
SOFTWARE METAPAPER

A Global Hydrologic Framework to Accelerate Scientific Discovery

Chris R. Vernon¹, Mohamad I. Hejazi², Sean W. D. Turner², Yaling Liu³, Caleb J. Braun², Xinya Li¹ and Robert P. Link²

¹ Pacific Northwest National Laboratory, Richland, WA, US

² Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, Maryland, US

³ Department of Earth and Environmental Engineering, Columbia University, New York, US

Corresponding author: Chris R. Vernon (chris.vernon@pnnl.gov)

With the ability to simulate historical and future global water availability on a monthly time step at a spatial resolution of 0.5 geographic degree, the Python package Xanthos version 1 provided a solid foundation for continuing advancements in global water dynamics science. The goal of Xanthos version 2 was to build upon previous investments by creating a Python framework where core components of the model (potential evapotranspiration (PET), runoff generation, and river routing) could be interchanged or extended without having to start from scratch. Xanthos 2 utilizes a component-style architecture which enables researchers to quickly incorporate and test cutting-edge research in a stable modeling environment prebuilt with diagnostics. Major advancements for Xanthos 2 were also achieved by the creation of a robust default configuration with a calibration module, hydropower modules, and new PET modules, which are now available to the scientific community.

Keywords: Python; SciPy; Global Hydrology; Hydrologic Framework; Xanthos; Runoff; River Routing; Potential Evapotranspiration

Funding statement: This research was supported by the U.S. Department of Energy, Office of Science, as part of research in Multi-Sector 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.

(1) Overview

Introduction

Research in water-energy-land systems (WEL) [1, 2] and understanding the implications of water scarcity on WEL systems [3–5] is rapidly progressing. Additional impacts to water resource management are now also being considered when addressing supply-demand cycles and reservoir effects from potential water shortages [6]. Xanthos 2 was created to accommodate these progressive advancements in water-energy-land science that support cutting-edge research in human-Earth systems interactions. Currently, Xanthos 2 provides water supply, accessibility, and hydropower potential to the Global Change Assessment Model (GCAM) which is an integrated human-Earth systems model [7–9]. Xanthos 2 was built upon previous development by Li et al. [10] by creating a Python framework where core components of the model (potential evapotranspiration (PET), runoff generation, and river routing) can be interchanged or extended without having to start from scratch. Xanthos 2 utilizes a component-style architecture which enables researchers to quickly incorporate and test cutting-edge research in

a stable modeling environment prebuilt with diagnostics to keep up with advancing science objectives. The new default configuration of Xanthos now supports the Penman-Monteith [11, 12] PET module which incorporates more comprehensive climate drivers as well as dynamic land cover data to allow land atmosphere interactions. The *abcd* water balance module was also added as the default runoff component to account for groundwater recharge and baseflow in runoff estimates [13]. Additionally, the following enhancements were also made: improved water velocity considerations for the Modified River Transport Model (MRTM) routing module [14, 15], hydropower production assessment and potential capacity modules, and a built-in differential evolution optimization module to calibrate *abcd* parameters based on benchmark global runoff.

Implementation and architecture

Xanthos 2 was created as a Python package to facilitate reuse. Xanthos runs are setup using a configuration file to describe which components of the model are used, data parameterization, run-specific settings, and more. A

GitHub Wiki explaining how to set up a Xanthos run, and the required project level settings of the configuration file is available on the Xanthos GitHub website.¹ A list of available components with references to the algorithms from which they are derived is described in **Table 1**; though, the GitHub Wiki details each component's configuration settings and variable descriptions for all additional model components.² The following is a simple example of how to call and run Xanthos in a Python script after it has been setup:

```
from xanthos import Xanthos

def run(ini):
    # instantiate model
    xth = Xanthos(ini)

    # run intended configuration
    xth.execute()

    return xth

if __name__ == "__main__":
    # full path to parameterized config file
    ini = '<the directory path to your example
          directory>/pm_abcd_mrtm.ini'

    # run the model
    xth = run(ini)
```

Calling the `execute()` method after instantiating the Xanthos class will start a run and initialize the logger. The user will be updated as the model progresses through each step and as outputs are generated. **Table 2** describes the expected outputs of a Xanthos run using the default core configuration and the accessible water, hydropower, diagnostics, and time series plots post-processing modules.

