High Mountain Asia CMIP6 Monthly and Yearly Water Balance Projections, 2016-2099 for Parts of Afghanistan, Tajikistan, Kyrgyzstan, and Pakistan, Version 1

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Lammers, R. B., A. A. Prusevich, & S. J. Glidden. 2023. *High Mountain Asia CMIP6 Monthly and Yearly Water Balance Projections, 2016-2099 for Parts of Afghanistan, Tajikistan, Kyrgyzstan, and Pakistan, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/F87QWGU8K14W. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/HMA2_WBP

TABLE OF CONTENTS

1 DATA DESCRIPTION

This High Mountain Asia (HMA) data set comprises a suite of monthly and yearly water balance model (WBM) projections for the years 2016 – 2099, covering parts of Afghanistan, Tajikistan, Kyrgyzstan, and Pakistan (primarily the headwaters of the Amu Darya and Indus River basins).

Projections are available for 12 CMIP6[1] global climate models (GCMs) and two Shared Socioeconomic Pathways (SSPs): SSP 2-4.5 and SSP 5-8.5. The data were generated using the University of New Hampshire WBM.

A historical run is also available for the years 1980 through 2018, generated from ERA5 reanalysis temperature data and ensemble precipitation estimates.

> ⓘ The Shared Socioeconomic Pathways, or SSPs, comprise five different projected climate futures based on both emissions scenarios and socioeconomic factors, such as population, economic growth, education, urbanization, and technological development (See Riahi et al 2017 for details.).

1.1 Parameters

The WBM was run for SSP 2-4.5 and SSP 5-8.5 using the output from each of the CMIP6 GCMs listed in Table 3 below. WBM outputs were saved at monthly time steps for the full projection period and then aggregated into annual values. Both the monthly and annual values are provided for each year of the projection (2016 – 2099) as separate files.

The historical data (1980 – 2018) were generated with a single model run spanning 1980 through 2018. These data are also provided at monthly and annual time steps for each year of the run.

1.1.1 Monthly

Table 1 below lists the parameters in the monthly data files. The first 26 parameters consist of WBM outputs, while the last four contain temporal and spatial information.

The monthly WBM output parameters are stored in 3D, 32-bit, floating-point format data fields, with dimensions of 12×105×194 (month × latitude × longitude). A value of -9999 indicates missing data.

[1] Coupled Model Intercomparison Project Phase 6. For more information, see "CMIP6: the next generation of climate models explained."

Table 1. Data Fields and Descriptions

Parameter	Description
AQF_H_Depth	Aquifer head depth (m)
Aqf_Storage	Aquifer groundwater storage (mm)
UGW	Accumulated unsustainable groundwater use (mm)
airT	Air temperature (interpolated input, °C)
baseflow	Baseflow (mm/day)
cropArea	Area of cropland (km^2)
discharge	River discharge (m^3/s)
domUseGross	Gross domestic water withdrawal (mm/day)
endoStrg	Endorheic lake water storage (m^3)
evapotrans	Evapotranspiration (mm/day)
glArea[2]	Glacier area (fraction of cell)
glMelt[2]	Glacier melt (m^3/s)
glVolume[2]	Glacier water storage volume (mm)
grdWater	Groundwater storage (mm)
indUseGross	Gross industrial water withdrawal (mm/day)
irrArea	Irrigated cropland area (km^2)
irrigationGross	Gross irrigation (mm/day)
precip	Precipitation (interpolated, mm/day)
resStorage	Reservoir storage (m^3)
runoff	Total runoff (mm/day)
snowFall	Snowfall (mm/day)
snowMelt	Snow melt (mm/day)
snowPack	Snowpack, as snow water equivalent (mm)
soilMoist	Soil moisture content (mm)
stkUseGross	Gross livestock water withdrawal (mm/day)
total_mass	Terrestrial water storage (mm)
time	Simulation time (32-bit integer, days since 1 January, 1900)
lat	Latitude (64-bit float, °N)
lon	Longitude (64-bit float, °E)
crs	Coordinate reference system details (string)

[2] Input from Rounce et al., 2022

1.1.2 Yearly

Yearly files include all the data fields in Table 1, plus 26 additional data fields that contain standard deviations for each of the 26 WBM outputs. Standard deviation data fields have “_sigma” appended to the name of their corresponding WBM output: for example, AQF_H_Depth_sigma, Aqf_Storage_sigma, etc.

