

---

SOFTWARE METAPAPER

# Tethys – A Python Package for Spatial and Temporal Downscaling of Global Water Withdrawals

Xinya Li<sup>1</sup>, Chris R. Vernon<sup>1</sup>, Mohamad I. Hejazi<sup>2</sup>, Robert P. Link<sup>2</sup>, Zhongwei Huang<sup>2</sup>, Lu Liu<sup>2,3</sup> and Leyang Feng<sup>2</sup>

<sup>1</sup> Pacific Northwest National Laboratory, Richland, WA 99352, US

<sup>2</sup> Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, Maryland, US

<sup>3</sup> University of Maryland, College Park, Maryland, US

Corresponding author: Xinya Li (Xinya.Li@pnnl.gov)

---

Downscaling of water withdrawals from regional/national to local scale is a fundamental step and also a common problem when integrating large scale economic and integrated assessment models with high-resolution detailed sectoral models. Tethys, an open-access software written in Python, is developed with statistical downscaling algorithms, to spatially and temporally downscale water withdrawal data to a finer scale. The spatial resolution will be downscaled from region/basin scale to grid (0.5 geographic degree) scale and the temporal resolution will be downscaled from year to month. Tethys is used to produce monthly global gridded water withdrawal products based on estimates from the Global Change Assessment Model (GCAM).

---

**Keywords:** Global Change Assessment Model (GCAM); Python; Water Disaggregation; Global Water Withdrawals; Spatial Downscaling; Temporal Downscaling

