_GreatLakes.tif_v1.1.zip
  ASCII raster format: rect_landcover_NACLMS_GreatLakes_raw-meter-grid_v1.1.txt.zip
- Landcover classes (≈ 30m; projected to degree-grid)
  ASCII raster format: rect_landcover_NACLMS_GreatLakes_projected-degree-grid_v1.1.txt.zip
- Landcover classes aggregated to RDRS-v2 forcing grid
  tabulated TXT format: landcover_NALCMS_GreatLakes_aggregated_v1.2.zip
  NetCDF format: landcover_NALCMS_GreatLakes_aggregated_v1.2_RDRS-v2.nc

A3. Routing product
The data provided in the routing product generated with the BasinMaker software (http://hydrology.uwaterloo.ca/basinmaker/index.html) are shapefiles that contain the spatial discretization of the watersheds as well as properties associated to each subbasin/HRU. The data also contain the Raven setups that can then be used to couple output of any model (gridded or subbasin-based surface runoff and possibly groundwater recharge) with the routing product as done by most (semi-) distributed models participating in this study, i.e., 'blended-raven', 'mesh-class-raven', 'mesh-svs-raven', 'swat-raven', 'vic-raven', 'watflood-raven'. A tutorial on how to use these data and couple them to your model can be found on the Raven website (http://raven.uwaterloo.ca) under "Downloads". A direct link to the routing tutorial is the following: http://raven.uwaterloo.ca/files/RavenTutorial6.zip.

A3a. Raven setups
Raven setups contain all information to run Raven in routing mode. Some adjustments might be required for your specific model output. Please refer to the tutorial for further details (http://raven.uwaterloo.ca/files/RavenTutorial6.zip).

Data under "A3_Routing_product/Raven_setups/"

- 1_Local_calibration Raven routing setup files for each of the 141 calibration watersheds individually; Contains only calibration locations

- 2_Regional_calibration — Raven routing setups for individual regions (six regions defined: HUR, ERI, MIC, ONT, OTT, SUP); Contains only calibration locations but all of them are included in one setup based on the calibration locations-merged routing network. This is for routing in only the case study calibration watersheds (141 in total) within a given merged region in one Raven model.
- 3_Global_calibration — Raven routing setup for routing of the entire Great Lakes watershed; Contains calibration and validation locations but all of them are included in one setup based on the entire Great Lakes routing network. This is for routing in the case study calibration and validation watersheds (212 in total) within the Great Lakes region.

NOTE: The regional and global setups are consistent among themselves such that they can both be used for regional or global calibration. The difference is the regional setup is based on the 141 calibration locations-merged shapefile but the global setup is based on the entire Great Lakes routing network shapefile. This makes the two setups differ in only the subbasin/HRU encoding (i.e., subbasin and HRU IDs) but they share the same attributes.

A3b. Shapefiles
Shapefiles of lumped calibration watersheds (1_Lumped_version) as well as shapefiles discretizing the domain into subbasins for only calibration locations (2_Regional_version) or the entire Great Lakes watershed including watersheds that are not modelled in this project (3_Global_version).

Data under “A3_Routing_product/ Shapefiles/”

- 1_Lumped_version — Shapefile containing lumped basin geometries of only calibration locations all merged in one file
- 2_Regional_version — Shapefile containing only calibration locations with each of them discretized into subbasins
- 3_Global_version — Shapefile containing subbasin-based discretization of the entire Great Lakes domain including regions that are not modelled in this project.

