Title

Global monthly sectoral water use for 2010-2100 at 0.5° resolution across alternative futures (Version 1).

Authors

Zarrar Khan[1], Isaac Thompson[1], Chris R. Vernon[1], Neal Graham[1], Tom B. Wild[1], Min Chen[2]

Affiliations

1. Joint Global Change Research Institute, Pacific Northwest National Laboratory, 5825 University Research Ct., Suite 3500, College Park, MD, 20740, USA

2. Department of Forest and Wildlife Ecology, College of Agriculture & Life Sciences, University of Wisconsin – Madison, Russell Labs, 1630 Linden Drive, Madison, WI 53706

*corresponding author(s): Zarrar Khan (Zarrar.khan@pnnl.gov)

Abstract

Future sectoral-specific water withdrawals at a temporal resolution capable of representing patterns in seasonality and a commonly used spatial resolution are an important factor to consider for energy, water, land and environmental research. Projected water withdrawals that are harmonized with assumptions for alternate futures that capture socioeconomic and climatic variation are critical for many modeling studies on future global and regional dynamics. Here we generate a novel global gridded water withdrawals dataset by coupling Global Change Analysis Model (GCAM) with a land use spatial downscaling model (Demeter), a global hydrologic framework (Xanthos) and a water withdrawal downscaling model (Tethys) for the five Shared Socioeconomic Pathways (SSPs) and four Representative Concentration Pathways (RCPs) scenarios. The dataset provides sectoral monthly data at 0.5° resolution for years 2010 to 2100. The presented dataset will be useful for both global and regional analysis looking at the impacts of socioeconomic, climate and technological futures as well as in characterizing the uncertainties associated with these impacts.

Background & Summary

This paper documents a global monthly gridded (0.5° resolution) sectoral water withdrawal and consumption dataset across a range of future socio-economic and climate scenarios from 2010 to 2100. The dataset was generated using Tethys[1] to spatially and temporally downscale outputs from the Global Change Analysis Model (GCAM)[2] coupled with land-use change projections from Demeter[3]. GCAM is an integrated assessment model and thus captures the impacts of both human and climate system changes in the future. Demeter is a high-resolution downscaling model that uses GCAM outputs to calculate global gridded monthly land-use change. With the combination of GCAM and Demeter, Tethys is able to project water withdrawal and consumption demands for 6 sectors (domestic, electricity generation, irrigation, livestock, industry and mining) with the irrigation sector further divided into 7 different crop types (biomass, corn, fiber crop, fodder grass, fodder herb, oil crops and miscellaneous crops). To capture a range of futures, 75 scenarios were used which comprised of a combination of 4 Representative Concentration Pathways (RCPs)[4], 5 Shared Socioeconomic Pathways (SSPs)[5] and 5 Global Climate Models (GCMs) as shown in Figure 1. 15 viable combinations of the SSPs and RCPs were combined with each of the 5 GCMs to arrive at the final 75 scenarios used. Details on the GCMs which are from the Inter-sectoral Impact Model Intercomparison Project (ISIMIP)[6] protocol 2b are provided in Graham et al. 2020[2] where the original data from the GCMs were used to run the GCAM simulations.

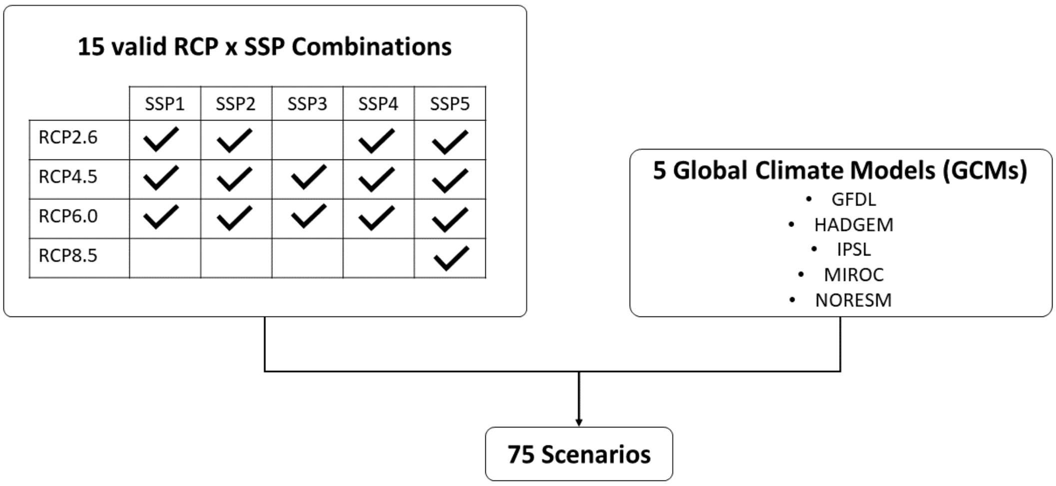

Figure 1. The 75 scenarios in the dataset are a combination of 4 Representative Concentration Pathways (RCPs), 5 Shared Socioeconomic Pathways (SSPs) and 5 Global Climate Models (GCMs). 15 viable combinations of SSPs and RCPs were combined with each of the 5 GCMs to arrive at the final 75 scenarios.

This dataset is important because it can help planners understand the sources of demand-side pressures on scarce water resources. Mekonnen & Hoekstra 2016[7] (also cited in the UN World Water Development Report 2022[8]) estimated that roughly 71% (4.1 billion people) of the world's population was exposed to water scarcity at least one month in the year over the period from 1996 to 2005. In their more recent study, Van Vliet et al. 2021[9] estimate global water scarcity over the period from 2000 to 2010 to range from 30% without water quality considered to 40% when also including water quality. Global water scarcity is expected to increase across the globe with critical development implications for sustainable development[2,10–13]. Recent studies highlight that future water scarcity is primarily driven by human water demands, more than by climate impacts on availability[2,14]. Additionally, irrigation water demands have been shown to have the largest impact on water scarcity[10,11,15]. Furthermore, water access, availability and demands are highly localized with large energy and economic costs associated with water transfers and thus a regional understanding of water use is essential[16,17]. This paper accounts for all these key factors by providing an a transparent and open-source dataset and accompanying methodology that captures the key drivers of future water scarcity (water use for human activities) at a fine spatio-temporal scale (0.5° resolution and monthly) and with added detail on irrigation water use by crop types.

Past studies that have looked at global water use at similar spatial (gridded) and temporal (monthly) scales have been limited to historical data, while future projections are provided for annual projections at aggregated country, basin or regional scales. Other studies with future projections use different scenarios and modelling techniques. Table 1 compares the key features in this study as compared to other similar studies that analyse global water use.

This study thus addresses the critical need for future projections of distributed water demand at a fine resolution so that local water managers can start to explore and plan for future water needs. The entire dataset can be easily downloaded from a dataverse online repository (https://doi.org/10.7910/DVN/VIQEAB) and is accompanied by a meta-repository (https://jgcri.github.io/khan-etal_2022_tethysSSPRCP/) which provides detailed figures and workflows for interested readers.

Table 1 Comparison of selected global water use studies

	Water Use Types	Sectors	Additional Sectors	Spatial Scope	Temporal Scope	Scenarios
Khan et al. 2022 (This study)	- Withdrawals - Consumption	- Mining - Domestic - Electricity - Livestock - Industry - Irrigation	(13 Crops) Biomass, Corn, Fiber Crop, Misc Crop, Oil Crop, Other Grain, Palm Fruit, Rice, Root Tuber, Sugar Crop, Wheat, Fodder Herb, and Fodder Grass	- Global - 0.5 deg gridded	- Historical: - 2010 - Monthly - Future: - 2015 to 2100 - Monthly	- Historical (2010) - Future: - SSPs 1 to 5 - RCP2.6, 4.5, 6.0, 8.5 - 5 CMIP5 GCMs (GFDL, HADGEM, IPSL, MIROC, NORESM)
Aqueduct (WRI) (2019, 2015)[18,19]	- Withdrawals - Consumption	- Domestic - Industry - Agriculture - Livestock	-	- Global - 0.083 deg (historical) - 0.5 deg (future)	- Historical: - 1990-2014 - Monthly - Future: - 2020, 2030, 2040 - Annual	- Historical: PCR-GLOBWB 2 Outputs - Future: - SSP2, SSP3 - RCP4.5, RCP8.5 - 6 CMIP5 GCMs (CCSM4, CNRM-CM5, GFDL-ESM2M, INMCM4, MPI-ESM-LR, MRI-CGCM3)
Huang et al. 2018[20]	- Withdrawals - Consumption	- Mining - Domestic - Electricity - Livestock - Industry - Irrigation	-	- Global - 0.5 deg gridded	- Historical: - 1971-2010 - Monthly	- Historical (4 GHMs: WaterGAP, H08, LPJml, PCR-GLOBWB)
Wada et al. 2014[21]	- Withdrawals - Consumption	- Domestic - Livestock - Industry - Irrigation	- Paddy - Non-paddy	- Global - 0.5 deg gridded	- 1979 - 2010 - Daily	- Historical (1979-2010)
Hanasaki et al. 2013[10]	- Withdrawals	- Municipal - Industry - Irrigation	-	- Global - 0.5 deg gridded	- 2000 to 2100 - Daily	- Historical (2000) - Future: - SSPs 1 -5 - RCP2.6, 4.5, 6.0, 8.5
Mekonnen & Hoekstra 2011[22]	- Consumption (blue water footprint)	- Total	- Additional datasets available for crops, industrial products and livestock[23–25]	- Global - 0.5 deg gridded	- Historical: - 1996 - 2005 - Monthly	- Historical outputs of water balance model