**Funding statement:** PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

---

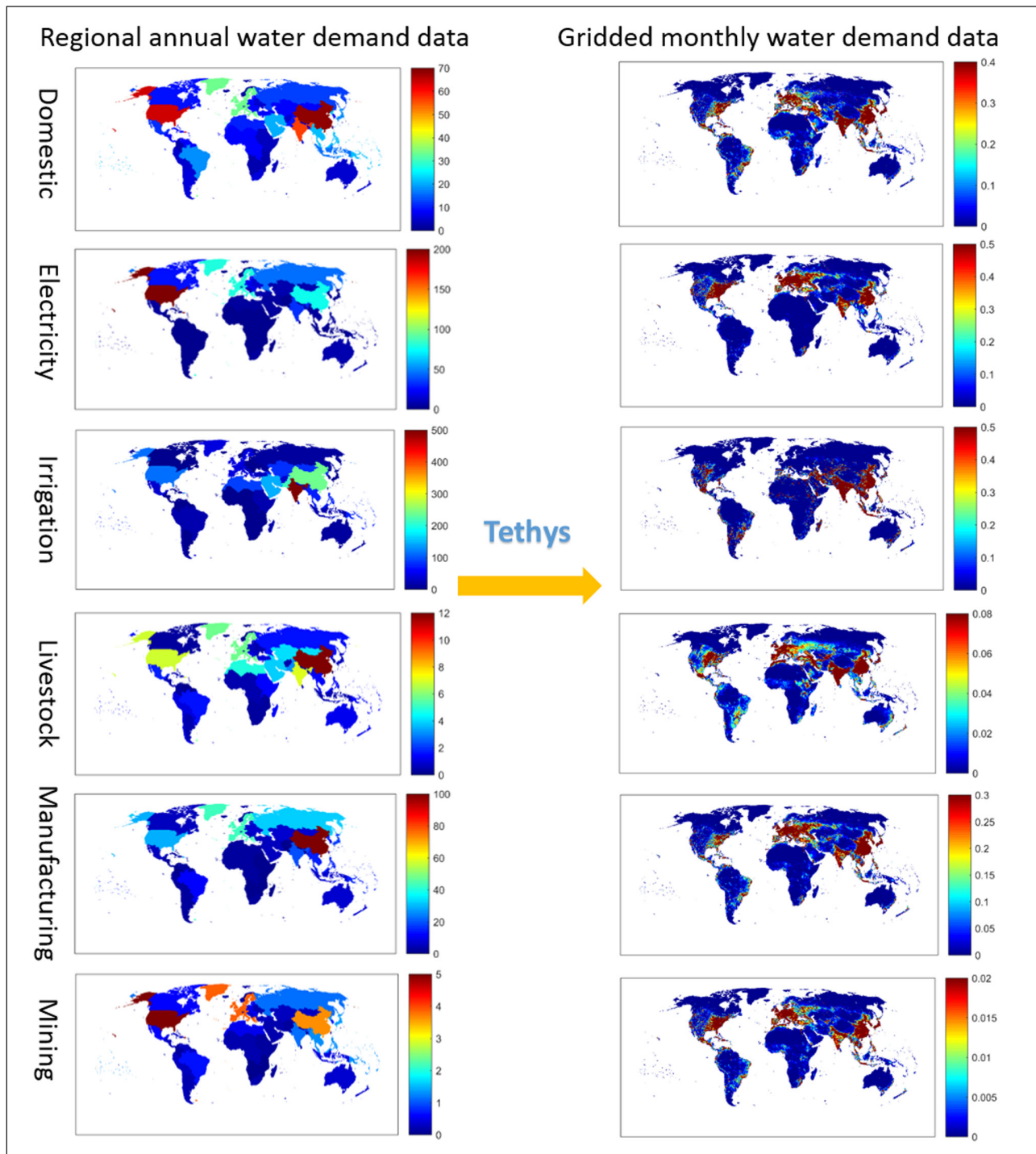
## (1) Overview

### Introduction

Tethys was constructed at the Joint Global Change Research Institute of the Pacific Northwest National Laboratory (<http://www.globalchange.umd.edu>). It served as a critical step to link Xanthos (a global hydrologic model) [1] and the Global Change Assessment Model (GCAM) [2–3]. The spatial resolution of GCAM is geopolitical regional scale for energy and economy systems (e.g. 32 regions), and river basins for the land, agriculture, and water systems (e.g. 235 water basins [21]). GCAM is often used as a boundary condition and coupled to sectoral models, such as the Community Land Model and Xanthos, which typically operate at finer spatial and temporal scales than GCAM [19, 20]. For example, Xanthos is a globally gridded hydrology model that operates at monthly scale. Resolving such a mismatch in spatial and temporal scales facilitated coupling these models together. It is also helpful for understanding seasonal patterns of water use and acquiring high resolution water use data [5]. The main objective of Tethys is to reconstruct global monthly gridded (0.5 geographic degree) water withdrawal datasets by spatial and temporal downscaling water withdrawal estimates at region/basin and annual scale

(**Figure 1**). As an open-access software, Tethys applies statistical downscaling algorithms, to spatially and temporally downscale water withdrawal data from annual region/basin scale into monthly grid scale. In our study, the water withdrawals are separated into six sectors: irrigation, livestock, domestic, electricity (generation), manufacturing and mining.

The algorithms for spatial downscaling were derived from research by Edmonds and Reilly [2, 4]. Non-agriculture (domestic, electricity, manufacturing and mining) sectors are downscaled based on global gridded population density maps [6]. Irrigation water withdrawal is downscaled using global coverage of gridded cropland areas equipped with irrigation [7, 8]. The gridded population maps (combined Historical Database of the Global Environment (HYDE) [9] and Gridded Population of the World (GPW) [10] data products) and gridded crop irrigation area maps (combined HYDE [9] and Food and Agriculture Organization (FAO) [11] data products) are updated in the algorithms over time by using historical datasets (the most recent available historical map is applied for future years). The gridded global maps of livestock in six types (cattle, buffalo, sheep, goats, pigs and poultry) [12] are used as proxy to downscale livestock water withdrawal [6, 13, 14].



**Figure 1:** Major inputs and outputs of Tethys by six sectors.

Different temporal downscaling algorithms were applied to different water withdrawal sectors [5]:

1. Irrigation: The monthly gridded irrigation water withdrawal was estimated by relying on monthly irrigation results from several global hydrological models (e.g. H08 [15–16], LPJmL [17], and PCR-GLOBWB [6, 18]) to quantify monthly weighting profiles of how irrigation is spread out within a year in a particular region and per crop type.
2. Domestic: Temporal downscaling of domestic water withdrawal from annual to monthly was based on

a formula from [6] and [19] and utilizing monthly temperature data; details of data sources were listed in [5].

3. Electricity: Temporal downscaling of electricity water withdrawal from annual to monthly was based on the assumption that the amount of water withdrawal for electricity generation is proportional to the amount of electricity generated [19, 20].
4. Livestock, manufacturing and mining: A uniform distribution was applied; i.e., the same water withdrawal amount was applied to each month within a year.

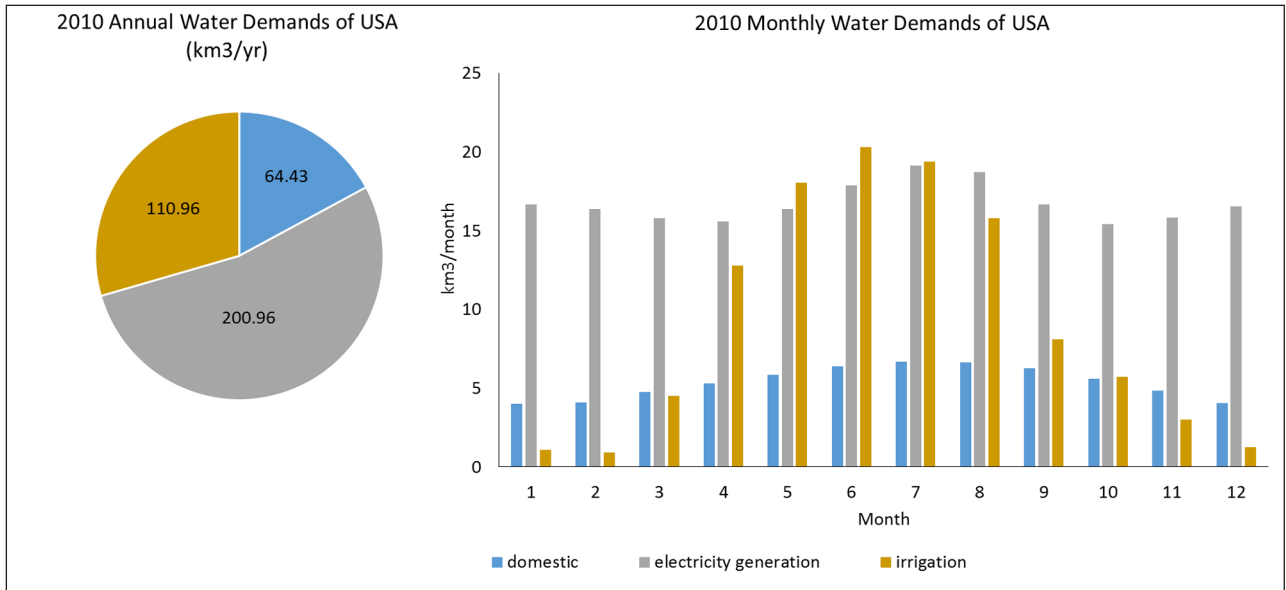
An example of data products from temporal downscaling was illustrated in **Figure 2**. Monthly profiles were estimated from annual water withdrawal estimates of USA in 2010 for domestic, electricity generation and irrigation sectors.