NOTE: Users who want to use the global version need to be aware of that there is a sub-domain named “Not_in_GRIP_GL”, which is not considered in this project, and those subbasins in this “Not_in_GRIP_GL” sub-domain have not been QA/QC’d when we produced this shapefile

A4. Forcing datasets
The dataset we use here is the regional deterministic reforecast system (RDRS-v2) product (Gasset et al., 2021) (https://doi.org/10.5194/hess-25-4917-2021). The RDRS-v2 dataset can be downloaded through CaSPAr (www.caspar-data.ca) under product “RDRS_v2”. The data have an hourly temporal resolution and a spatial resolution of about 10 km. It is available for 18 years from 2000 to 2017. The RDRS-v2 dataset covers North and Central America. The domain of the Great Lakes was cropped for this project, and this cropped RDRS-v2 dataset is archived here. The following data have been prepared for this project. They are available as NetCDF on a rotated lat-lon grid. Be aware these data are in UTC time; a time shift of 5h to local time is required before comparing simulations to observed streamflow.

Data under “A4_Forcing_datasets/RDRS_v2/2000-2010/”

- Quantity of precipitation (CaPA) (unit: [m], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_A_PR0_SFC.nc
- Air temperature (unit: [°C], vertical level: 1.5m)
  NetCDF format: RDRS_v2_forcings_final_P_TT_1.5m.nc
- Downward solar flux (unit: [W/m2], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_FB_SFC.nc
- Surface incoming infrared flux (unit: [W/m2], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_FI_SFC.nc
- Relative humidity (fraction) (unit: [-], vertical level: 1.5m)
  NetCDF format: RDRS_v2_forcings_final_P_HR_1.5m.nc
- Specific humidity (unit: [kg/kg], vertical level: ~40m)
  NetCDF format: RDRS_v2_forcings_final_P_HU_09944.nc
- Surface pressure (unit: [mb], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_P0_SFC.nc
- U-comp. of the wind (along the grid X axis) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_UU_10m.nc
- V-comp. of the wind (along the grid Y axis) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_VV_10m.nc

- Wind Modulus (derived using UU and VV) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_UVC_10m.nc

The above gridded data have been aggregated to lumped forcings for the convenience of several modelers (Raven, HYMOD2, LBRM, LSTM). All these data are in local time.

Data under "A4_Forcing_datasets/RDRS_v2/aggregated/"

- Aggregated (daily) forcings for **HYMOD2-lumped** (PRECIP, TEMP_DAILY_MIN, TEMP_DAILY_MAX, RELHUMIDITY)
  ASCII files: cal-hymod2-lumped.zip
- Aggregated (daily) forcings for **LBRM-CC-lumped** (TEMP_DAILY_AVE, TEMP_DAILY_MIN, TEMP_DAILY_MAX, PRECIP)
  ASCII files: cal-lbrm-cc-lumped.zip
- Aggregated (hourly) forcings for **LSTM-lumped** (RDRS_v2_A_PR0_SFC_m, RDRS_v2_P_FB_SFC_W_m2, RDRS_v2_P_FI_SFC_W_m2, RDRS_v2_P_HU_09944_kg_kg, RDRS_v2_P_P0_SFC_mb, RDRS_v2_P_TT_1.5m_degC, RDRS_v2_P_UU_10m_kts, RDRS_v2_P_UUC_10m_kts, RDRS_v2_P_UVC_10m_kts, RDRS_v2_P_VV_10m_kts, RDRS_v2_P_VVC_10m_kts)
  ASCII files: cal-lstm-lumped.zip
- Aggregated (daily) forcings for lumped Raven models, i.e., **GR4J-lumped, HMETS-lumped, Blended-lumped** (TEMP_DAILY_AVE, TEMP_DAILY_MIN, TEMP_DAILY_MAX, PRECIP)
  ASCII files: cal-raven-lumped.zip
- Aggregated (hourly) forcings for semi-distributed Raven models, i.e., **Blended-Raven** (PRECIP, TEMP)
  NetCDF files: cal-raven-subbasin.zip

NOTE: The gridded data are in UTC time. The local time is UTC-5h. Please make sure to shift the forcings or your simulated (hourly) streamflow simulations accordingly before comparing to (local) streamflow observations.

NOTE: The data are only shared for the warmup and calibration period (2000 and 2001-2010, respectively) to allow for blind validation experiments of additional models in future. The forcing data for the validation period will be shared as soon as your the streamflow time series of the calibrated model have been shared.