Methods

Tethys v1.3.1[26] was used to downscale the water withdrawals and consumption outputs from GCAM. GCAM produces water withdrawals and consumptions outputs for 32 regions for the domestic, mining, power generation, industry, and livestock sectors and for 434 region-basin intersections for the irrigation sector as shown in Figure 2. These outputs are downscaled onto a 0.5° by 0.5° grid as shown in Figure 3. Of the 259,200 possible grid cells at this resolution (360 x 720), only the 67,420 cells categorized as land are considered.

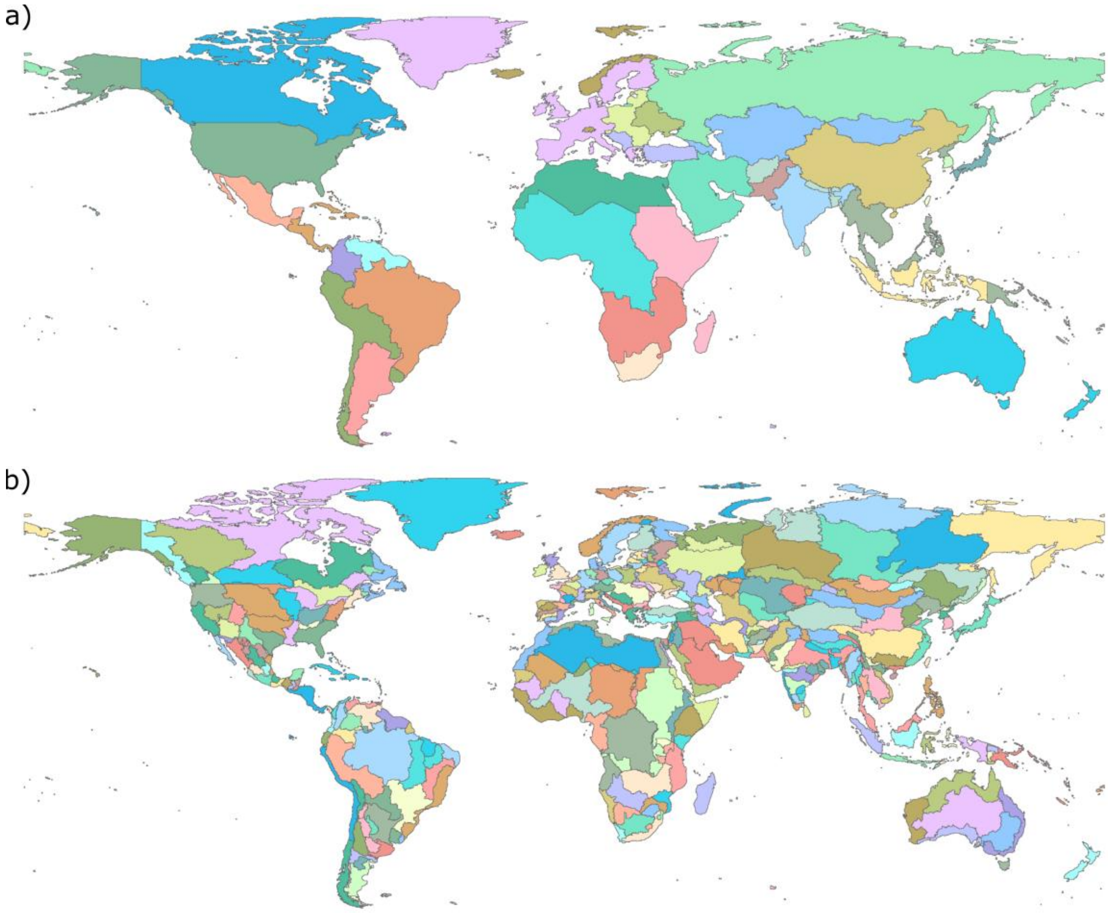

Figure 2. Water withdrawals and consumption from GCAM by a) 32 GCAM regions for domestic, mining, power generation, industry, and livestock sectors and b) 434 GCAM region and basin intersections for the irrigation sector.

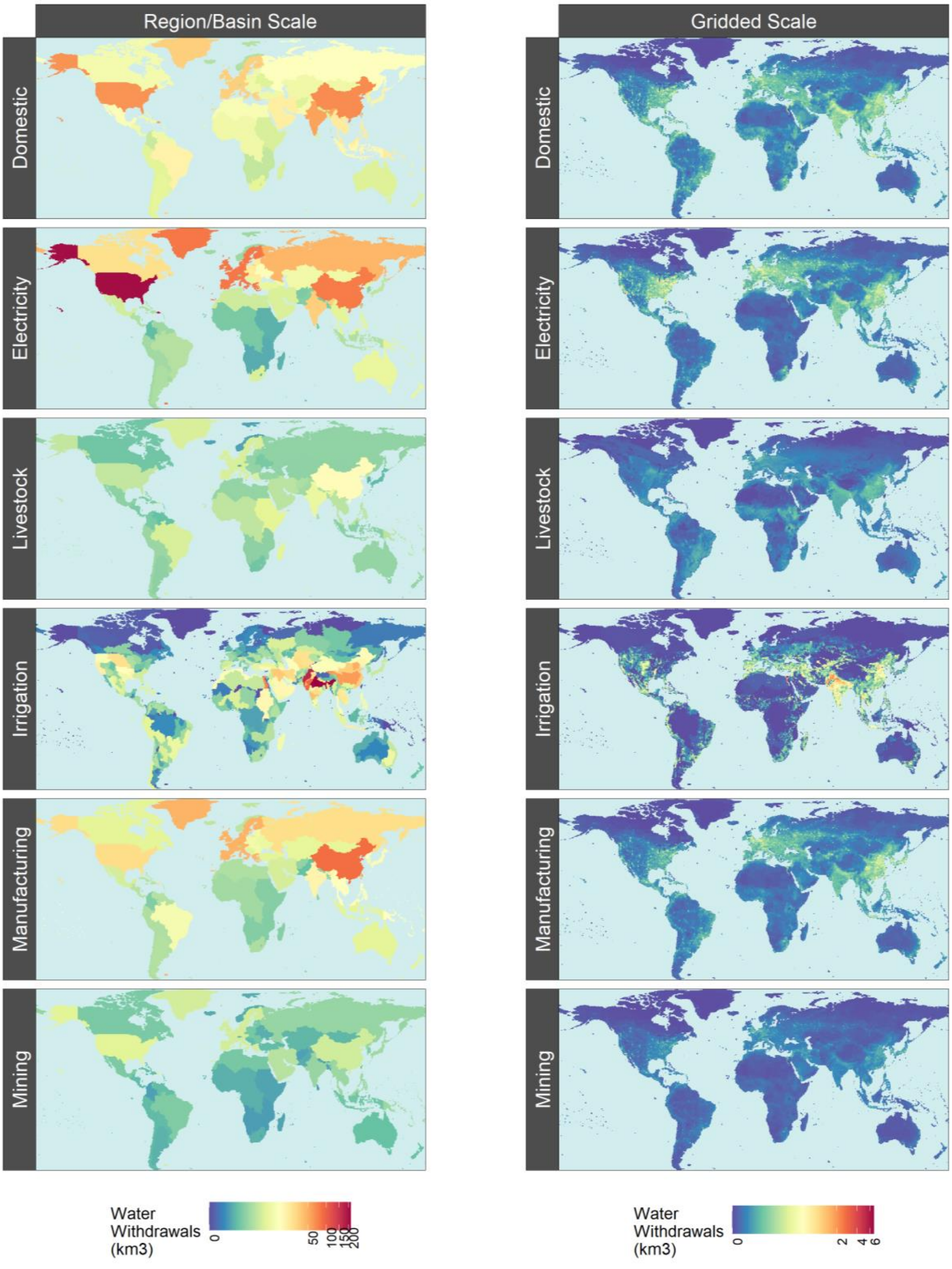

Figure 3. Example outputs of Tethys spatial downscaling of 2010 water withdrawals by sector from GCAM regions and basins to 0.5° x 0.5° grids.