Tethys is written in Python (version 2.7) with scientific libraries. Besides the modules, it also provides collected and consolidated data from various sources as inputs. Each of the datasets used by Tethys has clear sources and references that will be beneficial for the users to update and create their own datasets.

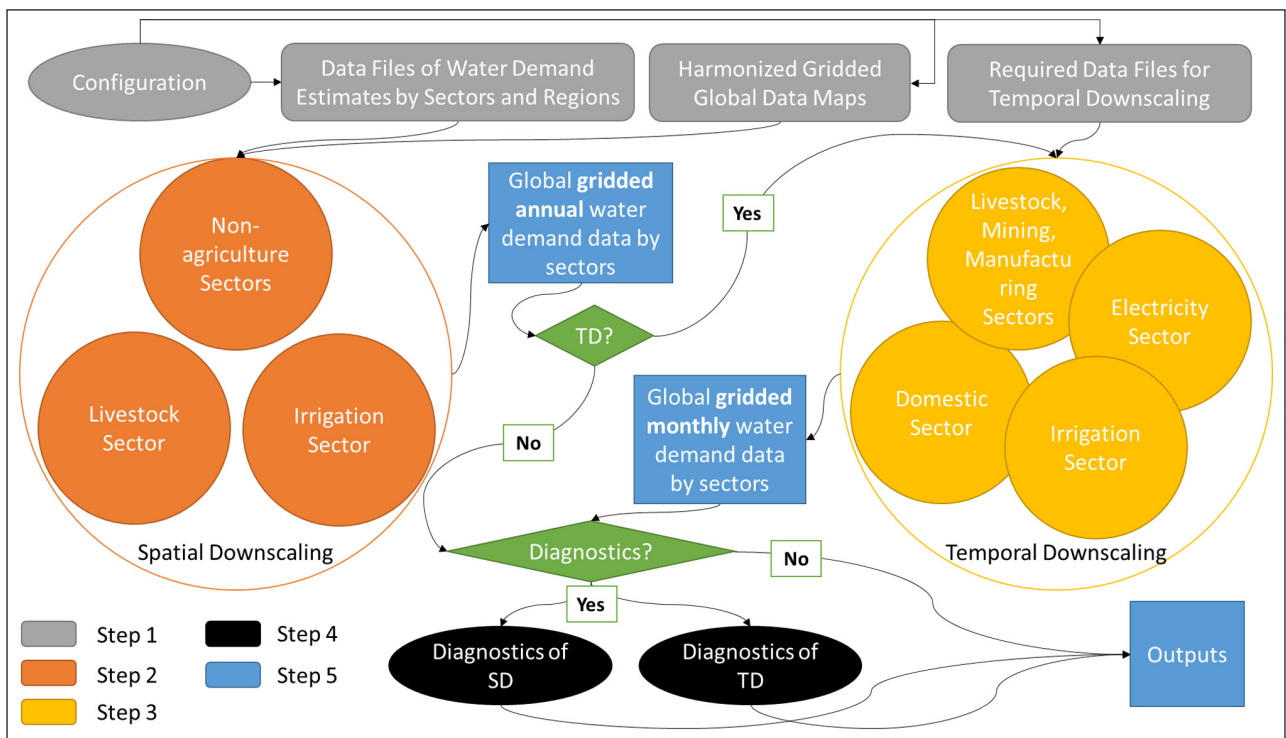
**Implementation and architecture**

Tethys as a downscaling tool follows a sequential flowchart (**Figure 3**):

- Step 1: Import needed data files (module package “tethys\DataReader”)
- Step 2: Spatial downscaling (module package “tethys\SpatialDownscaling”)
- Step 3: Temporal downscaling (module package “tethys\TemporalDownscaling”)



**Figure 2:** Downscaled sectoral (domestic, electricity generation and irrigation) monthly distributions of water withdrawals in USA from annual estimates in 2010.



**Figure 3:** Flowchart of Tethys.

Step 4: Diagnostics of spatial and temporal downscaling (module package “tethys\Diagnostics”)

Step 5: Output all the results of Step 2–5 (module package “tethys\DataWriter”)

For each step, the corresponding module package is also listed. Spatial downscaling (Step 2) is the core of computation flow in Tethys while temporal downscaling (Step 3) is an additional step. The outputs of Step 2, global gridded annual water withdrawal data by sectors, are the inputs of Step 3.

The term “grid” is used to describe the spatial resolution of 0.5 geographic degrees. A global full data map contains a total of 259,200 grid cells ( $360 \times 720$ ) of which 67,420 grid cells are categorized as “land grids” and are considered valid for simulation purposes. In this study, the land grid cells are used to define a “gridded” map according to the coordinates and the indexes of the 67,420 cells on the  $360 \times 720$  grid. To aggregate the gridded data into basin/country/region scale for outputs and diagnostics, certain commonly used global data maps such as IDs of basins/countries/regions are harmonized into the gridded format required by Tethys. The inputs converted using the 67,420 grid cells according to the coordinate data file are called harmonized inputs.

The input interface of Tethys is controlled by the user through the configuration file (e.g. “\*.ini” file). Each downscaling simulation is initiated by importing a single configuration file into Tethys. There are four sections included in the configuration file:

1. Project (Required): This section defines the paths of input and output folders, the output formatting, along with two important options 1) “PerformDiagnostics” determines if diagnostics will be performed; 2) “PerformTemporal” determines if temporal downscaling will be performed.
2. GCAM (Required): As described previously, two formats are allowed 1) GCAM database format; 2) GCAM csv format. The related parameters need to be defined when switching between options for “UseGCAMDatabase”.
3. GriddedMap (Required): This section defines the required global data maps, such as population, irrigation area, and livestock counts for each grid.
4. TemporaDownscaling (Optional, required only if “PerformTemporal = 1” in “Project” section): All the required data files for temporal downscaling are defined in this section. The time period of the data files should be uniformed (e.g. 1971–2010). When “TemporalInterpolation = 1”, Tethys will linearly interpolate the downscaling results when the input data sets are not annual.