NOTE: Additional variables that none of the GRIP-GL modelling teams used (but could have) are available in the original RDRS-v2 product (see https://github.com/julemai/CaSPAr/wiki/Available-products#list-of-available-variables-in-rdrs-v2-and-rdrs-v21) and can be shared upon request.

A5. Streamflow observation datasets

The streamflow observations have been downloaded from Water Survey Canada (WSC) and United States Geological Survey (USGS) and are available in the original downloaded format (one file per gauge; CSV from WSC in [m^3/s] and TXT from USGS in [ft^3/s]) as well as in standardized NetCDF format (one file containing all gauges; all in [m^3/s]). The data for only the calibration stations are shared; the data for validation stations will be shared as soon as calibrated model outputs are provided for the calibration stations (see Part B of this document below). This is to ensure a fair comparison with the current models from GRIP-GL that all underwent a blind validation experiment. The data available for each station here includes the entire period of record for the station not just the GRIP-GL calibration period, including time periods that might be later used for validation). The gauges are grouped into "objective 1" gauges which are only stations that can be classified as low-human impact (in total 66 stations) and "objective 2" gauges which are gauges that are most downstream to a lake or Ottawa River (in total 104 stations).

Data under "A5_Streamflow_observation_datasets/calibration/"

- Metadata for gauges
  CSV files: objective_1/gauge_info.csv
  CSV files: objective_2/gauge_info.csv
- Raw observations downloaded from WSC in [m^3/s]
  ASCII files: objective_1/<gauge-ID>.csv
  ASCII files: objective_2/<gauge-ID>.csv
- Raw observations downloaded from USGS in [ft^3/s]
  ASCII files: objective_1/<gauge-ID>.txt
  ASCII files: objective_2/<gauge-ID>.txt
- Streamflow with consistent unit [m^3/s] and missing value
  NetCDF files: objective_1/all_gauges.nc
  NetCDF files: objective_2/all_gauges.nc

A6. Auxiliary reference datasets
We used three additional datasets to allow models to be evaluated. This means these datasets are not used for calibration. We used data for surface soil moisture, evapotranspiration, and snow water equivalent. The original data were downloaded and then cropped to the GRIP-GL domain. Only data available between 2000 and 2017 were kept in the dataset. The datasets were converted to a standardized NetCDF format. No datapoints were perturbed or filtered during this process.

A6a. Surface Soil Moisture
As a reference dataset for surface soil moisture, we used the GLEAM product (raw format: NetCDF). The dataset has a 0.25° spatial and daily temporal resolution. It is available from 2003 (January 1st) to 2020 (July 31st). The satellite data obtained meteorologic variables and vegetation are used to derive evapotranspiration estimates using the Priestley-Taylor model. The raw data can be downloaded here (product v3.5b; variable #10 - SMsurf): https://www.gleam.eu/#datasets. For further information on this dataset please refer to:

1. Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geoscientific Model Development*, 10, 1903–1925, doi: 10.5194/gmd-10-1903-2017, 2017.
2. Miralles, D.G., Holmes, T.R.H., de Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J.: Global land-surface evaporation estimated from satellite-based observations, *Hydrology and Earth System Sciences*, 15, 453–469, doi: 10.5194/hess-15-453-2011, 2011.

The data have been processed for the GRIP-GL project by cropping the domain to the Great Lakes including the Ottawa River (no temporal cropping of the dataset).