Spatial Downscaling – Non-Agriculture: Spatial downscaling for non-agricultural (domestic, electricity, manufacturing, and mining), water withdrawals and consumption in each grid cell are assumed to be proportional to that cell's population as compared to the larger GCAM region within which that grid cell falls. The population data set used for this paper is from "Gridded Population of the World" (SEDAC, 2016)[27]. Tethys uses the nearest available year, which for this paper was 2010 in 2010, and 2015 in all other years. Each region's population is determined by taking the sum of population over all cells belonging to that region. For each of these sectors, Tethys calculates the water withdrawals and consumption as shown in Equation (1) and Equation (2) for a given cell by:

$$withdrawal_{cell} = withdrawal_{region} \times \frac{population_{cell}}{population_{region}} \tag{1}$$

$$consumption_{cell} = consumption_{region} \times \frac{population_{cell}}{population_{region}} \tag{2}$$

Spatial Downscaling – Livestock: Spatial downscaling of livestock water use, is calculated using gridded global maps from the FAO gridded livestock of the world (Wint and Robinson, 2007)[28] dataset for six types of livestock (cattle, buffalo, sheep, goats, pigs, and poultry). GCAM outputs are organized into five types (beef, dairy, pork, poultry, and "sheepgoat"), so these are first reorganized to match the six types from Wint and Robinson, 2007[28] using ratios for each region estimated from the dataset. The ratios are stored in two files that are used as inputs to Tethys: bfracFAO2005.csv ("buffalo fraction") and gfracFAO2005.csv ("goat fraction"). The following formulas are used to map the five GCAM livestock types to the six livestock types from Wint and Robinson, 2007[28] for each region:

$$buffalo = (beef + dairy) \times buffalo\_fraction \tag{3}$$

$$cattle = (beef + dairy) \times (1 - buffalo\_fraction) \tag{4}$$

$$goat = (sheepgoat) \times goat\_fraction \tag{5}$$

$$sheep = (sheepgoat) \times (1 - goat\_fraction) \tag{6}$$

No adjustment is required for pork (pigs) or poultry. After this, downscaling for each livestock type is very similar to downscaling the nonagricultural sectors, except the respective livestock population (heads) is used as the proxy instead of human population.

$$withdrawal_{animal,cell} = withdrawal_{animal,region} \times \frac{heads_{animal,cell}}{heads_{animal,region}} \tag{7}$$

$$consumption_{animal,cell} = consumption_{animal,region} \times \frac{heads_{animal,cell}}{heads_{animal,region}} \tag{8}$$

The results for each of the six types are then added together to get the total livestock withdrawal and consumption for each cell:

$$withdrawal_{livestock,cell} = \begin{pmatrix} withdrawal_{cattle,cell} + \\ withdrawal_{buffalo,cell} + \\ withdrawal_{sheep,cell} + \\ withdrawal_{goat,cell} + \\ withdrawal_{pigs,cell} + \\ withdrawal_{poultry,cell} \end{pmatrix} \tag{9}$$

$$consumption_{livestock,cell} = \begin{pmatrix} consumption_{cattle,cell} + \\ consumption_{buffalo,cell} + \\ consumption_{sheep,cell} + \\ consumption_{goat,cell} + \\ consumption_{pigs,cell} + \\ consumption_{poultry,cell} \end{pmatrix} \tag{10}$$

Spatial Downscaling – Irrigation: GCAM irrigation withdrawal and consumption outputs are organized by 13 crop types: Biomass, Corn, Fiber Crop, Misc Crop, Oil Crop, Other Grain, Palm Fruit, Rice, Root Tuber, Sugar Crop, Wheat, Fodder Herb, and Fodder Grass. Demeter[29] provides a spatial landcover breakdown by all crop types except biomass, which is downscaled uniformly within a region-basin intersection (with respect to land area).

$$withdrawal_{biomass,cell} = withdrawal_{biomass,region} \times \frac{area_{cell}}{area_{region}} \tag{11}$$

$$consumption_{biomass,cell} = consumption_{biomass,region} \times \frac{area_{cell}}{area_{region}} \tag{12}$$

When possible, the other 12 crops are downscaled in proportion to the crop land area maps from Demeter, which have been reaggregated to the target resolution of 0.5 degrees. There are certain exceptions:

- If the GCAM withdrawal or consumption value for a crop in some region-basin is nonzero, but Demeter does not show any cells with that crop type in that region-basin, it will be downscaled uniformly, the same as biomass.
- Additionally, it is possible for GCAM and Demeter to have different total crop irrigation areas for a region basin intersection, so applying the raw Demeter ratios to irrigation withdrawals or consumption (which are directly related to irrigation areas) could result in cell withdrawal values that imply larger irrigation area than total cell area. In order to avoid this situation, excess irrigation area in cells that are above capacity is assigned evenly among irrigated cells with capacity remaining if there are any, otherwise it is assigned evenly among the remaining cells in the region-basin. Should there still be excess after those cells have been filled, it would be dropped.

Using these adjusted irrigation area values for each crop, cell withdrawal values are given by:

$$withdrawal_{crop,cell} = withdrawal_{biomass,region} \times \frac{area_{crop,cell}}{area_{crop,region}} \tag{13}$$

$$consumption_{crop,cell} = consumption_{biomass,region} \times \frac{area_{crop,cell}}{area_{crop,region}} \tag{14}$$

The total irrigation sector value for a cell is the sum of that cell's values for all 13 crops.

Temporal Downscaling – Domestic:

Temporally downscaling domestic withdrawal uses the following formula from Wada et al., 2011[30].The R parameter described below is from Huang et al. 2018[20] and temperature data is from Weedon et al.2014[31] Withdrawals for each month of a year for each cell are given by the formula:

$$withdrawal_{month} = \frac{withdrawal_{year}}{12}\left[\left(\frac{temp_{month} - temp_{min}}{temp_{max} - temp_{max}}\right)R + 1\right] \tag{15}$$

Where:

$temp_{month}$	=	Average temperature for the month
$temp_{mean}$	=	Mean monthly temperature for the year
$temp_{max}$	=	Max monthly temperature for the year
$temp_{min}$	=	Min monthly temperature for the year
R	=	Parameter representing the relative difference of water withdrawals between the warmest and coolest months of the year

Temporal Downscaling – Electricity Generation:

Water withdrawal and consumption for electricity generation each month are assumed to be proportional to the amount of electricity generated, using the formula developed in Voisin et al., 2013[32]:

$$withdrawal_{month} = withdrawal_{year}\left[\rho_b\begin{pmatrix}\rho_h \dfrac{HDD_{month}}{HDD_{year}} + \\ \rho_c \dfrac{CDD_{month}}{CDD_{year}} + \\ \rho_u \dfrac{1}{12}\end{pmatrix} + \rho_{it}\frac{1}{12}\right] \tag{16}$$

$$consumption_{month} = consumption_{year}\left[\rho_b\begin{pmatrix}\rho_h \dfrac{HDD_{month}}{HDD_{year}} + \\ \rho_c \dfrac{CDD_{month}}{CDD_{year}} + \\ \rho_u \dfrac{1}{12}\end{pmatrix} + \rho_{it}\frac{1}{12}\right] \tag{17}$$

Where:

ρ_b = Proportion of electricity used for buildings
ρ_{it} = Proportion of electricity used for industry and transportation
$\rho_b + \rho_{it}$ = 1
ρ_h = Proportion of electricity used for buildings heating
ρ_c = Proportion of electricity used for buildings cooling
ρ_u = Proportion of electricity used for buildings other
$\rho_h + \rho_c + \rho_u$ = 1
HDD = Heating Degree Days
CDD = Cooling Degree Days

Heating degree days (HDD) and cooling degree days (CDD) are indicators for the amount of electricity used to heat and cool buildings, and are calculated from mean daily outdoor air temperature. HDD for a month is the sum of $(18^{o}\mathrm{C}-\mathrm{temperature}_{\mathrm{day}})$ across all days where temperature is less than 18 degrees Celsius. CDD is the sum of $(\mathrm{temperature}_{\mathrm{day}}-18^{o}\mathrm{C})$ across all days where temperature is greater than 18. Annual HDD and CDD are the sum of their respective monthly values.

Tethys uses HDD, CDD, and pp values for each cell from the nearest available year in the input files listed at the end of this subsection, which is 2010 for this data set.

The formula is modified for cells with low annual HDD or CDD as described in Huang et al., 2018[20], since these may not have heating or cooling services despite nonzero values of ρ_h or ρ_c.

