

# Basin Futures: Supporting water planning in data poor basins

Peter Taylor<sup>a</sup>, Joel Stewart<sup>b</sup>, Joel Rahman<sup>c</sup>, Amit Parashar<sup>d</sup>, Carmel Pollino<sup>d</sup>, Geoff Podger<sup>d</sup>

<sup>a</sup> CSIRO Data61

<sup>b</sup> Catchment Research Pty Ltd

<sup>c</sup> Flowmatters Pty Ltd

<sup>d</sup> CSIRO Land & Water

Email: [peter.taylor@data61.csiro.au](mailto:peter.taylor@data61.csiro.au)

**Abstract:** Water resources are under growing pressures globally. Sustainable water management requires good understanding of water resources, e.g. where water is stored and used, and how future developments may impact on water availability. Developing this understanding requires investment in data, models, and experts, often over multi-year timeframes.

In the context of developing countries, in particular those with institutional gaps in the water sector, these challenges are difficult to overcome. In this paper we discuss Basin Futures, a response to four issues faced by institutions and water policy makers in developing countries: data availability, capability, capacity and costs.

In Basin Futures, we aim to accelerate the progression toward detailed basin planning by offering access to data sets and models that facilitate the initial stages of basin planning.

Basin Futures is a web based modelling platform that brings together global and local datasets to support water planning in developing countries. The system is designed to leverage investment in existing data (global and local), and then use this to empower decision-makers to understand the opportunities and constraints in managing their water resources. Basin Futures provides an enabling environment for planning, cooperation and participation in water management. It allows defensible decisions to be made at a level that is supported by available data.

With the increasing availability and range of global data, e.g. from satellites, and technologies such as cloud computing, there is an opportunity to perform faster water assessments as an early stage of a large water assessment cycle. We have developed a cloud-based approach that aims to give users a quicker route to an initial understanding of water resources in basins.

Basin Futures integrates a range of global datasets, include HydroSHEDS, global precipitation, and river discharge observations, with models built for use in a cloud environment. Together with a web-based application, this allows users to make an initial estimate of current water resources. Core to the value of Basin Futures is that it provides user-customisable scenarios that allow dynamic exploration of future scenarios, for example, construction of new storages, changes to irrigation practices, population growth, and climate change scenarios. Users are then able to visualise impacts on water-related security indices.

The model engine is a reach-based model - rather than a node-link model - where the demands are only associated with local (reach) storages and there is no support for upstream ordering of water or for exploring ground and surface water interactions. This is a deliberate choice to create a system that runs fast enough in web time and provides the user with an initial understanding of the water resources in unregulated systems. This streamlined approach to water resource assessments gives users a powerful starting point for positioning future investments in new knowledge and planning activities. The outputs will also guide where more definitive answers, and data, will be required by the appropriate use of more powerful software.

This paper details the driving requirements for the Basin Futures, the underlying model, and the implementation using modern web technologies.

**Keywords:** *Basin planning, global data, cloud processing, decision support systems*

## 1. INTRODUCTION

Water management challenges are more commonplace in developing countries (Dasgupta and Mäler 2000) due to increasing pressure on water resources. Understanding the state of the water system and its related parts (agriculture, soils etc.) is crucial to improving management practices. The technical modeling required to inform decision making is complex and time-consuming. Our experience in developing countries suggests there is a need for approaches that lower the barrier to entry to water modelling and associated tools.

The development of highly detailed water balance models requires a large investment in data acquisition, interpretation, restructuring, and processing. This increases the cost of performing water modelling activities. Basin Futures is our response to attempt to provide easier and lower costs solutions to inform water planning.

Our approach aims to help countries to build capacity in water modeling by providing fit for purpose tools. We do this by making use of global data sets, cloud-based technologies and interactive visualisations. We want to allow users to explore how the water balance changes under different scenarios. Our approach uses visualisations to show the changes in distribution of the basin water resources through time and space, as well as changes to water, food, energy security, the environment

The value proposition of our approach has three core aspects: global application, development of scenarios, and lower barrier to entry. This is where it differs from existing systems, such as those focused in different water management areas: for example, flood forecasting (Svatoň *et al.* 2017)(Castrogiovanni, La Loggia, and Noto 2005)(Al-Sabhan, Mulligan, and Blackburn 2003); or those providing spatial data portals (Blaschek *et al.* 2015).