The example data files for inputs are all included in the “example\Input” folder while they are divided by subfolders according to the sections described above. The metadata (data source, format, related pre-processing, etc.) of all the input files are described in a document

**Table 1:** Input file names and their corresponding sectors.

Name	Content
pop_tot.csv	Population
irrA.csv	irrigated area for each region, AEZ and crop type
withd_irrV.csv	Water Withdrawal of Irrigation
withd_dom.csv	Water Withdrawal of Domestic
withd_elec.csv	Water Withdrawal of Industrial-Electricity
withd_liv.csv	Water Withdrawal of Livestock
rgn_tot_withd_liv.csv	Water Withdrawal of Livestock (total)
withd_manuf.csv	Water Withdrawal of Industrial-Manufacturing
withd_mining.csv	Water Withdrawal of Resource Extraction

**Table 2:** Output file names and their corresponding sectors.

Sector	SD Results	TD results
Domestic	wddom	twddom
Electricity Generation	wdelec	twdelec
Irrigation	wdirr	twdirr
Livestock	wdliv	twdliv
Manufacturing	wdmfg	twdmfg
Mining	wdmin	twdmin
Non-Agriculture	wdnonag	–
Total	wdtotal	–

called “ReadMe\_IO\_Data.pdf”, which is included in the document folder “docs”.

As described previously, data files of water withdrawal by sectors and region are imported in Tethys, representing the datasets to be downscaled. Since Tethys was originally designed to link to GCAM, a GCAM reader was developed to query information from GCAM database (BaseX format). To extend the usability of Tethys to the wider community, a series of csv files can be prepared following the GCAM csv format as inputs (**Table 1**). The user is required to provide formatted data files for each sector. The format for each file and how to prepare them are introduced in “ReadMe\_IO\_Data.pdf”.

The results after the spatial downscaling step (**Figure 3**), i.e., global annual gridded water withdrawal by sectors, are the default outputs of Tethys. Temporal downscaling is optional and if temporal downscaling step is selected, the results of global monthly gridded water withdrawals by sectors will be additionally outputted (**Table 2**). The outputs can be formatted as classic NetCDF [22] file. The alternative output format is CSV (comma-separated values). The default option generates results in both formats. The default unit is billion  $m^3$  and another optional unit is mm.



Tables and plots from the diagnostics step will also be stored in the output folder if the diagnostics option is selected.

### Quality control

Tethys is a software controlled by inputs. As described in the introduction, these population/livestock/irrigation area data sets used as proxies to spatially downscale different sectors were adopted from widely used open-source databases, which directly determine the quality of the

downscaled results. These population/livestock/irrigation area data sets are widely used and high-quality. The users are encouraged to replace the current input data sets and apply preferred data sets.

A straightforward method to verify the success of the spatial downscaling step is to compare the downscaled results with the original inputs. For example, the following information showed the comparison between the global total values of spatially downscaled results and aggregated results of the original GCAM outputs:

```
---Spatial Downscaling Diagnostics (Global): downscaled results vs. aggregated results from GCAM
(Total Water, km3/yr)
Year 2005 : 3019.53988001      3019.55000639      Diff= -0.0101263749998
Year 2010 : 3253.31261669      3253.32433411      Diff= -0.0117174209977
Year 2015 : 3446.70647763      3446.71935673      Diff= -0.0128790970007
Year 2020 : 3563.76181958      3563.77567633      Diff= -0.0138567450035
Year 2025 : 3730.10510977      3730.12000467      Diff= -0.014894899004
-----Diagnostics information is saved to:
../../Output/Test001/Diagnostics_Spatial_Downscaling.csv
```

The differences were insignificant indicating that water withdrawals at large scale (e.g. region/basin) are simulated at local scale (e.g. grid). A full table of comparison (“Diagnostics\_Spatial\_Downscaling.csv”) can be found in the output folder, which will help the user to examine the downscaled results by year, region and sector in case large differences are observed.