When $HDD_{year} < 650$, the HDD term is removed and ρ_h is reallocated to the cooling proportion, giving:

$$withdrawal_{month} = withdrawal_{year}\left[\rho_b\begin{pmatrix}(\rho_h+\rho_c)\dfrac{CDD_{month}}{CDD_{year}}+\\ \rho_u\dfrac{1}{12}\end{pmatrix}+\rho_{it}\right. \tag{18}$$

$$consumption_{month} = consumption_{year}\left[\rho_b\begin{pmatrix}(\rho_h+\rho_c)\dfrac{CDD_{month}}{CDD_{year}}+\\ \rho_u\dfrac{1}{12}\end{pmatrix}+\right. \tag{19}$$

When $CDD_{year} < 450$, the CDD term is removed and ρ_c is reallocated to the cooling proportion, giving:

$$withdrawal_{month} = withdrawal_{year}\left[\rho_b\begin{pmatrix}(\rho_h+\rho_c)\dfrac{HDD_{month}}{HDD_{year}}+\\ \rho_u\dfrac{1}{12}\end{pmatrix}+\rho_{it}\right. \tag{20}$$

$$consumption_{month} = consumption_{year}\left[\rho_b\left(\begin{matrix}(\rho_h+\rho_c)\dfrac{HDD_{month}}{HDD_{year}}+\\ \rho_u\dfrac{1}{12}\end{matrix}\right)+\right. \tag{21}$$

When annual HDD and CDD are both below their respective thresholds, all sources of monthly variation vanish and the formula reduces to

$$withdrawal_{month} = \frac{withdrawal_{year}}{12} \tag{22}$$

$$consumption_{month} = \frac{consumption_{year}}{12} \tag{23}$$

Temporal Downscaling – Livestock, Manufacturing and Mining:

For livestock, manufacturing, and mining, a uniform distribution is applied. The withdrawal or consumption for the year is divided between months according to the number of days.

$$withdrawal_{month} = withdrawal_{year} \times \frac{days_{month}}{days_{year}} \tag{24}$$

$$consumption_{month} = consumption_{year} \times \frac{days_{month}}{days_{year}} \tag{25}$$

Temporal Downscaling – Irrigation:

Temporal downscaling for irrigation water withdrawal and consumption is based on weighted irrigation profiles for each of the 235 basins. Gridded monthly irrigation withdrawal values from the PCR-GLOBWB global hydrological (from Huang et al. 2018[20], original data from ISIMIP[33]) model are averaged across the years 1971-2010, then aggregated to the basin scale. The monthly irrigation withdrawal percentages for a basin are applied to all crops in each of its cells:

$$withdrawal_{month} = withdrawal_{year} \times percent_{basin,month} \tag{26}$$

$$consumption_{month} = consumption_{year} \times percent_{basin,month} \tag{27}$$

In the event that the model has no monthly data for a basin with nonzero irrigation, the profile of the nearest available basin is used.

Data Records

Data outputs from this experiment have been minted and are available in the repository indicated in Table 2. A meta-repository with detailed information on the workflows to produce the data is also available and shown in Table 2.

Table 2 Data records

Record	Details	Location
Output Dataset	Data outputs from experiment	https://doi.org/10.7910/DVN/VIQEAB
Supporting Meta-repository	Meta-repository with detailed workflows for experiment	https://jgcri.github.io/khan-etal_2022_tethysSSPRCP/index.html

The dataset contains separate files with names which start with a combination of the following SSP, RCP, GCM and water usage type:

- **SSP:** ssp1, ssp2, spp3, spp4, spp5
- **RCP:** rcp26, rcp45, rcp60, rcp85
- **GCM:** gfdl, hadgem, ipsl, miroc, noresm
- **Water use type:** consumption, withdrawals

Example 1: ssp1_rcp26_gfdl_consumption_XXX
Example 2: ssp1_rcp26_gfdl_withdrawal_XXX

The datasets files have been then divided into sub-sets to manage their size. The following list shows the file structure for one of the SSP, RCP, GCM combinations:

- ssp1_rcp26_gfdl_consumption_crops_annual.zip
- ssp1_rcp26_gfdl_consumption_crops_monthly_1.zip
- ssp1_rcp26_gfdl_consumption_crops_monthly_2.zip
- ssp1_rcp26_gfdl_consumption_sectors_annual.zip
- ssp1_rcp26_gfdl_consumption_sectors_monthly_1.zip
- ssp1_rcp26_gfdl_consumption_sectors_monthly_2.zip

The files with "_crops_" in their names include data for individual crops while the files with "_sectors_" in their name include data for other aggregated sectors. The following expanded list shows the individual files inside the zipped files for the example ssp1_rcp26_gfdl cases. ***"cd" stands for "consumption downscaled" and "tcd" stands for "temporal consumption downscaled"***:

- ssp1_rcp26_gfdl_consumption_crops_annual.zip
    - crops_cdirr_biomass_km3peryr.csv
    - crops_cdirr_Corn_km3peryr.csv
    - crops_cdirr_FiberCrop_km3peryr.csv
    - crops_cdirr_FodderGrass_km3peryr.csv
    - crops_cdirr_FodderHerb_km3peryr.csv
    - crops_cdirr_MiscCrop_km3peryr.csv
    - crops_cdirr_OilCrop_km3peryr.csv
    - crops_cdirr_OtherGrain_km3peryr.csv
    - crops_cdirr_PalmFruit_km3peryr.csv
    - crops_cdirr_Rice_km3peryr.csv
    - crops_cdirr_Root_Tuber_km3peryr.csv
    - crops_cdirr_SugarCrop_km3peryr.csv
    - crops_cdirr_Wheat_km3peryr.csv
- ssp1_rcp26_gfdl_consumption_crops_monthly_1.zip
    - crops_tcdirr_biomass_km3peryr.csv

    - crops_tcdirr_Corn_km3peryr.csv
    - crops_tcdirr_FiberCrop_km3peryr.csv
    - crops_tcdirr_FodderGrass_km3peryr.csv
    - crops_tcdirr_FodderHerb_km3peryr.csv
    - crops_tcdirr_MiscCrop_km3peryr.csv
    - crops_tcdirr_OilCrop_km3peryr.csv
- ssp1_rcp26_gfdl_consumption_crops_monthly_2.zip
    - crops_tcdirr_OtherGrain_km3peryr.csv
    - crops_tcdirr_PalmFruit_km3peryr.csv
    - crops_tcdirr_Rice_km3peryr.csv
    - crops_tcdirr_Root_Tuber_km3peryr.csv
    - crops_tcdirr_SugarCrop_km3peryr.csv
    - crops_tcdirr_Wheat_km3peryr.csv
- ssp1_rcp26_gfdl_consumption_sectors_annual.zip
    - cddom_km3peryr.csv*(Domestic)*
    - cdelec_km3peryr.csv*(Electricity Generation)*
    - cdirr_km3peryr.csv*(Irrigation)*
    - cdliv_km3peryr.csv*(Livestock)*
    - cdmfg_km3peryr.csv*(Industry & manufacturing)*
    - cdmin_km3peryr.csv*(Mining)*
    - cdnonag_km3peryr.csv*(Aggregated non-agriculture)*
    - cdtotal_km3peryr.csv*(Total)*
- ssp1_rcp26_gfdl_consumption_sectors_monthly_1.zip
    - tcddom_km3peryr.csv*(Domestic)*
    - tcdelec_km3peryr.csv*(Electricity Generation)*
    - tcdirr_km3peryr.csv*(Irrigation)*
- ssp1_rcp26_gfdl_consumption_sectors_monthly_2.zip
    - tcdliv_km3peryr.csv*(Livestock)*
    - tcdmfg_km3peryr.csv*(Industry & manufacturing)*
    - tcdmin_km3peryr.csv*(Mining)*

Technical Validation

Results of the model were validated by re-aggregating spatial and temporal downscaled model outputs and comparing to the original aggregated inputs. Figure 4 shows how the disaggregated water withdrawal values in km^3 equal back to the original values both spatially for GCAM regions and temporally for annual values across sectors and crops. Figure 5 shows the same validation for how the disaggregated water consumption values in km^3 equal back to the original values both spatially for GCAM regions and temporally for annual values across sectors and crops.

