How to reproduce

This document is meant to guide anyone who wishes to reproduce the entirety or a part of the scripts used in Wiersma et al. (2022). I have done my best to keep it somewhat structured, but I should give a disclaimer that I'm a scientist, not a programmer, and this is without doubt reflected in the quality of the code. Following this guide alone is probably complicated, so I strongly encourage anyone wishing to work with this script to contact me beforehand.

Different scripts and files

Script	**What does it do**
1_GHMGGM_preprocessing.py	• Helps to select GRDC stations and to create subbasins and clonemaps based on their location • rasterizes and resamples GloGEM data • prevents their spilling into other basins • saves the result in a NetCDF file
2_GHMGGM_Modelruns.py	Loads the NetCDF-files of (1/3): Preprocessing and couples them to PCR-GLOBWB 2 • Loads and adjusts the PCRGLOB inifile • Calls the PCRGLOB docker and initializes the model • Has a setting for doing only spinup • Runs PRCGLOB coupled or uncoupled for one basin • Saves modeled discharge at gauging station • Saves daily parameter fields into netcdf (optional) • Can be made into a function and run together with eval_inputs.py • for job arrays

3_GHMGGM_Hydrographs.py	Loads the modelled hydrographs and plots them
	• Single basin all years
	• One year with a selection of basins
	Calculates the objective functions which can be plotted in the next script
	• ND
	• Ratio
	• RRD
	• Benchmark efficiency
	• Flow duration curve efficiency
	• Mass balance efficiency
	• Standalone GHM evaluation
	  ○ NSE
	  ○ Calendar benchmark efficiency (Schaefli&Gupta2007)

4_GHMGGM_ObjectiveFunctionPlots.py	Loads the objective functions (from script 3) and plots them
5_GHMGGM_MassImbalance.py	Calculates the annual runoff difference between the coupled model and the benchmark and the contribution of • Snow towers • Glacier mass loss • Spilling prevention to this annual runoff difference for the basins Alsek, Columbia, Oelfusa and Rhone
aux_Adjust_PCRGLOB_ landcover_fractions.py	Takes the PCRGLOB grassland (short vegetation) landcover fraction map and subtracts the RGI glacier cover fraction from it.
aux_Plot_Worlmap.py	This script loads the basin shapefiles and plots them on a world map with the glacierization degree in blue hue

Files needed at the start

- External files

PCR-GLOBWB 2 datasets	https://github.com/UU-Hydro/PCR-GLOBWB_model https://doi.org/10.5194/gmd-11-2429-2018
GloGEM outputs	On request from Huss and Hock (2018), a more recent version is now available https://doi.org/10.1038/s41558-017-0049-x
ERA-Interim T & P	Through eWaterCycle or manual download, future studies ought better work with ERA5
Hydrosheds (lvl 12)	https://www.hydrosheds.org/page/hydrobasins
GRDC metadata & runoff data	https://www.bafg.de/GRDC/EN
Hydrobanque Rhone discharge	https://www.hydro.eaufrance.fr/
Randolph Glacier Inventory	https://doi.org/10.7265/N5-RGI-60 https://www.glims.org/RGI/
River vectors from Yan et al. 2019 (optional)	https://doi.org/10.6084/m9.figshare.8044184.v5

- Custom files (at Zenodo)
    - GHMGGM_basin_info.csv
    - Yml files to create the conda environments with the necessary packages
        - GHMGGM_ewc.yml for the eWaterCycle runs on Linux
        - GHMGGM.yml for all other scripts
    - Glob_glac_fac_shp
    - global_masked_vectorgrid.shp
    - Clonemaps
    - Basin and subbasin shapefiles
    - Adjusted landcover fractions (to skip the aux_Adjust_PCRGLOB_landcover_fractions.py)
    - PCR-GLOBWB 2 ini file (setup_05min_non-natural.ini)

Folder structure

- Code
    - Files
        - basin_geojsons
        - subbasin_geojsons
        - isout
        - isglac
        - isbasin
        - clones
        - globglacfrac
            - Containing *glob_glac_frac.shp* and *global_masked_vectorgrid.shp*
        - HH2018_validation
            - Containing per basin *Discharge_catchment, Overview* and *Annual_balance*
        - GG_rasterized_monthly (auxiliary)
        - GG_resampled_daily (auxiliary)
        - GG_spilling_prevented (auxiliary)
        - GG_final (auxiliary)

        - ERA_Interim
        - glaciers_nc
        - Hydrosheds
        - Yan2019rivers
        - grassf_adjusted_landcover_PCRGLOB
        - *GRDC_Stations.csv*
        - *GHMGGM_basin_info.csv*
        - *Setup_05min_non-natural.ini*
    - Figures
        - Basin_maps
        - Ensembleplots
        - Main_plots
    - Output

eWaterCycle version

This study was done with an old version of eWaterCycle, with the following package versions

- Ewatercycle-parametersetdb = 1.0.0
- Grpcio 1.27.2
- Grpc4bmi 0.3

This older version was installed with the following (outdated) instructions https://github.com/jeromaerts/external_user_guide/tree/pcrglobwb_install_instructions.

Any future runs ought better use the newer eWaterCycle package (https://ewatercycle.readthedocs.io/en/latest/system_setup.html). The necessary changes in the script should be minor. Keep in mind that eWaterCycle should be run on Linux, the GHMGGM_ewc.yml file contains a conda environment that should support the eWaterCycle model runs.