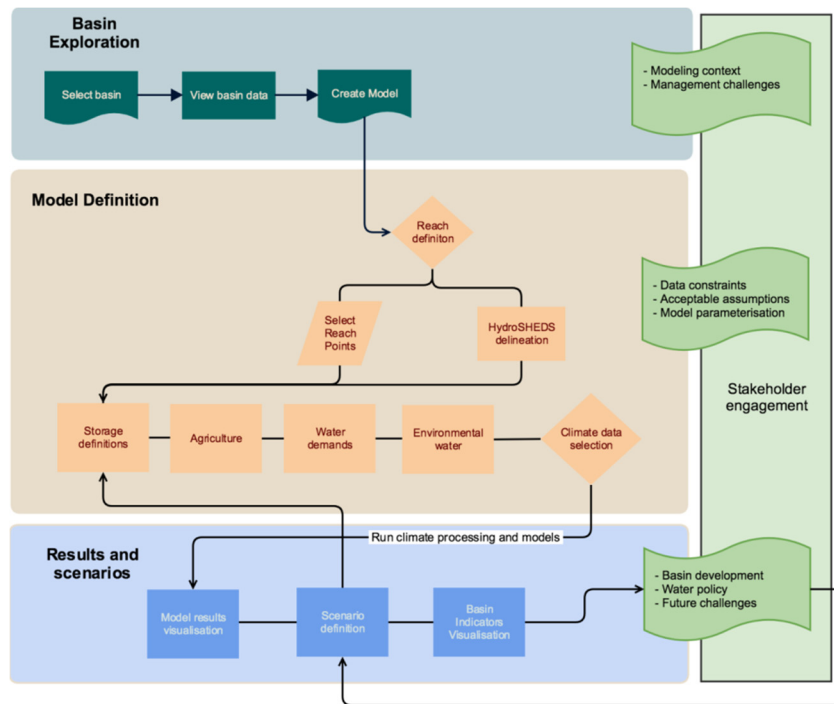
We specifically address the accessibility challenges of modelling systems identified by Al-Sabhan (Al-Sabhan, Mulligan, and Blackburn 2003): interfacing is difficult, knowledge of GIS is required, platform dependency, customisation of models is difficult, and limited accessibility of data. Using a combination of modernization of web-based visualisation, GIS systems, cloud computing and global data, Basin Futures provides a focused to help users work through basin assessments.

Basin Futures is aimed at providing an easy starting point for water assessments. The intention is to provide a stepping stone to other modelling platforms that can be used for further, more in depth modelling. Given the use of global and coarse data sets as a starting point, we recognise the need to consider the inherent uncertainties of the system (Beven 2007) – we are investigating how best to visualize results and communicate uncertainty to users.

This paper begins by detailing our approach to user interaction and the workflow at the core of the application. We then detail the underlying models, followed by the system implementation and data used, and finish with a discussion on current state and future work.

## 2. LOWERING THE BARRIER TO SCENARIO DEVELOPMENT

There are three phases in the overall Basin Futures interaction: basin exploration; model definition; and results and scenario exploration, as shown in Figure 1. Stakeholder engagement is a key principle of the Basin Futures value proposition – users interact with the system at multiple stages to ensure appropriate understanding is obtained as the model matures. The green side bar in Figure 1 shows the type of discussions that should be happening as users interact with the system. The intention is to have a simplified (80% of the value at 20% of the cost) system, where users and experts can interact and openly discuss the process of modeling and interpretation.



**Figure 1.** The phases of user interaction in Basin Futures

Another principle of Basin Futures is to make the best use of global and locally relevant data when it is needed by a user. These data are used as recommendations at each phase of development and to set reasonable defaults. Our intention is to significantly reduce the data processing required to develop a baseline model from which scenarios can be developed. We do this by making data and processing available in the form required for modelling analysis and tailored for the type of model under development.

### 3. MODEL OVERVIEW AND STRUCTURE

Basin Futures currently has two core modelling components: Rainfall Runoff via GR4J (Perrin, Michel, and Andréassian 2003) to generate the runoff in a basin, and a custom-built Reach Model Engine for undertaking core water management and transfer activities such as water storage, hydropower generation, routing of flows through the basin and crop modelling.

The GR4J model currently operates at the sub-basin scale and takes the daily rainfall and PET data sets generated at the sub-basin scale to estimate daily runoff time series. Users have the option of default parameter sets or user defined parameter sets for GR4J. Model development underway is focused upon disaggregating this model process further through more distributed runoff generation or user defined runoff input time series for more data rich models.