The yearly WBM output parameters (and their standard deviations) are stored in 2D, 32-bit, floating-point format, with dimensions of 105 × 194 (latitude × longitude). A value of -9999 indicates missing data.

1.2 File Information

1.2.1 Format

NetCDF-4

1.2.2 Naming Convention

Data files utilize the following naming convention:

Convention
HMA2_WBP_[GCM]_SSP[nnn]_[resolution]_[yyyy]_V[xx.x].nc
HMA2_WBP_HISTORICAL_[resolution]_[yyyy]_V01.0.nc

Example
HMA2_WBP_BCC-CSM2-MR_SSP245_MONTHLY_2018_V01.0.nc
HMA2_WBP_BCC-CSM2-MR_SSP245_YEARLY_2018_V01.0.nc
HMA2_WBP_BCC-CSM2-MR_SSP585_MONTHLY_2018_V01.0.nc
HMA2_WBP_BCC-CSM2-MR_SSP585_YEARLY_2018_V01.0.nc

HMA2_WBP_HISTORICAL_MONTHLY_1980_V01.0.nc
HMA2_WBP_HISTORICAL_YEARLY_1980_V01.0.nc

The following table describes the variables in the file naming convention:

Table 2. File Naming Convention Variables and Descriptions

Variable	Description
HMA2_WBP	High Mountain Asia CMIP6 Monthly and Yearly Water Balance Projections, 2016-2099 for Parts of Afghanistan, Tajikistan, Kyrgyzstan, and Pakistan data set
GCM	Global climate model (see Table 3)
SSPnnn	SSP245 (SSP 2.45) or SSP585 (SSP 5.85). See Riahi et al 2017 for details.
resolution	MONTHLY or YEARLY
yyyy	Model run year
Vxx.x	Version number. E.g., V01.0 is Version 1.

Table 3 lists each CMIP6 GCM used as input, with a link to its description on the World Climate Research Programme web site.

Table 3. CMIP6 Global Climate Models

BCC-CSM2-MR	INM-CM4-8
EC-Earth3	MRI-ESM2-0
GFDL-ESM4	CESM2-WACCM
MPI-ESM1-2-HR	FGOALS-f3-L
CESM2	INM-CM5-0
EC-Earth3-Veg	NorESM2-MM

1.3 Spatial Information

1.3.1 Coverage

N: 39.83333° N

S: 31.08333° N

E: 81.75° E

W: 65.58333° E

The model domain (see Figure 1) was derived from the global MERIT PLUS river network (Version 1) in Prusevich et al., 2022, which was based on Eilander et al., 2021.