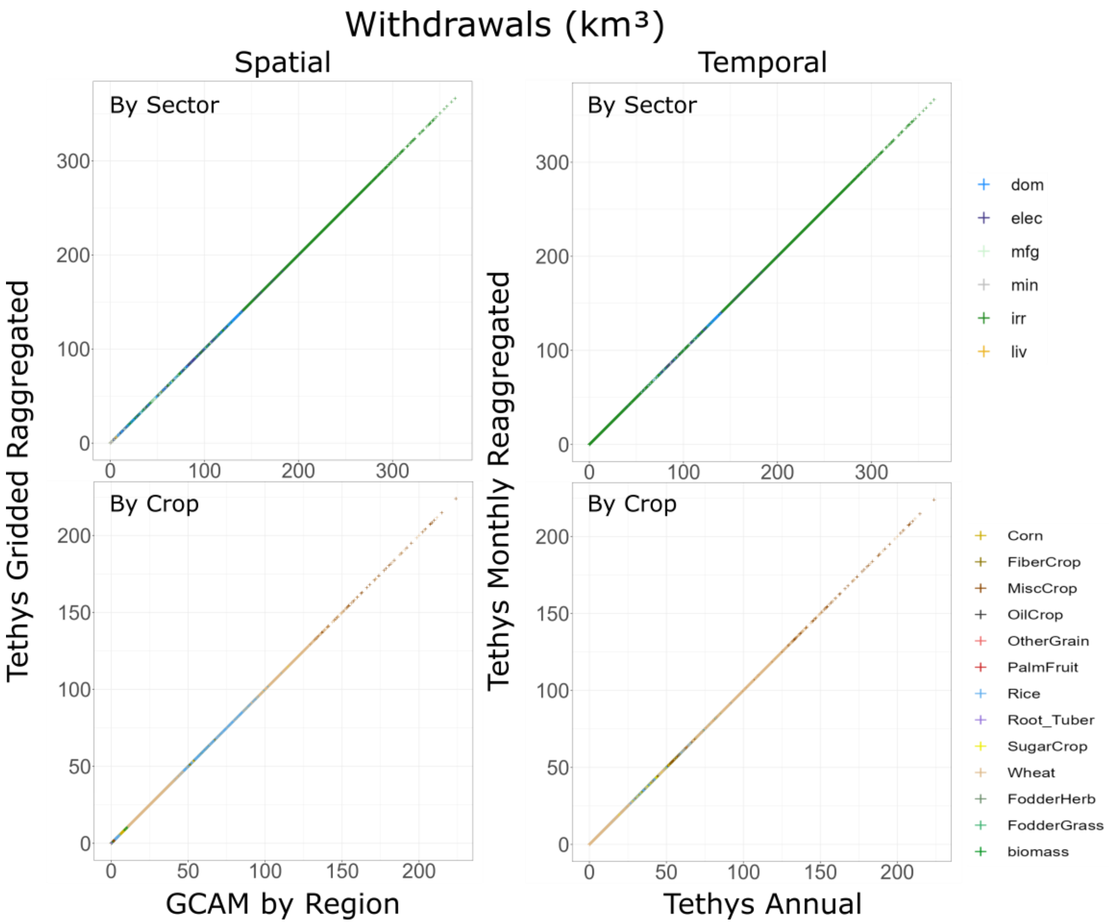

Figure 4 Validation of downscaled spatial and temporal Tethys water withdrawals (km³)

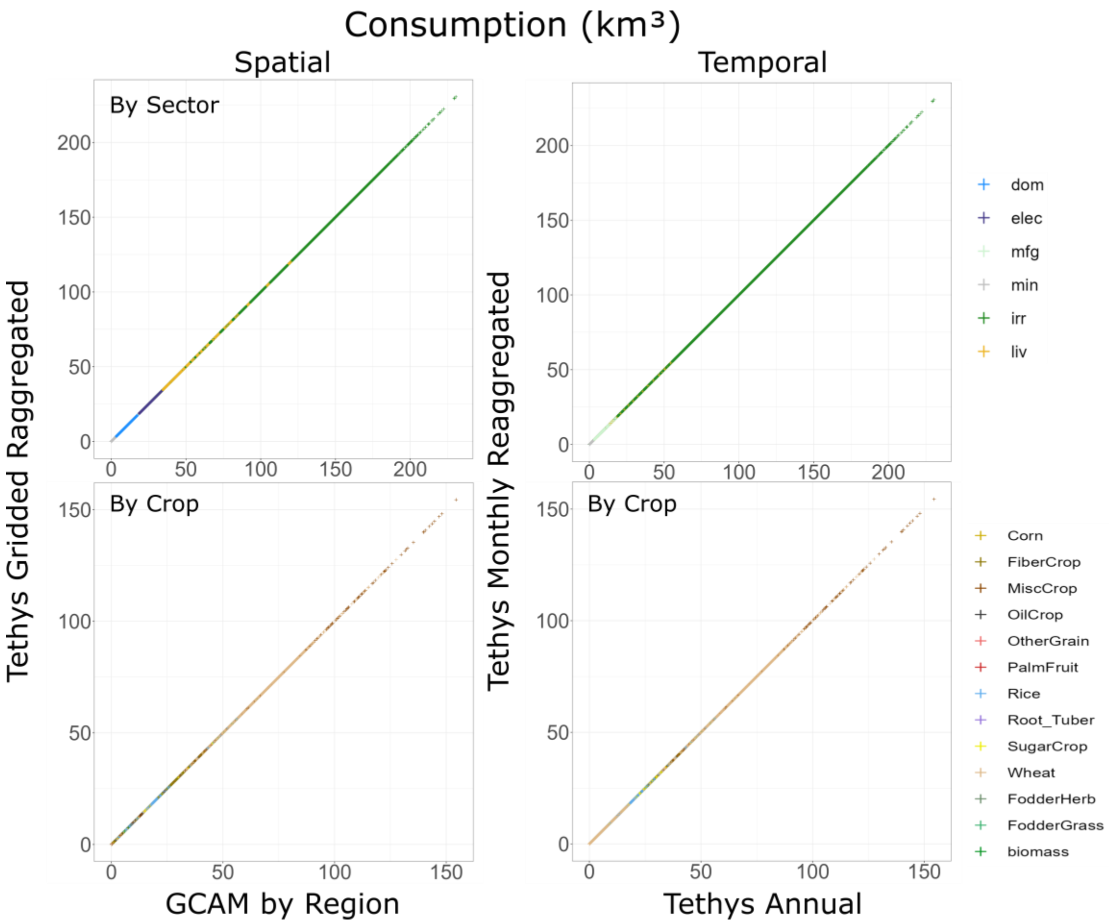

Figure 5 Validation of downscaled spatial and temporal Tethys water consumption (km^3)

Additionally, Tethys outputs were also compared to results from other two other studies Huang et al. 2018[20] and Mekonnen, M.M. and Hoekstra, A.Y. 2011[22]. Given the larger number of variables and assumptions for future scenarios we limit the validation with other studies to historical data. Since this work is primarily concerned with the downscaling of existing projections to a gridded monthly scale, we look at how spatial and temporal patterns in the year 2010 (for which all scenarios are identical) compare to those of the chosen datasets.

Huang et al. 2018[20], uses an earlier version of Tethys on historical data from 1971-2010. The underlying data have more regions and different totals, but many of the downscaling methods are identical, leading to similar results. For the non-agricultural sectors (domestic, electricity, manufacturing, and mining), the same underlying population map is used to downscale water use. This means that within each intersection of the different region schemes, the distribution of demand is identical, i.e., the ratio of their values to ours is constant for each grid cell in that intersection. For irrigation Huang et al. 2018[20] use USGS and FAO AQUASTAT irrigation data, whereas the current version of Tethys uses crop landcover maps from Demeter. Consumption and withdrawals generally showed similar spatial patterns, with differences in assumptions regarding each region's and sector's consumption to withdrawal ratios accounting for some differences. There are also some differences in accounting. For example, in this study hydropower is included in the consumption for electricity generation category, which by itself is several times greater than the entire water consumption for electricity generation in Huang et al. 2018[20].

The second data set we compared with is from Mekonnen, M.M. and Hoekstra, A.Y. 2011[22]. It contains monthly total blue water consumption values representing an average of years 1996-2005 which we compare to the base year values from 2010 from this study. The sectoral breakdown is different between the two datasets but at the same spatial resolution, so we compare monthly totals for each grid cell. We see some agreement between the two data sets. As the largest sector, differences in irrigation downscaling are likely responsible for the variation.

As seen in Figure 6 we see a general agreement in the sub-regional patterns across the data sets. Figure 7 also shows similar sub-annual patterns across the dataset with some differences in total values being attributed to underlying data and year of the study. Additional details on differences between the datasets are discussed in the meta-repository https://jgcri.github.io/khan-etal_2022_tethysSSPRCP/index.html.