The water system modelling component unique to Basin Futures is the Reach Model Engine (Figure 2). In this model, a Reach is defined as a segment of river that transfers water from one sub-basin to another. Contained within the Reach are some of the key water management processes such as water storage, hydropower generation and extraction for crop production. The general model structure places the storage (if present in a reach) at the top of the reach. Inflows to the storage come from upstream reaches; outflows from a storage are governed by spillway and valve configurations defined by the user and generally related to the level-volume-area relationship for the storage. In addition to storage mass balance modelling, the Reach Model Engine routes flows within a reach (storage routing or piecewise linear routing) incorporates catchment inflow, reach losses, models water demand from crop models (water balance for multiple crop types simultaneously) and other demands (urban, industrial, environmental). Water demands are passed up a reach to the reach storage where water will be released (if available) incorporating the travel time and losses to attempt to fulfill the downstream demand. The model architecture prohibits the propagation of water orders beyond the reach to upstream reaches diminishing the need for iterative solvers to calculate the water demand cascade between all reaches. This model structure decision allows linear computation sequence and thereby much faster computation times.

Sustainability and security indicators have been linked to the Reach Model Engine. Hydropower production is a function of discharge, head and turbine efficiency which is also linked to basin population to derive an energy security indicator. Likewise, the economic value and estimated calorie and protein content of crop production is also linked to basin population and thereby food and economic security indicators. Water availability (timing and quantity) for crops, urban, rural and industrial demand are combined to derive water security indices.

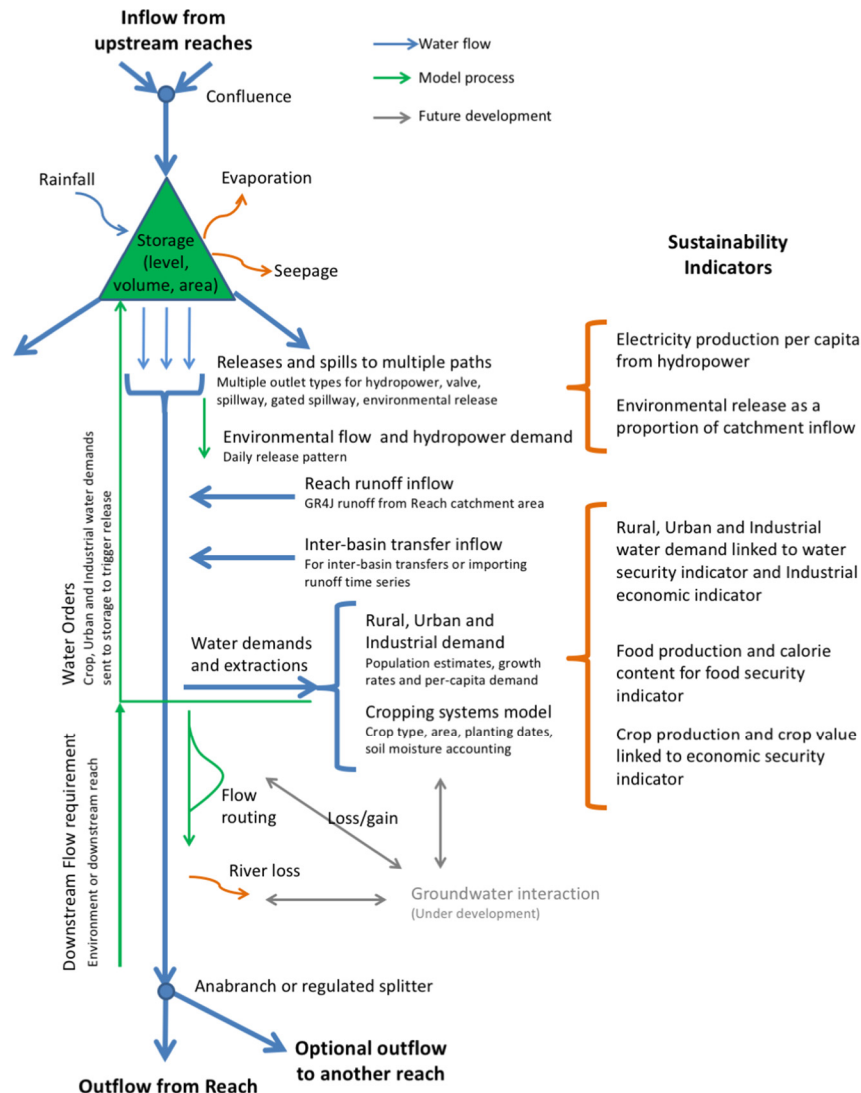
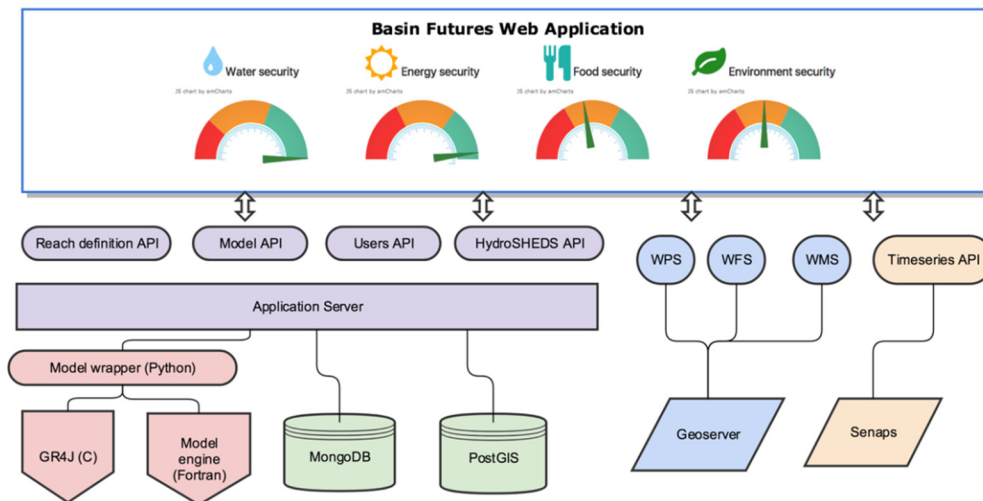