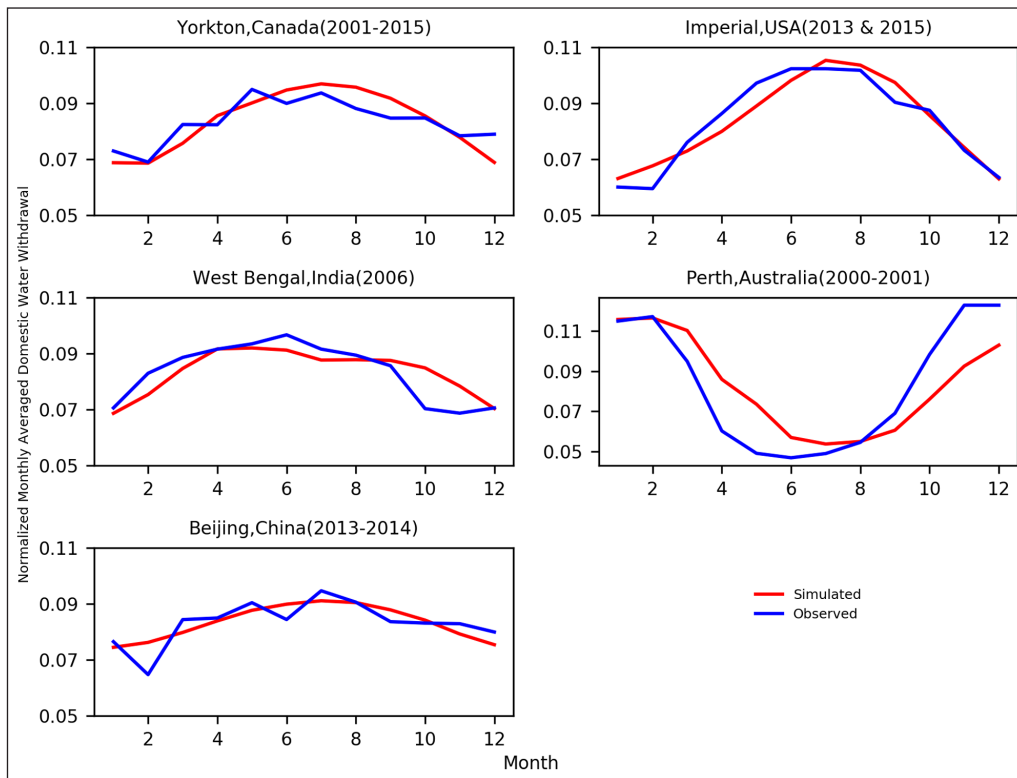
Since the temporal downscaling step was performed using different algorithms among sectors, the

diagnostics module provides different methods to examine the quality of the downscaling results. Results of livestock, mining and manufacturing are not considered for diagnostics while downscaled results of irrigation, domestic and electricity generation are inspected. Similar to spatial downscaling, the global total values of temporal downscaled results and aggregated results before temporal downscaling are compared:

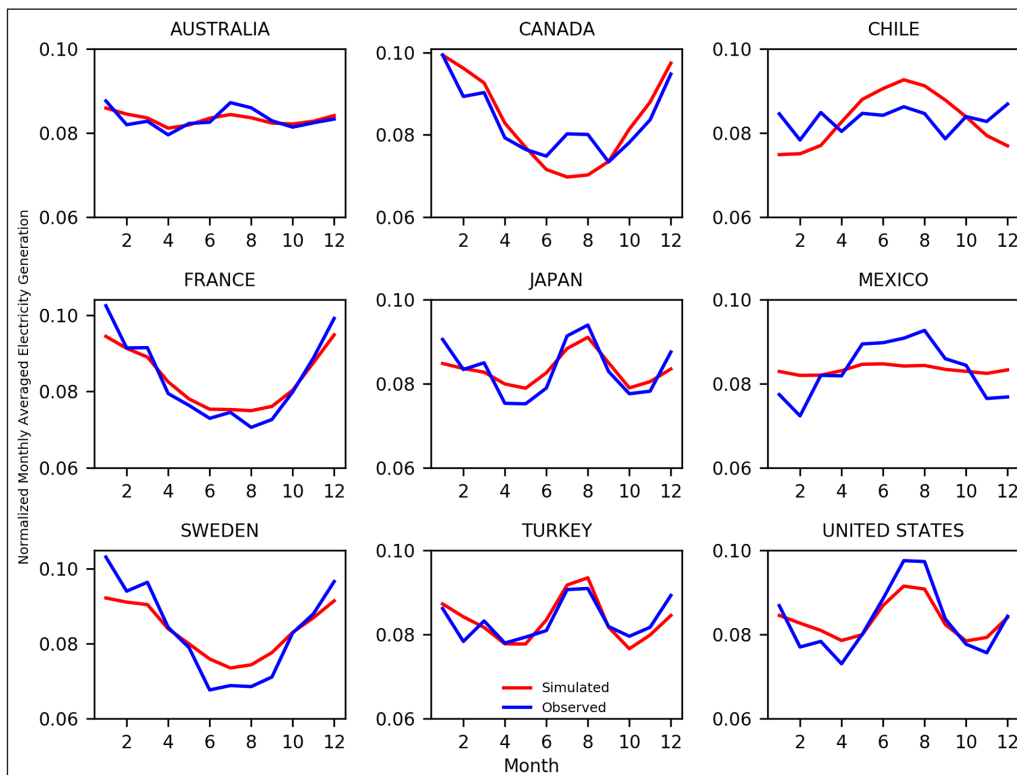
```
---Temporal Downscaling Diagnostics (Global): downscaled results vs. results before temporal
downscaling (Total Water, km3/yr)
-----Irrigation-----
Year 2005 : 1611.86438331      1611.86438331      Diff= 2.27373675443e-13
Year 2006 : 1642.38442693      1642.38442693      Diff= -4.54747350886e-13
Year 2007 : 1672.90447055      1672.90447055      Diff= -4.54747350886e-13
Year 2008 : 1703.42451417      1703.42451417      Diff= 2.27373675443e-13
Year 2009 : 1733.94455779      1733.94455779      Diff= 0.0
Year 2010 : 1764.46460142      1764.46460142      Diff= -6.8212102633e-13
-----Domestic-----
Year 2005 : 456.71      456.71      Diff= 0.0
Year 2006 : 460.118      460.118      Diff= -1.70530256582e-13
Year 2007 : 463.526      463.526      Diff= 0.0
Year 2008 : 466.934      466.934      Diff= -1.70530256582e-13
Year 2009 : 470.342      470.342      Diff= 5.68434188608e-14
Year 2010 : 473.75      473.75      Diff= 0.0
-----Electricity Generation-----
Year 2005 : 540.376128006      540.37612801      Diff= -3.8929783841e-09
Year 2006 : 544.776521342      544.776521326      Diff= 1.61905973073e-08
Year 2007 : 549.176914654      549.176914641      Diff= 1.27258772409e-08
Year 2008 : 553.577307938      553.577307957      Diff= -1.83796373676e-08
Year 2009 : 557.977701031      557.977701272      Diff= -2.40958343056e-07
Year 2010 : 562.378094473      562.378094588      Diff= -1.15136913337e-07
```