Xanthos 2 provides an extensible framework to interchange or add additional PET, runoff, and routing components of the core monthly water balance module

(**Figure 1**). Post-processing modules (e.g., accessible water, hydropower potential, etc.) can also be added to create output products. A complete workflow for integrating a new component into Xanthos is provided in detail in an associated GitHub Wiki.³

Calibration module

The new default configuration of Xanthos contains the *abcd* water balance module to generate runoff as interpreted by Liu et al. [13]. The *abcd* module is driven by five aptly-named parameters: *a* which controls the amount of runoff that occurs when soils are under-saturated; *b* which controls the saturation level of the soil; *c* which defines the degree of recharge to groundwater; *d* which controls the rate of groundwater discharge; and *m* for snowpack accumulation and snowmelt estimates which was added by Martinez and Gupta [22]. The performance of the *abcd* module is influenced by per-basin *a, b, c, d, m* parameters which need to be calibrated based on the chosen climate forcing, PET, and runoff routines.

To accommodate the need for recalibration depending on experimental design, we chose to include a preconfigured calibration module in Xanthos 2 that uses SciPy's differential evolution (DE) optimization function [23] based off of the algorithm presented in Storn and Price [21]. The DE method is useful for global optimization problems and is a stochastic population-based optimization method [23]. The objective function provided in Xanthos 2 evaluates the Kling-Gupta efficiency (KGE) which is a goodness-of-fit measure between simulated and observed runoff [24]; however, alternate methods could be utilized.

Figure 2 demonstrates the optimization module's ability to calibrate Xanthos 2 runoff to the complex Variable Infiltration Capacity (VIC) model's monthly

Table 1: Available components for Xanthos 2, their component types, and references to the supporting algorithms.

Component	Type	Algorithm Reference
Penman-Monteith	PET	McMahon et al. 2012; Monteith 1965 [11, 12]
Hargreaves-Samani	PET	Hargreaves and Samani 1982 [16]
Hargreaves	PET	Hargreaves et al. 1985 [17]
Thorntwaite	PET	Thorntwaite 1948 [28]
<i>abcd</i>	Runoff	Liu et al. 2018 [13]
Global Water Availability Model (GWAM)	Runoff	Hejazi et al. 2014 [3]
Modified River Transport Model (MRTM)	Routing	Zhou et al. 2015; Li et al. 2015 [14, 15]
Diagnostics	Post-processing	Li et al. 2017 [10]
Time Series Plotting	Post-processing	Li et al. 2017 [10]
Accessible Water	Post-processing	Kim et al. 2016 [18]
Hydropower – Potential	Post-processing	Zhou et al. 2015 [14]
Hydropower – Actual	Post-processing	Turner et al. 2017a; Turner et al. 2017b [19, 20]
Calibration – <i>abcd</i>	Ancillary	Storn and Price 1997 [21]

Table 2: Xanthos expected outputs when using the default configuration and the accessible water, hydropower – potential, hydropower – actual, diagnostics, and time series plots post-processing modules.