Figure 2. Model structure

#### 4. IMPLEMENTATION

We have designed the Basin Futures platform using three technical principles: Use service-based components; maximise component and library reuse by building on past experience; containerise all services using Docker to isolate dependencies and simplify deployment. The components used in the implementation of the platform are consistent with the ecosystem of tools recently described by Palomino, Muellerklein, and Kelly (2017). Specifically, we use open source, geospatial, environmental, and data processing tools and standards that have been widely adopted in research communities. An overall system design is shown in Figure 3. Each part of the system is described further in the following sections.



**Figure 3.** Overview of the system design

We have integrated a range of datasets. HydroBASINS (Lehner and Grill 2013) is used as a global database of catchment delineations at 12 scales. These data underpin the initial basin exploration and selection. HydroBASINS additionally provides flow direction grids that are used to dynamically delineate catchments based on user-selected locations. This is part of the reach definition process (see model description). HydroLAKES (Messenger *et al.* 2016) provides a global spatial database of water storages and their characteristics, e.g. volume and surface area.

A range of climate data is provided by the platform, including The Climate Hazards group Infrared Precipitation with Stations (CHIRPS) (Funk *et al.* 2015), The Climate Data Guide: GPCP (Daily): Global Precipitation Climatology Project (Pendergrass and National Center for Atmospheric Research Staff 2016), a CSIRO derived version of gridded PET from Princeton University (Princeton University 2015), and the Global Runoff Data Centre (GRDC) discharge data (The Global Runoff Data Centre 2017). We also provide the option for user-uploaded data where local data is available.