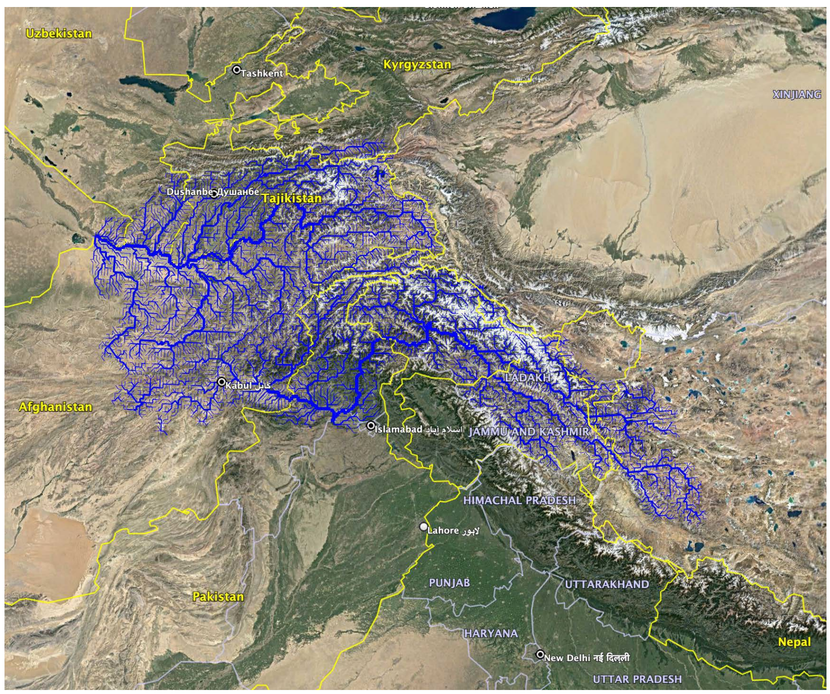

Figure 1. HMA Water Balance Projection Model Domain

1.3.2 Resolution

0.0833333° (5 arcmin)

1.3.3 Geolocation

The following tables provide information for geolocating this data set:

Table 4. Geolocation Details

Geographic coordinate system	World Geodetic System 1984
Projected coordinate system	N/A
Longitude of true origin	Prime Meridian, Greenwich
Latitude of true origin	N/A
Scale factor at longitude of true origin	N/A
Datum	WGS 84
Ellipsoid/spheroid	WGS 84
Units	degree
False easting	N/A
False northing	N/A
EPSG code	EPSG:4326
PROJ4 string	+proj=longlat +datum=WGS84 +no_defs
Reference	https://epsg.io/4326

Table 5. Grid Details

Grid cell size (x, y pixel dimensions)	0.0833333° × 0.0833333°
Number of rows	105
Number of columns	194
Geolocated lower left point in grid	31.0833333° N 65.5833333° E
Nominal gridded resolution	0.0833333°
Grid rotation	N/A
ulxmap – x-axis map coordinate of the center of the upper-left pixel (XLLCORNER for ASCII data)	65.625° E
ulymap – y-axis map coordinate of the center of the upper-left pixel (YLLCORNER for ASCII data)	39.7916667° N

1.4 Temporal Information

1.4.1 Coverage

01 January 2016 – 31 December 2099

01 January 1980 – 31 December 2018

1.4.2 Resolution

1 month

1 year

2 DATA ACQUISITION AND PROCESSING

2.1 Acquisition

The water balance projections (2016 – 2099) were generated using v.22.12.0 of the University of New Hampshire Water Balance Model, a process-based, modular, gridded, macro-scale hydrologic model that simulates land surface and groundwater hydrology. It accounts for human influences on the hydrological cycle due to dams and reservoirs (Wisser et al. 2010); inter-basin hydrological transfers (Lammers 2022 and Zaveri et al. 2016); and agricultural water use (Grogan et al. 2017 and Zaveri et al. 2016). The latest open source release, utilities, and complete documentation are available on GitHub.

The historical model run (1980 – 2018) was generated using temperatures from the European Centre for Medium-Range Weather Forecasts (ECMWF), fifth generation atmospheric reanalysis (ERA5) and ensemble precipitation estimates for High Mountain Asia developed by Maina et al., 2022.

2.2 Processing

Water balance projections for SSP 2-4.5 and SSP 5-8.5 were run using each CMIP6 GMC (see Table 3) as input. Model output was saved as monthly time steps and then aggregated to annual values.

The model runs included the recently developed 3D groundwater flow module for WBM based on principle formulations from the USGS MODFLOW software (Langevin et al., 2017) which allows a full run-time coupling of land and river flow hydrology with deep groundwater aquifers. The module can simulate both unconfined and confined aquifers using formulations similar to those in de Graaf et al. (2015, 2017).

These runs also utilized the WBM's human modules, which included: dams and reservoirs; inter-basin hydrological transfers (although none exist in this domain); impervious surfaces; cropland irrigation; livestock water use; and other water use and consumption from domestic and industrial sectors.

Lastly, the WBM outputs total water storage, which allows these data to be compared with large-scale groundwater changes derived from observations obtained by NASA's twin Gravity Recovery and Climate Experiment (GRACE) satellites.