The comparison details for irrigation can be found in a csv file in the output folder (“Diagnostics\_Temporal\_Downscaling\_Irrigation.csv”). Two figures adopted from [5] are plotted to monitor domestic and electricity generation sectors, since the downscale algorithms are not based on proxies or uniform

distribution. The simulated mean monthly domestic water withdrawals were displayed in **Figure 4**, with reasonable agreement with collected observations in some listed urban centres and countries [5]. **Figure 5** shows the comparison between simulated and observed monthly water withdrawals for electricity



**Figure 4:** Example of diagnostics plot for comparison between observed and simulated monthly averaged domestic water withdrawal (normalized) in five cities.



**Figure 5:** Example of diagnostics plot for comparison between observed and simulated monthly averaged electricity generation (normalized) in nine countries.

generation during 2000–2012 in 9 OECD countries[5]. It is found that the simulations agree well with observations in most of the countries. Perfect matches in **Figures 4** and **5** are not expected considering the

inherent uncertainties [5] in estimating monthly profiles of water withdrawals.

The user is able to get familiar with the features and I/O interface of Tethys by a comprehensive example case. This

case teaches how to spatially and temporally downscale a datasets of 32 regions and 5 years in 2005, 2010, 2015, 2020 and 2025. The available input data for temporal downscaling is in the period of 1971–2010. Thus, the interpolated temporal downscaling results will be saved for 72 months from 2005 to 2010 (2005/01, 2005/02 ... 2010/11, 2010/12). The name of the configuration file is “config.ini” and the outputs are saved in the folder of “example\Output\Test001”. The example will print the following messages at the beginning and at the end into the log file when it runs successfully:

```
Project Name       : Test001
Input Folder      : ../../Input/
Output Folder     : ../../Output/Test001/
GCAM CSV Folder  : ../../Input/GCAM/CSV/
                  Case001/
Region Info Folder : ../../Input/rgn32/
Start Run_Disaggregation...
.....
End Run_Disaggregation...
---Disaggregation: 103.512000084 seconds ---
Save the gridded water usage results for each
withdrawal category in NetCDF format
(Unit: km3/yr)
Save the monthly water usage results for each
withdrawal category (Unit: km3/month)
---Output: 75.7409999371 seconds ---
('End Project:  ', 'Test001')
```

An automatically created log file will be saved in the output folder, that lists:

- 1) model settings;
- 2) progress and time cost for each step;
- 3) information of regions, years, and adjustment to region maps;
- 4) used population and irrigation data for each year;
- 5) information of unassigned GCAM data during downscaling of livestock and irrigation;
- 6) diagnostics (the comparison results showed above will be printed into the log file);
- 7) output format and unit;
- 8) warnings and errors if applicable.

## (2) Availability

### Operating system

Tethys has been tested successfully on Linux (64-bit), Windows 7 and Mac OS X.

### Programming language

Python (2.7.11)

### Additional system requirements

As modules using enormous global gridded datasets, a minimum memory size of 8GB is recommended and memory capacity determines how fast the code is able to run.

### Dependencies

- NumPy (version 1.13.1)
- Scipy (version 0.18.1)
- Matplotlib (version 2.0.2)
- Pandas (version 0.19.2)
- configobj (version 5.0.6)

## Software location

### Archive

**Name:** GitHub

**Persistent identifier:** <https://github.com/JGCRI/tethys/releases>

**Licence:** BSD 2-Clause

**Publisher:** Chris R Vernon

**Version published:** v1.0.0

**Date published:** 22/09/2017

### Code repository

**Name:** GitHub

**Identifier:** <https://github.com/JGCRI/tethys>

**Licence:** BSD 2-Clause <https://github.com/JGCRI/tethys/blob/master/LICENSE>

**Date published:** 22/09/2017

### Language

English

### Contact

For questions, technical supporting and user contribution, please contact:

Li, Xinya [Xinya.Li@pnnl.gov](mailto:Xinya.Li@pnnl.gov)

Vernon, Chris R [Chris.Vernon@pnnl.gov](mailto:Chris.Vernon@pnnl.gov)

Link, Robert P [Robert.Link@pnnl.gov](mailto:Robert.Link@pnnl.gov)

Hejazi, Mohamad I [Mohamad.Hejazi@pnnl.gov](mailto:Mohamad.Hejazi@pnnl.gov)

### Installation

The “InstallationRequirements” file in “docs” on the repository is to help the user set up the Python environment for a proper run. It explains the steps required for a user to download and install the software with all its dependencies. Also, “setup.py” file is included in the repository.

## (3) Reuse potential

The Python language and the dependent library packages used are all open-source. Tethys is highly modularized and designed for easy installation. The modules can be used independently by the user, which also allows the future development and feasibility of user contribution with least effort. Modification of a certain step could be restricted to the corresponding module. Extension of the model is achievable by adding a new module to an existing sub-folder or a new sub-folder.