Data under "A6_Auxiliary_reference_datasets/ssm/gleam-smsurf"

- Gridded reference dataset
  NetCDF files: gleam-3.5b-smsurf.nc
- Reference dataset aggregated to basin-level (only 141 calibration basins)
  NetCDF files: gleam-3.5b-smsurf_141_basins.nc
- Reference dataset aggregated to basin-level (only 71 validation basins)
  NetCDF files: gleam-3.5b-smsurf_71_basins.nc
- Reference dataset aggregated to basin-level (all basins)
  NetCDF files: gleam-3.5b-smsurf_212_basins.nc

A6b. Evapotranspiration
As a reference dataset for evapotranspiration, we used the GLEAM product (raw format: NetCDF). The dataset has a 0.25° spatial and daily temporal resolution. It is available from 2003 (January 1st) to 2020 (July 31st). The satellite data obtained meteorologic variables and vegetation are used to derive evapotranspiration estimates using the Priestley-Taylor model. The raw data can be downloaded here (product v3.5b; variable #1 - E): https://www.gleam.eu/#datasets. For further information on this dataset please refer to:

1. Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geoscientific Model Development*, 10, 1903–1925, doi: 10.5194/gmd-10-1903-2017, 2017.
2. Miralles, D.G., Holmes, T.R.H., de Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J.: Global land-surface evaporation estimated from satellite-based observations, *Hydrology and Earth System Sciences*, 15, 453–469, doi: 10.5194/hess-15-453-2011, 2011.

The data have been processed for the GRIP-GL project by cropping the domain to the Great Lakes including the Ottawa River (no temporal cropping of the dataset).

Data under "A6_Auxiliary_reference_datasets/aet/gleam-e"

- Gridded reference dataset
  NetCDF files: gleam-3.5b-e.nc
- Reference dataset aggregated to basin-level (only 141 calibration basins)
  NetCDF files: gleam-3.5b-e_141_basins.nc
- Reference dataset aggregated to basin-level (only 71 validation basins)
  NetCDF files: gleam-3.5b-e_71_basins.nc
- Reference dataset aggregated to basin-level (all basins)

NetCDF files: gleam-3.5b-e_212_basins.nc

A6c. Snow water equivalent

As a reference dataset for snow water equivalent, we used the ERA5-Land product (raw format: NetCDF). The dataset has a 0.1° spatial and daily temporal resolution. ERA5-Land combines many historical observations into global estimates of surface variables using advanced modelling and data assimilation systems. ERA5-Land is available for the entire study period and domain on a daily temporal scale and a 0.1° regular grid. The data are given in [m water equivalent] and have been converted to [kg/m2]=[mm] to allow comparison with model outputs. The raw data can be downloaded here (variable – "Snow depth water equivalent"): https://doi.org/10.24381/cds.e2161bac. The files were generated using Copernicus Climate Change Service information (2022). The data have been processed for the GRIP-GL project by cropping the domain to the Great Lakes including the Ottawa River. The time period was cropped to 2000 to 2017 from ERA5-Land hourly data from 1950 to present.

Data under "A6_Auxiliary_reference_datasets/swe/era5-swe"

- Gridded reference dataset
  NetCDF files: era5-swe-sd.nc
- Reference dataset aggregated to basin-level (only 141 calibration basins)
  NetCDF files: era5-swe-sd_141_basins.nc
- Reference dataset aggregated to basin-level (only 71 validation basins)
  NetCDF files: era5-swe-sd_71_basins.nc
- Reference dataset aggregated to basin-level (all basins)
  NetCDF files: era5-swe-sd_212_basins.nc

A7. Model outputs for streamflow

Simulated daily streamflow for models participating.

Model names are (13 models):

'blended-lumped', 'blended-raven', 'gem-hydro-watroute', 'gr4j-lumped', 'hmets-lumped', 'hymod2-lumped', 'lbrm-cc-lumped', 'mesh-class-raven', 'mesh-svs-raven', 'lstm-lumped', 'swat-raven', 'vic-raven', 'watflood-raven'

Phases are:

1 - calibration (66 stations for obj 1, 104 stations for obj 2; period: 2001-2010)
2 - temporal validation (66 stations for obj 1, 104 stations for obj 2; period: 2011-2017)
3 - spatial validation (33 stations for obj 1, 52 stations for obj 2; period: 2001-2010)
4 - spatio-temporal validation (33 stations for obj 1, 52 stations for obj 2; period: 2011-2017)