Output	Source	Units
Potential Evapotranspiration (PET) per grid cell	PET	Monthly or annually; mm/time or km ³ /time
Actual Evapotranspiration (AET) per grid cell	Runoff	Monthly or annually; mm/time or km ³ /time
Gridded runoff per grid cell	Runoff	Monthly or annually; mm/time or km ³ /time
Aggregated total runoff per basin	Runoff	Monthly or annually; mm/time or km ³ /time
Aggregated total runoff per GCAM country	Runoff	Monthly or annually; mm/time or km ³ /time
Aggregated total runoff per GCAM region	Runoff	Monthly or annually; mm/time or km ³ /time
Soil moisture per grid cell	Runoff	Monthly or annually; mm/time or km ³ /time
Average channel flow per grid cell	Routing	Monthly; m ³ /second
Accessible water	Accessible water	Monthly or annually; mm/time or km ³ /time
Exploitable hydropower potential by GCAM region	Hydropower – Potential	Monthly or annually; EJ/time
Technical hydropower potential by GCAM region	Hydropower – Potential	Monthly or annually; EJ/time
Actual hydropower production	Hydropower – Actual	Monthly or annually; EJ/time
Diagnostics – runoff by basin	Diagnostics	Annually; km ³ /year
Diagnostics – runoff by GCAM country	Diagnostics	Annually; km ³ /year
Diagnostics – runoff by GCAM region	Diagnostics	Annually; km ³ /year
Time series plots – streamflow by basin	Time Series Plots	Monthly or annually; mm/time or km ³ /time
Time series plots – streamflow by GCAM country	Time Series Plots	Monthly or annually; mm/time or km ³ /time
Time series plots – streamflow by GCAM region	Time Series Plots	Monthly or annually; mm/time or km ³ /time

runoff projections [25] when forced by WATER and global Change (WATCH) [26] climate data for years 1971–1991. The resulting median KGE for all basins was a respectable 0.9 which is an improvement to the calibration results of 0.79 generated by Liu et al. [13] when using MATLAB's built-in genetic algorithm (GA) [27]. A tutorial describing how to setup and use the calibration module in Xanthos can be viewed on our GitHub repository.³

Installation

Xanthos can be installed using the following steps:

1. This repository uses the Git Large File Storage (LFS) extension (see <https://git-lfs.github.com/> for details). Please install GitLFS and run the following command before cloning if you do not already have Git LFS initialized: `git lfs install`.
2. Clone Xanthos into your desired location `git clone https://github.com/JGCRI/xan-`

`thos.git`. Some Windows users have had better luck with `git lfs clone https://github.com/JGCRI/xanthos.git`.

3. Make sure that `setuptools` is installed for your Python version. This is what will be used to support the installation of the Xanthos package.
4. From the directory you cloned Xanthos into run `python setup.py install`. This will install Xanthos as a Python package on your machine and install the needed dependencies. If installing in a high-performance computing environment, a community user advised that it is best to install the `anaconda` environment before running the installation command. HPC environments may also require the use of the `--user` flag in the install command to avoid permissions errors.
5. Setup your configuration file (.ini). Examples are located in the "example" directory. Be sure to change