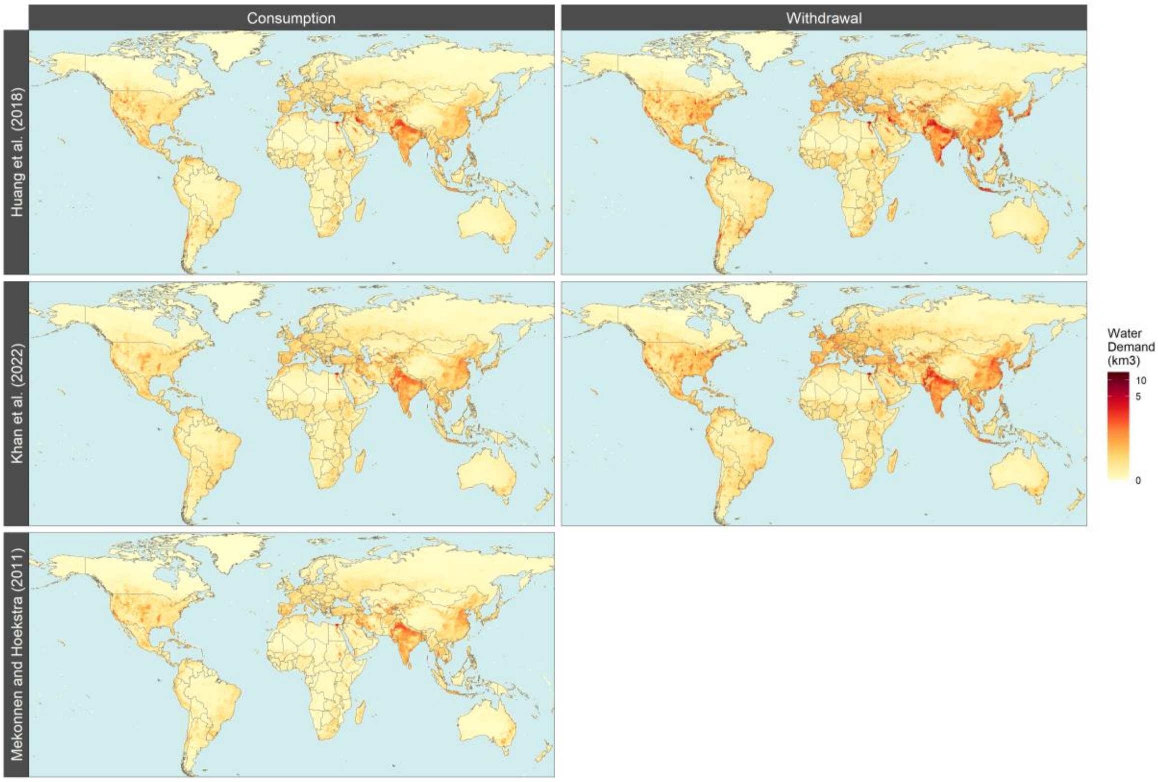

Figure 6 Spatial distribution of water withdrawals and consumption across this study, Huang et al. 2018[20] and Mekonnen, M.M. and Hoekstra, A.Y. 2011[22]

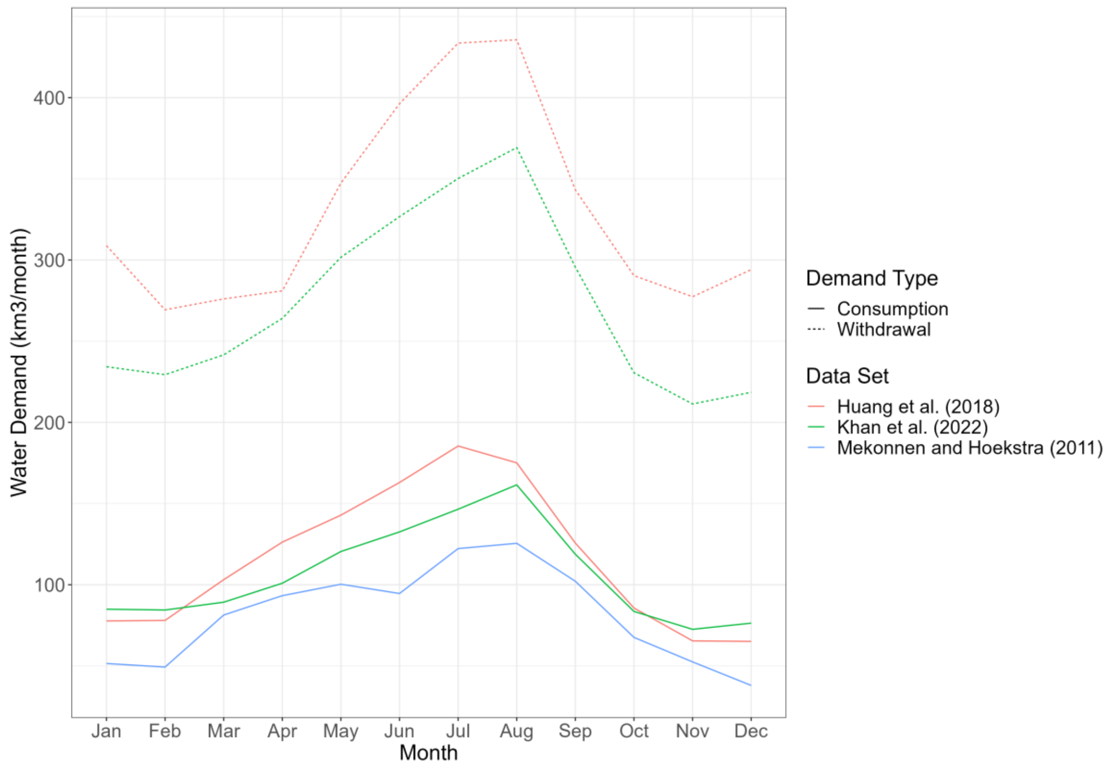

Figure 7 Temporal distribution of water withdrawals and consumption across this study, Huang et al. 2018[20] *and Mekonnen, M.M. and Hoekstra, A.Y. 2011*[22]

Usage Notes

Users are encouraged to explore the accompanying meta-repository (https://jgcri.github.io/khan-etal_2022_tethysSSPRCP/index.html) which provides detailed visualization across the various scenarios, sectors and time periods. Users can then download specific datasets for water withdrawal or consumption for relevant sectors, crops and desired SSP, RCP or GCM from the accompanying dataset repository (https://doi.org/10.7910/DVN/VIQEAB) to analyze the raw data. Some example figures from the meta-repository are presented in this section.

Figure 8 shows the total water withdrawals by sector for each SSP-RCP-GCM combination. Similar figures are available for consumption as well as by crop. Figure 9 shows the sub-annual temporal distribution across the same set of scenarios for 2010 and for 2100. Patterns such as an increase in summer water withdrawals can be seen in such figures.

The meta-repository also includes details on three selected basins: the Indus, Nile and Colorado. These are used to show how the data can be used to explore trends and patterns at this finer resolution. Figure 10 is an example showing how land-use change impacts which type of crop becomes the dominant water user in the Indus basin over time for the SSP1-RCP2.6-GFDL scenario. Figure 11 shows the accompanying distribution of total water withdrawals both spatially and temporally. Similar figures are provided in the meta-repository for water consumption as well as for other sectors, crops and scenarios.