Objectives are:

1 - low-human impact watersheds
2 - gauges that are most downstream to a lake or Ottawa River (can be low-human impact or not)

Data under "A7_Model_outputs_for_streamflow/"

- Simulated streamflow per model
  NetCDF files: <model-name>_phase_<phase>_objective_<objective>.nc

NOTE: The files can contain longer time series depending on how the files were provided. The performance was only derived based on the available data for the time period indicated, i.e., 2001-2010 for calibration (phase 1) and spatial validation (phase 3) as well as 2011-2017 for temporal validation (phase 2) and spatio-temporal validation (phase 4). Note that the published manuscript did not assess results considering objective 1 versus 2 designations (each of these lumped together in all published analyses).

A8. Model outputs for auxiliary variables

Model outputs for the three additional variables evaluated in this project.

Model names are (12 models; no "lstm-lumped'):

'blended-lumped', 'blended-raven', 'gem-hydro-watroute', 'gr4j-lumped', 'hmets-lumped', 'hymod2-lumped', 'lbrm-cc-lumped', 'mesh-class-raven', 'mesh-svs-raven', 'swat-raven', 'vic-raven', 'watflood-raven'

The variables are:

aet – actual evapotranspiration
ssm – surface soil moisture
swe – snow water equivalent

The model outputs are compared to reference datasets (see Sec. A6) based on two different spatial resolutions:

basin – model outputs are aggregated to basin-averages (each basin is a calibration/validation watershed)
gridcell – model outputs are regridded to grid of reference datasets (see Sec. A6)

The original model output was provided as lumped (model discretization=basin), semi-distributed (model discretization=subbasin) or gridded (model discretization=grid). All these outputs were aggregated to "basin-wise" or "gridcell-wise".

Data under "A8_Model_outputs_for_auxiliary_variables/basin-wise/" (3 variables x 12 models = 36 files total):

- Model outputs of actual evapotranspiration as basin averages (NetCDF variable is "aet")
  NetCDF files: <model-name>_aet_<model-discretization>2basin.nc
- Model outputs of surface soil moisture as basin averages (NetCDF variable is "ssm")
  NetCDF files: <model-name>_ssm_<model-discretization>2basin.nc
- Model outputs of snow water equivalent as basin averages (NetCDF variable is "swe")
  NetCDF files: <model-name>_swe_<model-discretization>2basin.nc

Data under "A8_Model_outputs_for_auxiliary_variables/gridcell-wise/" (3 variables x 12 models = 36 files total):

- Model outputs of actual evapotranspiration as basin averages (NetCDF variable is "aet")
  NetCDF files: <model-name>_aet_<model-discretization>2grid.nc
- Model outputs of surface soil moisture as basin averages (NetCDF variable is "ssm")
  NetCDF files: <model-name>_ssm_<model-discretization>2grid.nc
- Model outputs of snow water equivalent as basin averages (NetCDF variable is "swe")
  NetCDF files: <model-name>_swe_<model-discretization>2grid.nc

A9. Scripts
These are the scripts used to create figures of publication. Data required for plots are generated from data provided in Sec.s A1 to A8.

Scripts under "A9_Scripts":

- plot.sh — Wrapper script to plot figures; set dofig<ID> to 1 to select figure for plotting
- figure_1.py — Plots figure 1 of publication
- figure_2.py — Plots figure 2 of publication
- figure_3.py — Plots figure 3 of publication and S3, S4, S5 of supplements
  figure_3.json — Data required for figure 3
  (if does not exist it will be created from data which just takes longer)
- figure_4.py — Plots figure 4 of publication
- figure_4.json — Data required for figure 4
  (if does not exist it will be created from data which just takes longer)
- figure_5-7.py — Plots figure 5-7 of publication
  figure_5-7_E.nc — Data required for figure 5
  (if does not exist it will be created from data which just takes longer)
  figure_5-7_SMsurf.nc — Data required for figure 6
  (if does not exist it will be created from data which just takes longer)
  figure_5-7_sd.nc — Data required for figure 7
  (if does not exist it will be created from data which just takes longer)
- figure_8.py — Plots figure 8 of publication
- figure_S1.py — Plots figure S1 of supplements
  figure_S1.nc — Data required for figure S1
- figure_S2/figure_S2.tex — Plots figure S2 of supplements
- lib/ — Additional convenience tools used from JAMS library
  (https://github.com/mcuntz/jams_python)
- MachineLearning/ — Folder containing configuration files and scripts to reproduce the LSTM setup and training
- compare_ERA5-swe_CanSWE/ — Folder containing scripts used to compare ERA5 SWE with station observations in CanSWE dataset.