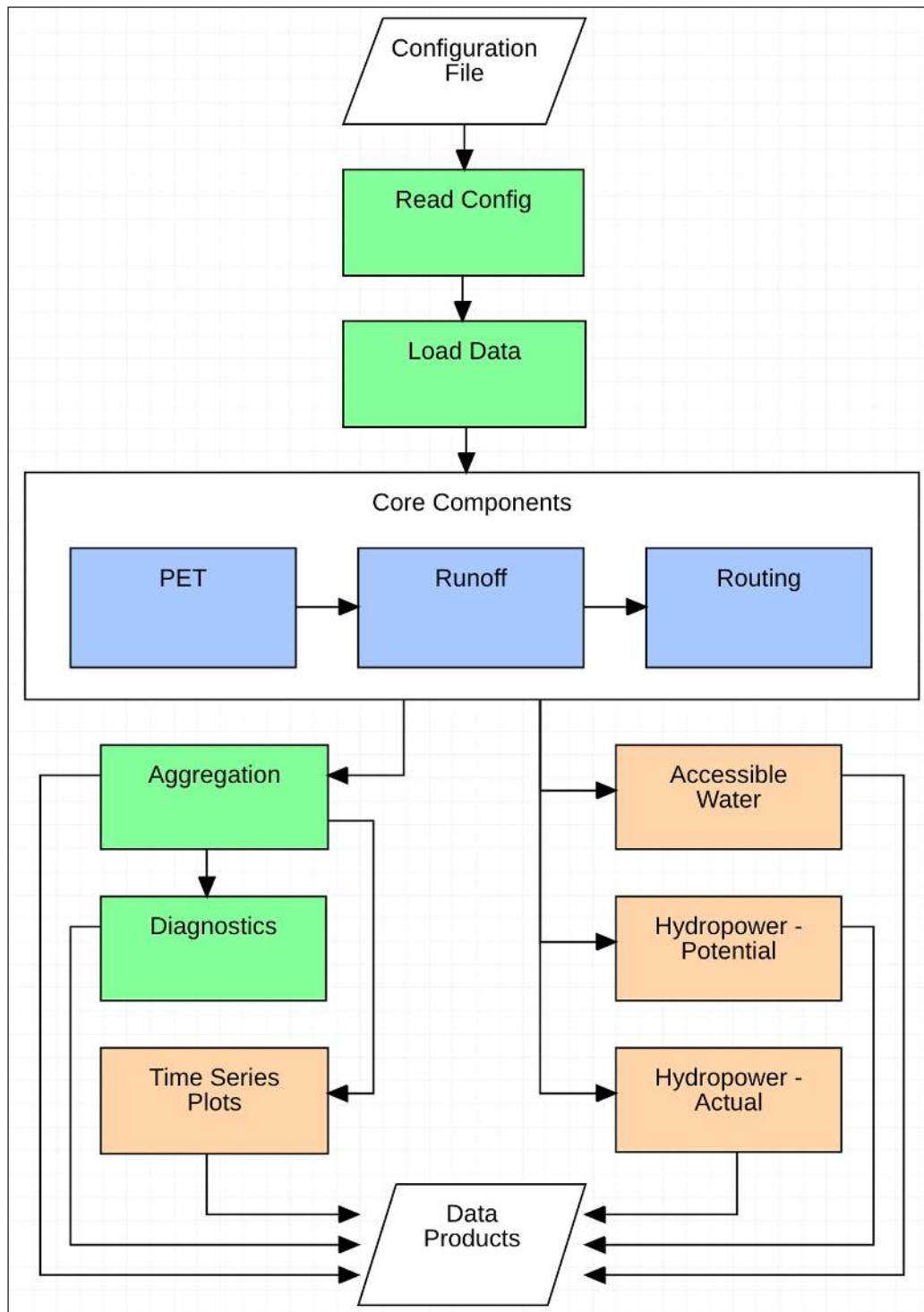


Figure 1: Xanthos 2 module diagram and workflow for the default configuration. Components in green provide built-in functionality; blue components are the interchangeable core modules; and orange components are post-processing modules that have been extended.

the root directory to the directory that holds your data (use the “xanthos/example” directory as an example).

6. If running Xanthos from an IDE: Be sure to include the path to your config file. See the “xanthos/example/example.py” script as a reference.
7. If running Xanthos from terminal: Run model.py found in xanthos/xanthos/model.py, which passes the full path to the config file as the only argument. (e.g., `python model.py <dirpath>/config.ini`).

Quality control

Xanthos has been functionally tested on Linux, Mac OSX, and Windows to ensure model reusability. Xanthos has a built-in diagnostics module that compares model outputs to expected output from other mainstream hydrologic models when forced by the same climate assumptions. The calibration module output has also been evaluated as described in the “Calibration module” section of this paper. Exception handling is present throughout the model to ensure data consistency and provide informed