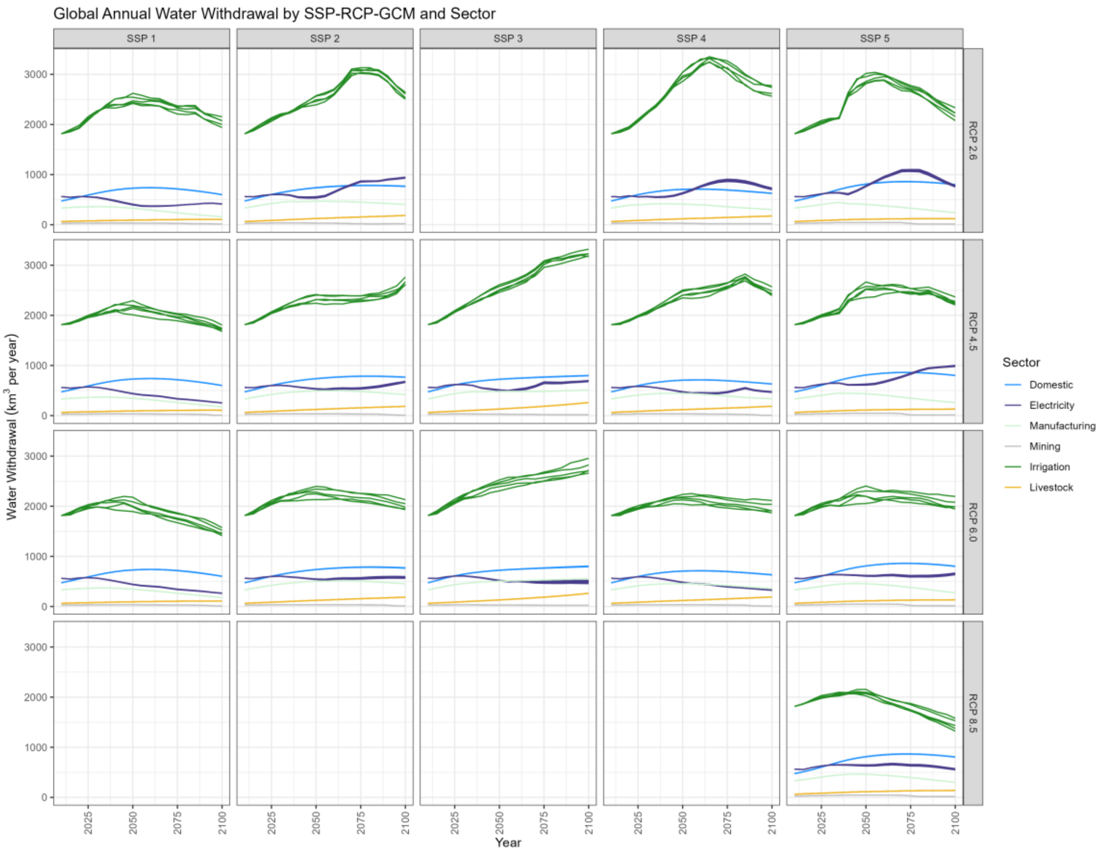

Figure 8 Global annual water withdrawals for SSP-RCP-GCM combinations by sector

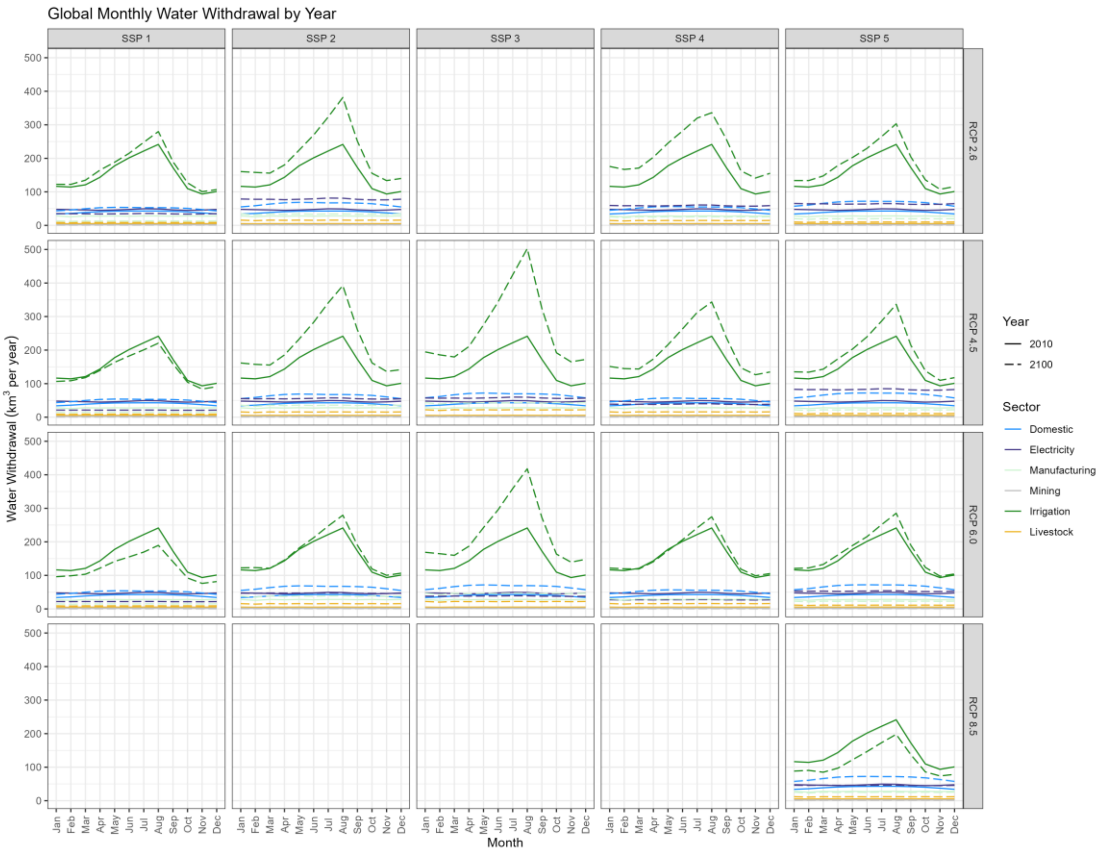

Figure 9 Global monthly water withdrawals for SSP-RCP-GCM combinations by sector for 2010 and 2100

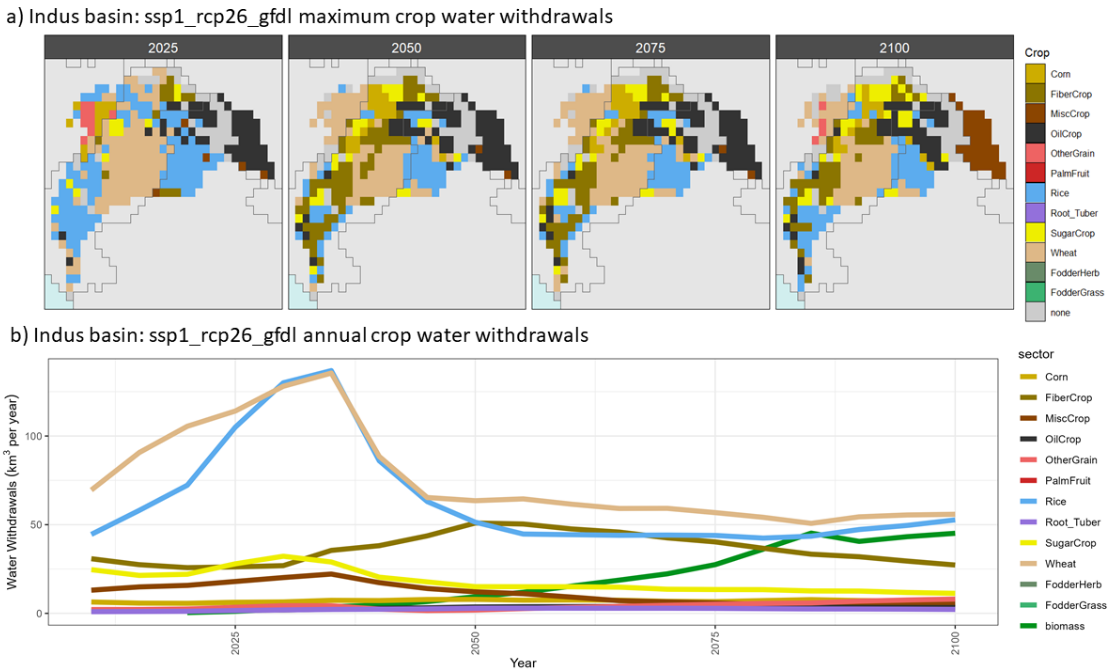

Figure 10 Indus Basin water withdrawals (km^3) by crop for scenario SSP 1, RCP 2.6, GCM GFDL. a) Showing which crop has the maximum water withdrawals (km^3) in each grid cell for years 2025, 2050, 2075 and 2100. b) Aggregated water withdrawals (km^3) by crop in the Indus Basin from 2015 to 2100.