All the source codes are in “tethys”. “example” folder contains inputs, outputs and configuration file of example cases. The documents are included in “docs”. The user is able to install Tethys as a Python package by running “setup.py” from terminal or command line:

```
$ python setup.py install
```

After installation, Tethys is able to be imported through “model” class as follows in a Python script:

```
from demeter_w.model import DemeterW
```

And the user is able to run the Tethys model and obtain the outputs as simple as follows in a Python script:

```
dmw = DemeterW('config.ini')
```

Another way to run the downscaling model is by calling different modules from the main function. In the source code package of Tethys, “demeter\_w\run\_disaggregation.py” contains the main function that executes the model steps described in “Implementation and architecture” section. The user is recommended to dig into this module to learn the workflow of Tethys. A simple example script of calling the main function directly is as follows:

```
import demeter_w.DataReader.IniReader as IniReader
from demeter_w.DataWriter.OUTWriter import
OutWriter
from demeter_w.Run_Disaggregation import run_
disaggregation as Disaggregation

# Read simulator settings from ini file.
settingFile = 'config.ini'
settings = IniReader.getSimulatorSettings(setting
File)
# Execute the main function
OUT, GISData = Disaggregation(settings)
# Output the results
OutWriter(settings, OUT, GISData)
```

The reusabilities of Demter-W were also determined by the availability of input data sets. A common desire is a different spatial resolution. Although, the default spatial resolution is 0.5-degree. Tethys can be adjusted to downscale to a different spatial resolution rather than 0.5-degree by updating inputs. The constraint of spatial resolution comes from the input data files not from the algorithm. In “Input” folder, there is a file named “coordinates.csv”. This file listed the 67420 grid cells and their corresponding indexes on the global map (360 × 720) of 0.5 geographic degrees. Tethys reads in this file and stores the grid information (67420 cells) as the “base” map. All the other required gridded maps as inputs should be converted according to this “base” map before being used by Tethys (e.g. All the data files in “harmonized\_inputs”). Thus, if another resolution is desired (e.g. 0.25 degree). The user needs to obtain or pre-process all the related input data files in that spatial resolution. For example, if 0.25 degree is desired. The global map size will be 720 × 1440 and 269680 land cells may be considered in the “coordinates.csv”.

The input formats of Tethys relate closely to GCAM and there are several limitations:

1. The required input data files for temporal downscaling. For example, the multiple electricity inputs are required by the downscaling model we used in Tethys (described in [5]). Thus, the user needs follow the data formats as described in configuration file to run temporal downscaling part.
2. As described previously, a series of formatted csv files (annual regional/basin data to be downscaled) are required and should be provided by the user.
3. Gridded livestock data maps are divided into six categories in “example\Input\harmonized\_inputs”. These categories might be not applicable for other studies.

Although Tethys was used intensively with GCAM data sets, it can still be used with other non-GCAM data sets at region/country/basin scale after reformatting of the input data sets and updating of the gridded data maps. For example, Tethys was applied to a non-global domain (e.g. US states data set was downscaled in [5]). The input csv files in the scale of US states rather than regions were generated and the corresponding region maps were replaced by USGS maps (Cells are not US states were excluded by assigning zeros). Again, Tethys is constrained and driven by input data files. It is essential for the users to get familiar with the formats of all the needed data files described in “tethys/docs/ReadMe\_IO\_Data.pdf”. So they are able to prepare their own data sets to execute Tethys if desired.

Documentation is organized through intensive comments inside the python code and the example configuration file. Execution will also produce a detailed log file lists model settings, the processing steps, CPU cost and warnings if applicable. The users can get support by contacting the authors when issues/bugs are found. The users may also contact the authors for contributions to the code base. The following guidance documents will help the users to get familiar with Tethys in applications:

1. The installation requirements can be referred in the pdf file “InstallationRequirements.pdf” in the “docs” folder on the repository.
2. Inside the “docs” folder, an introduction file (“ReadMe\_IO\_Data.pdf”) is included helping the user to get familiar with the data source and format of each input data file.

Tethys is founded as a member of an integrated modelling software for global water withdrawal, supply, and scarcity, which the authors’ team is continuing to develop. To make Tethys more general and more reusable is the major goal by increasing the flexibility in the modules to deal with different data limitations. A post-processing module package is under development that is able to plot gridded data on global maps by requirements, which can be linked directly to Tethys and other softwares to provide visualized gridded data results to users. And also the team is collecting available open-source water demand data from other studies in fine scale for certain domains (global data may not be achievable, reported historical estimates of sectoral water withdrawals are often sparse and incomplete) to provide comparison references for historical estimates and future predictions.

#### Acknowledgements

This research was supported by the Office of Science of the U.S. Department of Energy through the Integrated Assessment Research Program.

#### Competing interests

The authors have no competing interests to declare.