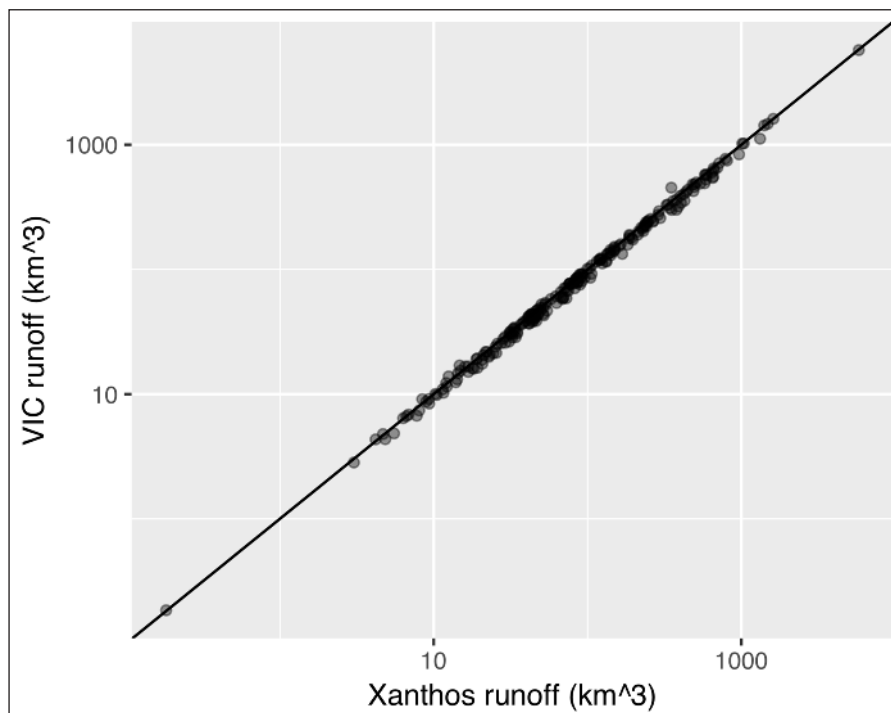


Figure 2: Performance of calibrated runoff by water basin (235 basins globally) from the *abcd* module using SciPy's differential evolution optimization function when compared to VIC model runoff forced by WATCH data for years 1971–1991. Note that runoff is converted to km^3 by multiplying by basin area, and each data point denotes the long term mean annual runoff of a water basin.

feedback to the user for cases where an assumption has been violated. Extensive tests are built-in that validate the configuration file, and the contents thereof, that is provided by the user. Both expected inputs and outputs for each component have been provided as a baseline for comparability if the user decides to extend or alter the code for an intended purpose.

(2) Availability

Operating system

Linux, Mac OSX, Windows 7 and 2012 server.

Programming language

Python 2.7, Python 3

Additional system requirements

We recommend a minimum RAM of 8GB due to routinely processing large datasets.

Dependencies

The following Python 3 packages are required for Xanthos:

- numpy >= 1.11
- scipy >= 0.18
- matplotlib >= 1.3.1, < 3.0
- pandas >= 0.19
- configobj >= 5.0.6
- joblib >= 0.11
- setuptools >= 24.2.0

Software location

Archive

Name: Zenodo

Persistent identifier: <https://doi.org/10.5281/zenodo.2035720>

Licence: BSD 2-Clause

Publisher: Chris R. Vernon

Version published: v2.2.0

Date published: December 7, 2018

Code repository

Name: GitHub

Identifier: <https://github.com/JGCRI/xanthos/tree/v2.2.0>

Licence: BSD 2-Clause

Date published: December 7, 2018

Language

English

(3) Reuse potential

Xanthos 2 was created to be extensible by design. Users can provide their own core and/or post-processing components as described in our GitHub Wiki.⁴ For example, a new PET module could be added into Xanthos and used instead of the default configuration. This allows those who are advancing the science of WEL research to both test their alternate process through integrated testing with the other core components of Xanthos, and then compare their results against the default Xanthos configuration and validate performance using the built-in diagnostics module. The core components of Xanthos are also interchangeable to accommodate research objectives. An example of extending a post-processing component could be adding a drought module that would use runoff as generated by Xanthos and any other required inputs to

examine timing and duration droughts through seamless and citable integration.

Xanthos provides a custom issue template, *Extending Xanthos Issues*,⁵ which guides the user to provide information that our team can use to quickly address any questions that may arise when extending Xanthos.