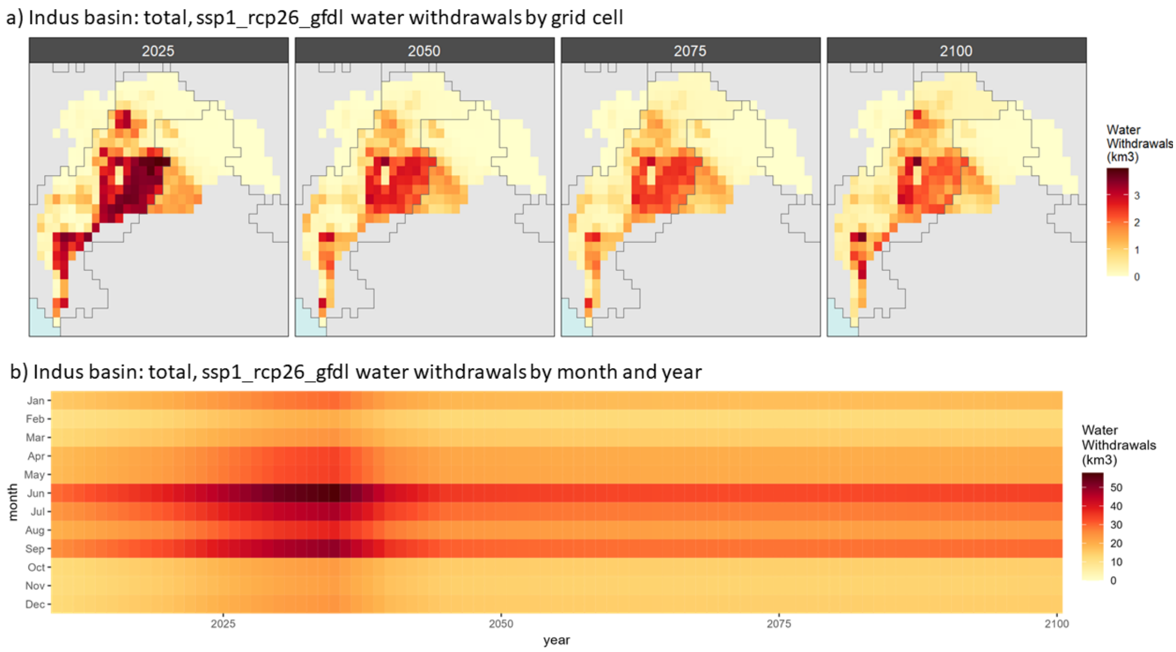

Figure 11 Indus Basin total water withdrawals (km^3) for scenario SSP 1, RCP 2.6, GCM GFDL. a) Showing total water withdrawals (km^3) in each grid cell for years 2025, 2050, 2075 and 2100. b) Aggregated total water withdrawals (km^3) in the Indus Basin from 2015 to 2100.

We highlight that several developments have been planned in the next release of Tethys to improve the methodologies used to downscale water use for the dataset in this paper. Some of the key planned developments include:

1. Improving the spatial distribution of powerplant water use based on actual and projected powerplant location instead of based on population.
2. Updating the output resolution to 1/8th degrees from the existing ½ degree resolution.

3. Including future population projections to improve on the current methodology which uses a static base year population map even for future years.
4. Improving the downscaling of biomass water use which is currently distributed equally within each region.
5. Making Tethys compatible with GCAM-USA which allow use of more accurate state-level water use data instead of using national data as inputs to Tethys.

Code Availability

The following table provides links to all models, data, versions and doi's used to generate this dataset.

Type	Details	Model Version	Data DOI	Model DOI
Tethys	Used to generate the data presented in this paper	v1.3.0	https://doi.org/10.7910/DVN/VIQEAB	https://doi.org/10.5281/zenodo.6399488
GCAM	Water use data used as inputs for Tethys	v4.3.chen	https://data.pnnl.gov/dataset/13224	http://doi.org/10.5281/zenodo.3713432
Demeter	Landuse change data used as input for Tethys	v1.chen	https://data.pnnl.gov/dataset/13192	http://doi.org/10.5281/zenodo.3713378

Acknowledgements

This research was supported by the U.S. Department of Energy, Office of Science, as part of research in MultiSector Dynamics, Earth and Environmental System Modeling Program. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. The views and opinions expressed in this paper are those of the authors alone.

Author contributions

Z.K., I.T., P.P., C.R.V., N.G., T.W., and M.C., designed the research.
Z.K. and I.T. ran Tethys to produce the outputs, prepared the figures and the data repository.
N.G. produced the GCAM data used as inputs for Tethys.
M.C. produced the Demeter data used as inputs for Tethys.
Z.K., I.T., P.P., C.R.V., N.G., T.W. all contributed to writing and reviewing the paper.

Competing interests

The authors declare no competing interests.

References

1. Li, X. *et al.* Tethys – A Python Package for Spatial and Temporal Downscaling of Global Water Withdrawals. *Journal of Open Research Software* **6**, (2018).
2. Graham, N. T. *et al.* Humans drive future water scarcity changes across all Shared Socioeconomic Pathways. *Environ. Res. Lett.* **15**, 014007 (2020).
3. Chen, M. *et al.* Global land use for 2015–2100 at 0.05° resolution under diverse socioeconomic and climate scenarios. *Sci Data* **7**, 320 (2020).
4. van Vuuren, D. P. *et al.* The representative concentration pathways: an overview. *Climatic Change* **109**, 5 (2011).
5. O'Neill, B. C. *et al.* The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change* **42**, 169–180 (2017).
6. ISIMIP. Inter Sectoral Impact Model Intercomparison (ISIMIP) - Input Data and Bias Correction. (2019).
7. Mekonnen, M. M. & Hoekstra, A. Y. Four billion people facing severe water scarcity. *Science Advances* **2**, e1500323 (2016).
8. UNESCO. *The United Nations World Water Development Report 2022: Groundwater: Making the invisible visible*. https://unesdoc.unesco.org/ark:/48223/pf0000380721 (2022).
9. Vliet, M. T. H. van *et al.* Global water scarcity including surface water quality and expansions of clean water technologies. *Environ. Res. Lett.* **16**, 024020 (2021).
10. Hanasaki, N. *et al.* A global water scarcity assessment under Shared Socio-economic Pathways – Part 1: Water use. *Hydrol. Earth Syst. Sci.* **17**, 2375–2391 (2013).
11. Hanasaki, N. *et al.* A global water scarcity assessment under Shared Socio-economic Pathways – Part 2: Water availability and scarcity. *Hydrol. Earth Syst. Sci.* **17**, 2393–2413 (2013).

12. Hejazi, M. I. *et al.* Integrated assessment of global water scarcity over the 21st century under multiple climate change mitigation policies. *Hydrology and Earth System Sciences* **18**, 2859–2883 (2014).

13. Wada, Y. & Bierkens, M. F. P. Sustainability of global water use: past reconstruction and future projections. *Environ. Res. Lett.* **9**, 104003 (2014).

14. Wada, Y., Beek, L. P. H. van, Wanders, N. & Bierkens, M. F. P. Human water consumption intensifies hydrological drought worldwide. *Environ. Res. Lett.* **8**, 034036 (2013).

15. Yoshikawa, S. *et al. An assessment of global net irrigation water requirements from various water supply sources to sustain irrigation: rivers and reservoirs (1960–2000 and 2050)*. https://hess.copernicus.org/preprints/10/1251/2013/hessd-10-1251-2013.pdf (2013) doi:10.5194/hessd-10-1251-2013.

16. Veldkamp, T. I. E. Water scarcity at the global and regional scales: unravelling its dominant drivers in historical and future time periods. (2017).

17. Wada, Y., de Graaf, I. E. M. & van Beek, L. P. H. High-resolution modeling of human and climate impacts on global water resources. *Journal of Advances in Modeling Earth Systems* **8**, 735–763 (2016).

18. Hofste, R. W. *et al.* Aqueduct 3.0: Updated decision-relevant global water risk indicators. *World Resources Institute: Washington, DC, USA* (2019).

19. WRI. WRI Aqueduct. (2021).

20. Huang, Z. *et al.* Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns. *Hydrology and Earth System Sciences* **22**, 2117–2133 (2018).

21. Wada, Y., Wisser, D. & Bierkens, M. F. P. Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics* **5**, 15–40 (2014).

22. Mekonnen, M. M. & Hoekstra, A. Y. *Total monthly blue water footprints of production at a 30 × 30 arc minute grid resolution (1996-2005)*. https://waterfootprint.org/en/resources/waterstat/monthly-gridded-blue-water-footprint-statistics/ (2011).

23. Mekonnen, M. & Hoekstra, A. National water footprint accounts: The green, blue and grey water footprint of production and consumption. Volume 1: Main Report. *Daugherty Water for Food Global Institute: Faculty Publications* (2011).

24. Mekonnen, M. M. & Hoekstra, A. Y. The green, blue and grey water footprint of crops and derived crops products. (2010).

25. Mekonnen, M. M. & Hoekstra, A. Y. A Global Assessment of the Water Footprint of Farm Animal Products. *Ecosystems* **15**, 401–415 (2012).

26. Khan, Z., Thompson, I. & Vernon, C. R. *Tethys v1.3.1*. (2022).

27. SEDAC. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. *Center for International Earth Science Information Network - CIESIN - Columbia University. 2018. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC)* https://doi.org/10.7927/H49C6VHW (2018).

28. Wint, W. & Robinson, T. *Gridded livestock of the world 2007*. (FAO, Roma (Italia), 2007).

29. Vernon, C. R. *et al.* Demeter – A Land Use and Land Cover Change Disaggregation Model. *Journal of Open Research Software* **6**, 15 (2018).

30. Wada, Y. *et al.* Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research* **47**, (2011).

31. Weedon, G. P. *et al.* The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research* **50**, 7505–7514 (2014).

32. Voisin, N. *et al.* One-way coupling of an integrated assessment model and a water resources model: evaluation and implications of future changes over the US Midwest. *Hydrology and Earth System Sciences* **17**, 4555–4575 (2013).

33. Warszawski, L. *et al.* The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. *PNAS* **111**, 3228–3232 (2014).