B. Data shared upon request (for blind validation)

These datasets are shared upon request in exchange for requestors to provide their simulated streamflow at the 141 calibration locations. This is supposed to ensure a blind validation, i.e., holding back forcings data covering validation period and observations at validation locations. Of course, there is no guarantee that future modellers do not access this validation data by other means.

Any future blind validation modelling results will only be considered for addition to the GRIP-GL website as post-publication blind validation results if modellers also confirm that their simulated results are truly validation results (e.g., their validation period/location performance was never used to influence model development or model calibration).

B3. Routing product
Additional Raven setup files for the validation locations as well as shapefiles labelling now also the validation locations. File formats, shapefile attributes, etc. are the same as provided under B3; it is solely dataset for 71 more stations.

B3a. Raven setups
Raven setups contain all information to run Raven in routing mode. Some adjustments might be required for your specific model output. A tutorial on how to use these data and couple them to your model can be found on the Raven website (http://raven.uwaterloo.ca) under "Downloads". A direct link to the routing tutorial is the following: http://raven.uwaterloo.ca/files/RavenTutorial6.zip. Please refer to this tutorial for further details.

Data under "A3_Routing_product/Raven_setups/"

- 1_Local_calibration — Raven routing setup files for each of the 71 calibration watersheds individually; Contains only validation locations
- 2_Regional_calibration — Raven routing setups for individual regions (six regions defined: HUR, ERI, MIC, ONT, OTT, SUP); Contains only validation locations but all of them are included in one setup based on the validation locations-merged routing network. This is for routing in only the case study validation watersheds (71 in total) within a given merged region in one Raven model.
- 3_Global_calibration — Raven routing setup for routing of the entire Great Lakes watershed; Contains calibration and validation locations but all of them are included in one setup based on the entire Great Lakes routing network. This is for routing in the case study calibration and validation watersheds (212 in total) within the Great Lakes region.

B3b. Shapefiles
Shapefiles of lumped validation watersheds (1_Lumped_version) as well as shapefiles discretizing the domain into subbasins for only validation locations (2_Regional_version) or the entire Great Lakes watershed including watersheds that are not modelled in this project (3_Global_version).

Data under "A3_Routing_product/ Shapefiles/"

- 1_Lumped_version — Shapefile containing lumped basin geometries of only validation locations all merged in one file
- 2_Regional_version — Shapefile containing only validation locations with each of them discretized into subbasins
- 3_Global_version — Shapefile containing subbasin-based discretization of the entire Great Lakes domain including regions that are not modelled in this project.

NOTE: Users who want to use the global version need to be aware that there is a sub-domain named "Not_in_GRIP_GL", which is not considered in this project, and those subbasins in this "Not_in_GRIP_GL" sub-domain have not been QA/QC'd when we produced this shapefile.

B4. Forcing datasets
Augmenting the forcing datasets by the validation period (2010-2017). Data formats, grid cells, NetCDF attributes, etc. are the same as the data provided under A4; it is solely a longer record.