Notes

- ¹ GitHub Wiki for setting up a Xanthos run: <https://github.com/JGCRI/xanthos/wiki/Tutorial-1:-Setting-up-a-Xanthos-run>.
- ² GitHub Wiki for available modules: <https://github.com/JGCRI/xanthos/wiki/Available-modules>.
- ³ GitHub Wiki for the *abcd* calibration module tutorial: <https://github.com/JGCRI/xanthos/wiki/Tutorial-3:-Calibrate-the-ABCD-runoff-model>.
- ⁴ GitHub Wiki for adding components: <https://github.com/JGCRI/xanthos/wiki/Tutorial-2:-Adding-new-components-to-Xanthos>.
- ⁵ Custom issue form for extending Xanthos: https://github.com/JGCRI/xanthos/blob/master/.github/ISSUE_TEMPLATE/extending-xanthos-issues.md.

Competing Interests

The authors have no competing interests to declare.

References

1. **Endo, A, Tsurita, I, Burnett, K and Orencio, P M** 2015 A Review of the Current State of Research on the Water, Energy, and Food Nexus. *J Hydrol Reg Stud*. DOI: <https://doi.org/10.1016/j.ejrh.2015.11.010>
2. **Cai, X, Wallington, K, Shafiee-Jood, M and Marston, L** 2018 Understanding and Managing the Food-Energy-Water Nexus – Opportunities for Water Resources Research. *Advances in Water Resources*, 111: 259–273. DOI: <https://doi.org/10.1016/j.advwatres.2017.11.014>
3. **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>
4. **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, Liu, Y, Patel, P, Pulsipher, T C, Rice, J S, Tesfa, T K, Vernon, C R and Zhou, Y** 2015 21st century United States emissions mitigation could increase water stress more than the climate change it is mitigating. *Proc Natl Acad Sci*, 112: 10635–10640. DOI: <https://doi.org/10.1073/pnas.1421675112>
5. **Pokhrel, Y N, Hanasaki, N, Wada, Y and Kim, H** 2016 Recent progresses in incorporating human land-water management into global land surface models toward their integration into Earth system models. *WIREs Water*, 3: 548–574. DOI: <https://doi.org/10.1002/wat2.1150>
6. **Di Baldassarre, G, Wanders, N, AghaKouchak, A, Kuil, L, Rangelcroft, S, Veldkamp, T I E, Garcia, M, van Oel, P R, Breinl, K and Van Loon, A F** 2018 Water shortages worsened by reservoir effects. *Nature Sustainability*, 1: 617–622. DOI: <https://doi.org/10.1038/s41893-018-0159-0>
7. **Edmonds, J and Reilly, J M** 1985 Global Energy: Assessing the Future. In: New York: Oxford University Press, 317.
8. **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>
9. **Kim, S H, Edmonds, J, Lurz, J, Smith, S J and Wise, M** 2006 The objECTS framework for integrated assessment: Hybrid modeling of transportation. *The Energy Journal*, (2): 51–80, Special Issue. DOI: <https://doi.org/10.5547/ISSN0195-6574-EJ-VolSI2006-NoSI2-4>
10. **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: 1–7. DOI: <https://doi.org/10.5334/jors.181>
11. **McMahon, T, Peel, M, Lowe, L, Srikanthan, R and McVicar, T** 2012 Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrology and Earth System Sciences Discussions*, 9: 11829–11910. DOI: <https://doi.org/10.5194/hessd-9-11829-2012>
12. **Monteith, J L** 1965 Evaporation and environment. *Symp Soc Exp Biol*, 19: 205–224.
13. **Liu, Y, Hejazi, M, Li, H, Zhang, X and Leng, G** 2018 A hydrological emulator for global applications – HE v1.0.0. *Geoscientific Model Development*, 11(3): 1077–1092. DOI: <https://doi.org/10.5194/gmd-11-1077-2018>
14. **Zhou, Y, Hejazi, M, Smith, S, Edmonds, J, Li, H, Clarke, L, Calvin, K and Thomson, A** 2015 A comprehensive view of global potential for hydro-generated electricity. *Energy & Environmental Science*, 8: 2622–2633. DOI: <https://doi.org/10.1039/C5EE00888C>