## References

1. **Li, X, Vernon, C R, Hejazi, M I, Link, R P, Feng, L, Liu, Y and Rauchenstein, L T** 2017 Xanthos – A Global Hydrologic Model. *Journal of Open Research Software*, 5(1): 21. DOI: <https://doi.org/10.5334/jors.181>
2. **Edmonds, J and Reilly, J M** 1985 *Global Energy: Assessing the Future*, 317. Oxford University Press, New York.
3. **Edmonds, J, Wise, M, Pitcher, H, Richels, R, Wigley, T and Maccracken, C** 1997 An integrated assessment of climate change and the accelerated introduction of advanced energy technologies—an application of MiniCAM 1.0. *Mitigation and adaptation strategies for global change*, 1(4): 311–339. DOI: <https://doi.org/10.1023/B:MITI.0000027386.34214.60>
4. **Hejazi, M I, Edmonds, J, Clarke, L, Kyle, P, Davies, E, Chaturvedi, V, Wise, M, Patel, P, Eom, J and Calvin, K** 2014 Integrated assessment of global water scarcity over the 21st century under multiple climate change mitigation policies. *Hydrology and Earth System Sciences*, 18: 2859–2883. DOI: <https://doi.org/10.5194/hess-18-2859-2014>
5. **Huang, Z, Hejazi, M, Li, X, Tang, Q, Leng, G, Liu, Y, Döll, P, Eisner, S, Gerten, D, Hanasaki, N and Wada, Y** 2017 Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns. *Hydrology and Earth System Sciences Discussions*. DOI: <https://doi.org/10.5194/hess-2017-551>
6. **Wada, Y, Van Beek, L P H, Viviroli, D, Dürr, H H, Weingartner, R and Bierkens, M F** 2011 Global monthly water stress: 2. Water withdrawal and severity of water stress. *Water Resources Research*, 47(7): W07518. DOI: <https://doi.org/10.1029/2010WR009792>
7. **Siebert, S, Döll, P, Feick, S, Hoogeveen, J and Frenken, K** 2007 Global map of irrigation areas version 4.0.1. Johann Wolfgang Goethe University, Frankfurt am Main, Germany/Food and Agriculture Organization of the United Nations, Rome, Italy.
8. **Portmann, F T, Siebert, S, Bauer, C and Döll, P** 2008 Global dataset of monthly growing areas of 26 irrigated crops: Version 1.0. University of Frankfurt, Germany.
9. **Klein Goldewijk, K, Beusen, A, Van Drecht, G and De Vos, M** 2011 The HYDE 3.1 spatially explicit database of human induced global land use change over the past 12,000 years. *Global Ecology and Biogeography*, 20(1): 73–86. DOI: <https://doi.org/10.1111/j.1466-8238.2010.00587.x>
10. **Center for International Earth Science Information Network (CIESIN) – Columbia University** 2016 Gridded Population of the World, Version 4 (GPWv4): Population Count. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY. DOI: <https://doi.org/10.7927/H4X63JVC>
11. **Siebert, S, Henrich, V, Frenken, K and Burke, J** 2013 Global Map of Irrigation Areas version 5. Rheinische Friedrich-Wilhelms-University, Bonn, Germany/Food and Agriculture Organization of the United Nations, Rome, Italy.
12. **Wint, W and Robinson, T** 2007 Gridded livestock of the world. *Food and Agriculture Organization (FAO), report 131*. Rome.
13. **Alcamo, J and Henrichs, T** 2002 Critical regions: A model-based estimation of world water resources sensitive to global changes. *Aquatic Sciences-Research Across Boundaries*, 64(4): 352–362. DOI: <https://doi.org/10.1007/PL00012591>
14. **Flörke, M and Alcamo, J** 2004 European outlook on water use. *Center for Environmental Systems Research, University of Kassel, Final Report*. EEA/RNC/03/007, 83.
15. **Hanasaki, N, Kanae, S, Oki, T, Masuda, K, Motoya, K, Shirakawa, N, Shen, Y and Tanaka, K** 2008 An integrated model for the assessment of global water resources—Part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences*, 12(4): 1007–1025. DOI: <https://doi.org/10.5194/hess-12-1007-2008>
16. **Hanasaki, N, Kanae, S, Oki, T, Masuda, K, Motoya, K, Shirakawa, N, Shen, Y and Tanaka, K** 2008 An integrated model for the assessment of global water resources—Part 2: Applications and assessments. *Hydrology and Earth System Sciences*, 12(4): 1027–1037. DOI: <https://doi.org/10.5194/hess-12-1027-2008>
17. **Rost, S, Gerten, D, Bondeau, A, Lucht, W, Rohwer, J and Schaphoff, S** 2008 Agricultural green and blue water consumption and its influence on the global watersystem. *Water Resources Research*, 44(9): W09405. DOI: <https://doi.org/10.1029/2007WR006331>
18. **Van Beek, L P H, Wada, Y and Bierkens, M F** 2011 Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, 47(7): W07517. DOI: <https://doi.org/10.1029/2010WR009791>
19. **Voisin, N, Liu, L, Hejazi, M, Tesfa, T, Li, H, Huang, M, Liu, Y and Leung, L R** 2013 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(11): 4555–4575. DOI: <https://doi.org/10.5194/hess-17-4555-2013>
20. **Hejazi, M I, Voisin, N, Liu, L, Bramer, L M, Fortin, D C, Hathaway, J E, Huang, M, Kyle, P, Leung, L R, Li, H Y and Liu, Y** 2015 21st century United States emissions mitigation could increase water stress more than the climate change it is mitigating. *Proceedings of the National Academy of Sciences*, 112(34): 10635–10640. DOI: <https://doi.org/10.1073/pnas.1421675112>
21. **Kim, S H, Hejazi, M, Liu, L, Calvin, K, Clarke, L, Edmonds, J, Kyle, P, Patel, P, Wise, M and Davies, E** 2016 Balancing global water availability and use at basin scale in an integrated assessment model. *Climatic Change*, 136(2): 217–231. DOI: <https://doi.org/10.1007/s10584-016-1604-6>
22. **An Introduction to NetCDF** [http://www.unidata.ucar.edu/software/netcdf/docs/netcdf\\_introduction.html](http://www.unidata.ucar.edu/software/netcdf/docs/netcdf_introduction.html).

**How to cite this article:** Li, X, Vernon, C R, Hejazi, M I, Link, R P, Huang, Z, Liu, L and Feng, L 2018 Tethys – A Python Package for Spatial and Temporal Downscaling of Global Water Withdrawals. *Journal of Open Research Software* 6: 9, DOI: <https://doi.org/10.5334/jors.197>

**Submitted:** 29 September 2017 **Accepted:** 07 December 2017 **Published:** 09 February 2018

**Copyright:** © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Journal of Open Research Software* is a peer-reviewed open access journal published by Ubiquity Press

OPEN ACCESS