2.3 Quality, Errors, and Limitations

The WBM uses simulated glacier data (Rounce et al., 2022) as an input. It does not make water balance calculations over any of the glaciers, but instead relies on the glacier model, PyGEM (Rounce et al., 2020), to provide water inputs to grid cells in which the glaciers are located. Within those grid cells, the WBM adjusts precipitation dynamically through time to account for how PyGEM handles glacier water balance and to compensate for changing glacier area.

Furthermore, in some high elevation grid cells, variations within the climate fields due to uncertainty in the GCMs, as well as downscaling to finer spatial and temporal resolutions, causes some grid cells to accumulate spurious snow throughout the simulation period.

3 VERSION HISTORY

Version 1 (initial release)

4 RELATED DATA SETS

- High Mountain Asia 5 Arc-Minute Hydrological Flow Direction for the Headwaters of the Amu Darya and Indus River Basins, Version 1
- Global PyGEM-OGGM Glacier Projections with RCP and SSP Scenarios, Version 1

5 REFERENCES

de Graaf, I. E. M., E. H. Sutanudjaja, L. P. H. van Beek, and M. F. P. Bierkens (2015). A high-resolution global-scale groundwater model, *Hydrology and Earth System Sciences*, *19*(2), 823-837. https://doi.org/10.5194/hess-19-823-2015

de Graaf, I. E. M., R. L. P. H. van Beek, T. Gleeson, N. Moosdorf, O. Schmitz, E. H. Sutanudjaja, and M. F. P. Bierkens (2017). A global-scale two-layer transient groundwater model: Development and application to groundwater depletion, *Advances in Water Resources*,*102*, 53-67. http://doi.org/10.1016/j.advwatres.2017.01.011

Eilander, D., van Verseveld, W., Yamazaki, D., Weerts, A., Winsemius, H. C., and Ward, P. J. (2021). A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models, *Hydrol. Earth Syst. Sci.*, 25(9), 5287-5313. https://doi.org/10.5194/hess-25-5287-2021

Grogan, D.S., D. Wisser, A. Prusevich, R.B. Lammers, and S. Frolking (2017). The use and re-use of unsustainable groundwater for irrigation: A global budget, *Environmental Research Letters*, 12(3), 034017. https://doi.org/10.1088/1748-9326/aa5fb2

Lammers, R. B (2022). Global Inter-Basin Hydrological Transfer Database (Version v22c) [Data set]. *MSD-LIVE Data Repository*. https://doi.org/10.57931/1905995

Maina, F. Z., S.V. Kumar, I.J. Dollan, and V. Maggioni (2022). Development and evaluation of ensemble consensus precipitation estimates over High Mountain Asia, *Journal of Hydrometeorology*, 23(9): 1469-1486. https://doi.org/10.1175/JHM-D-21-0196.1

Prusevich, A., R. Lammers, and S. Glidden (2022). MERIT-Plus Dataset: Delineation of endorheic basins in 5 and 15 min upscaled river networks (Version v1) [Data set]. MSD-LIVE Data Repository. https://doi.org/10.57931/1904379

Rounce, D., R. Hock, and F. Maussion (2022). Global PyGEM-OGGM Glacier Projections with RCP and SSP Scenarios, Version 1 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/P8BN9VO9N5C7. Date Accessed 03-20-2023.

Wisser D, BM Fekete, CJ Vörösmarty, and A.H. Schumann (2010). Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), *Hydrology and Earth System Science*, 14, 1-24. https://doi.org/10.5194/hess-14-1-2010

Zaveri, E., D.S. Grogan, K. Fisher-Vanden, S. Frolking, R.B. Lammers, D.H. Wrenn, A. Prusevich, and R.E. Nicholas (2016). Invisible water, visible impact: Groundwater use and Indian agriculture under climate change, *Environmental Research Letters*, 11, 084005. https://doi.org/10.1088/1748-9326/11/8/084005

6 DOCUMENT INFORMATION

6.1 Publication Date

August 2023

6.2 Date Last Updated

August 2023