15. **Li, H, Leung, L R, Getirana, A, Huang, M, Wu, H, Xu, Y, Guo, J and Voisin, N** 2015 Evaluating Global Streamflow Simulations by a Physically-based Routing Model Coupled with the Community Land Model. *J. of Hydromet*, 16(2): 948–971. DOI: <https://doi.org/10.1175/JHM-D-14-0079.1>
16. **Hargreaves, G H and Samani, Z A** 1982 Estimating potential evapotranspiration. *J. Irrig. Drain. Div*, 108: 225–230.
17. **Hargreaves, G L, Hargreaves, G H and Riley, J P** 1985 Irrigation water requirements for Senegal River basin. *Journal of Irrigation and Drainage Engineering*, 111(3): 265–275. DOI: [https://doi.org/10.1061/\(ASCE\)0733-9437\(1985\)111:3\(265\)](https://doi.org/10.1061/(ASCE)0733-9437(1985)111:3(265))
18. **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>
19. **Turner, S W, Ng, J Y and Galelli, S** 2017a Examining global electricity supply vulnerability to climate change using a high-fidelity hydropower dam model. *Science of the Total Environment*, 590: 663–675. DOI: <https://doi.org/10.1016/j.scitotenv.2017.03.022>
 20. **Turner, S W, Hejazi, M, Kim, S H, Clarke, L and Edmonds, J** 2017b Climate impacts on hydropower and consequences for global electricity supply investment needs. *Energy*, 141: 2081–2090. DOI: <https://doi.org/10.1016/j.energy.2017.11.089>
 21. **Storn, R and Price, K** 1997 Differential Evolution – a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11: 341–359. DOI: <https://doi.org/10.1023/A:1008202821328>
 22. **Martinez, G F and Gupta, H V** 2010 Toward improved identification of hydrological models: A diagnostic evaluation of the “abcd” monthly water balance model for the conterminous United States. *Water Resour. Res.*, 46. DOI: <https://doi.org/10.1029/2009WR008294>
 23. **SciPy** 2018 [ONLINE] *SciPy-Optimize-Differential Evolution*. https://docs.scipy.org/doc/scipy-0.17.0/reference/generated/scipy.optimize.differential_evolution.html#r140 [Accessed 27 July 2018].
 24. **Gupta, H V, Kling, H, Yilmaz, K K and Martinez, G F** 2009 Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(20): 80–91. DOI: <https://doi.org/10.1016/j.jhydrol.2009.08.003>
 25. **Leng, G, Tang, Q and Rayburg, S** 2015 Climate change impacts on meteorological, agricultural and hydrological droughts in China. *Global Planet. Change*, 126: 23–34. DOI: <https://doi.org/10.1016/j.gloplacha.2015.01.003>
 26. **Weedon, G P, Gomes, S, Viterbo, P, Shuttleworth, W J, Blyth, E, Österle, H, Adam, J C, Bellouin, N, Boucher, O and Best, M** 2011 Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *J. Hydrometeorol*, 12: 823–848. DOI: <https://doi.org/10.1175/2011JHM1369.1>
 27. **Deb, K, Pratap, A, Agarwal, S and Meyerivan, T** 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE T. Evolut. Comput*, 6: 182–197. DOI: <https://doi.org/10.1109/4235.996017>
 28. **Thorntwaite, C** 1948 An approach toward a rational classification of climate. *Geographical Review*, 38(1): 55–94. DOI: <https://doi.org/10.2307/210739>

How to cite this article: Chris R. Vernon, Mohamad I. Hejazi, Sean W. D. Turner, Yaling Liu, Caleb J. Braun, Xinya Li and Robert P. Link 2019 A Global Hydrologic Framework to Accelerate Scientific Discovery. *Journal of Open Research Software*, 7: 1. DOI: <https://doi.org/10.5334/jors.245>

Submitted: 20 August 2018 **Accepted:** 10 December 2018 **Published:** 07 January 2019

Copyright: © 2019 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/>.