Data under "B4_Forcing_datasets/RDRS-v2/2000-2017/"

- Quantity of precipitation (CaPA) (unit: [m], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_A_PR0_SFC.nc
- Air temperature (unit: [°C], vertical level: 1.5m)
  NetCDF format: RDRS_v2_forcings_final_P_TT_1.5m.nc
- Downward solar flux (unit: [W/m2], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_FB_SFC.nc
- Surface incoming infrared flux (unit: [W/m2], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_FI_SFC.nc
- Relative humidity (fraction) (unit: [-], vertical level: 1.5m)
  NetCDF format: RDRS_v2_forcings_final_P_HR_1.5m.nc
- Specific humidity (unit: [kg/kg], vertical level: ~40m)

NetCDF format: RDRS_v2_forcings_final_P_HU_09944.nc

- Surface pressure (unit: [mb], vertical level: SFC)
  NetCDF format: RDRS_v2_forcings_final_P_P0_SFC.nc
- U-comp. of the wind (along the grid X axis) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_UU_10m.nc
- V-comp. of the wind (along the grid Y axis) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_VV_10m.nc
- Wind Modulus (derived using UU and VV) (unit: [kts], vertical level: 10m)
  NetCDF format: RDRS_v2_forcings_final_P_UVC_10m.nc

Data under "B4_Forcing_datasets/RDRS-v2/aggregated/"

- Aggregated (daily) forcings for **HYMOD2-lumped** (PRECIP, TEMP_DAILY_MIN, TEMP_DAILY_MAX, RELHUMIDITY)
  ASCII files: val-temporal-hymod2-lumped.zip
  ASCII files: val-spatial-hymod2-lumped.zip
- Aggregated (daily) forcings for **LBRM-CC-lumped** (TEMP_DAILY_AVE, TEMP_DAILY_MIN, TEMP_DAILY_MAX, PRECIP)
  ASCII files: val-temporal -lbrm-cc-lumped.zip
  ASCII files: val-spatial -lbrm-cc-lumped.zip
- Aggregated (hourly) forcings for **LSTM-lumped** (RDRS_v2_A_PR0_SFC_m, RDRS_v2_P_FB_SFC_W_m2, RDRS_v2_P_FI_SFC_W_m2, RDRS_v2_P_HU_09944_kg_kg, RDRS_v2_P_P0_SFC_mb, RDRS_v2_P_TT_1.5m_degC, RDRS_v2_P_UU_10m_kts, RDRS_v2_P_UUC_10m_kts, RDRS_v2_P_UVC_10m_kts, RDRS_v2_P_VV_10m_kts, RDRS_v2_P_VVC_10m_kts)
  ASCII files: val-temporal -lstm-lumped.zip
  ASCII files: val-spatial -lstm-lumped.zip
- Aggregated (daily) forcings for lumped Raven models, i.e. **GR4J-lumped, HMETS-lumped, Blended-lumped** (TEMP_DAILY_AVE, TEMP_DAILY_MIN, TEMP_DAILY_MAX, PRECIP)
  ASCII files: val-temporal-raven-lumped.zip
  ASCII files: val-spatial-raven-lumped.zip
- Aggregated (hourly) forcings for semi-distributed Raven models, i.e. **Blended-Raven** (PRECIP, TEMP)
  NetCDF files: val-temporal-raven-subbasin.zip
  NetCDF files: val-spatial-raven-subbasin.zip

B5. Streamflow observation datasets

Streamflow observations for the validation locations. Data formats, units, file formats, NetCDF attributes, etc. are the same as the data provided under B4; it is solely data at 71 additional locations.

Data under "B5_Streamflow_observation_datasets/validation/"

- Raw observations downloaded from WSC in [m^3/s]
  ASCII files: objective_1/<gauge-ID>.csv
  ASCII files: objective_2/<gauge-ID>.csv
- Raw observations downloaded from USGS in [ft^3/s]
  ASCII files: objective_1/<gauge-ID>.txt
  ASCII files: objective_2/<gauge-ID>.txt
- Streamflow with consistent unit [m^3/s] and missing value
  NetCDF files: objective_1/all_gauges.nc
  NetCDF files: objective_2/all_gauges.nc