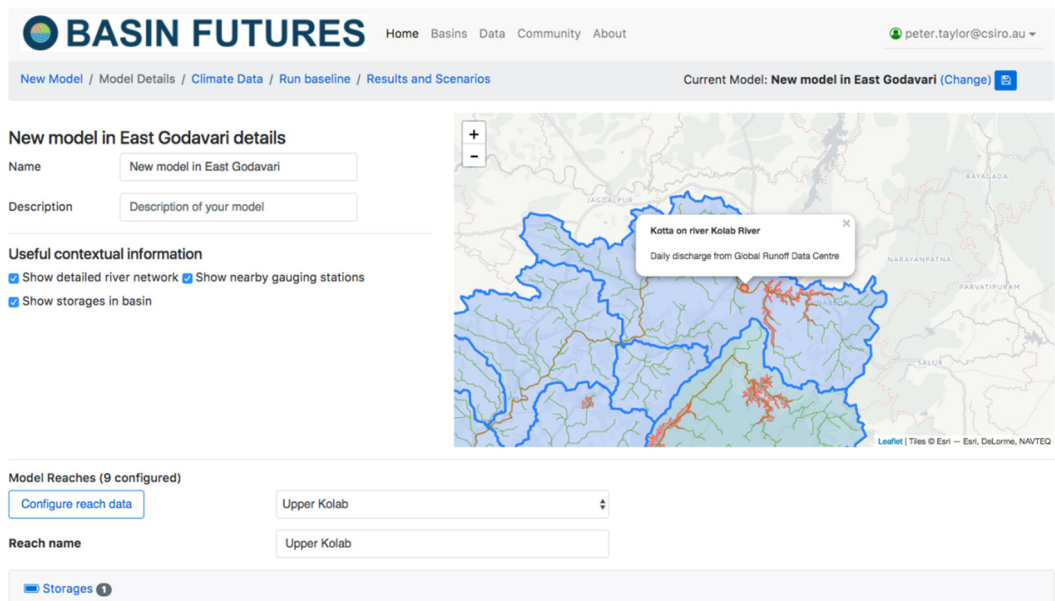
We use the GADM database of Global Administrative Areas (Robert Hijmans 2017) for a global database of administrative area names and types at varying levels. The Gridded Population of the World, Version 4 (Center For International Earth Science Information Network-CIESIN- Columbia 2016) is used to calculate population estimates on a sub-basin scale, and subsequently estimates of water use demand.

The application server is responsible for the interaction between the web application and backend services. It exposes an API for querying and accessing HydroSHEDS and other spatial data; reach definition processes; model definition and operation; user accounts; and climate data processing. It is implemented in Python using Python Flask, and Python Eve. Each of the APIs handles the interactions with specific internal and 3<sup>rd</sup> party libraries and data. For example, the model API handles the interaction with the Python Fortran model wrapper; the HydroSHEDS API interacts with the HydroSHEDS data in Postgis. The model is implemented in Fortran with a Python wrapper that marshals function calls using ctypes (Heller 2017).

Geoserver (Deoliveira 2008) is used to serve Web Mapping Service (WMS), Web Feature Service (WFS), and Web Processing Service (WPS) interfaces to the client application. It is linked to the PostGIS database, which contains the HydroSHEDS data. The web application makes use of the WFS and WPS services to query and extract relevant features from HydroSHEDS. The climate data are stored in NetCDF and processed using a Python library (Rahman 2017), which makes use of the rasterstats library (Perry 2017). Derived time-series from both climate data and model processing are stored in the Senaps platform (Coombe *et al.* 2017) through its API.

The web application has been developed using the Angular framework and the Typescript programming language. The user interface provides spatial context to users through relevant maps and spatial data. An example screenshot from the reach configuration stage is shown in

Figure 4.



**Figure 4.** Reach configuration interface

## 5. DISCUSSION AND FUTURE WORK

The current implementation of Basin Futures is an advanced proof of concept. We are developing a number of case studies in different countries to determine the effectiveness of the system. In addition to these case studies, development work is ongoing in a number of areas. We are in the process of integrating an approach to calibrate the model components (rainfall runoff, river routing, storages etc.), which will likely involve the use of existing tools (Doherty and others 1994) (Duan, Gupta, and Sorooshian 1993).

Existing work within CSIRO has developed a dam site assessment method (Petheram, Gallant, and Read 2017) that we are hoping to integrate with the platform. This would provide feasible locations for dam construction, which could then be used under development scenarios to understand potential benefits and effects on the downstream system.

Groundwater is a significant source of water for rural, urban and industrial demand in addition to crop production in developing countries. While the Reach model has some appropriate hooks built in for interaction between the surface water component and groundwater system, significant further work is required to implement a module to service these interactions.

Work is currently underway within CSIRO to investigate appropriate structures and data needs for a water quality module to interface with GR4J, the Reach module and the planned groundwater module. The Sustainability Indicators currently linked to the Reach module are currently under further refinement by CSIRO to more closely align with more widely accepted Sustainable Development Goals linked to water and basin scale processes.

Basin Futures aims to provide an integrated online tool to support rapid water resource assessments using global data, simple models, and visualization of scoping-stage development scenarios. It aims to lower the barrier to entry to the existing range of basin modelling tools available, by kick starting a basic user's understanding of water resource availability at basin scale. Basin Futures is not aimed at modelling regulated systems or supporting water trading or water sharing. Its key goal is to enable a cohort of users requiring early stage understanding of water resources to support planning outcomes. The outputs will also guide where more definitive answers, and data, will be required by the appropriate use of more powerful software.

## ACKNOWLEDGMENTS

The Basin Futures concept and development has been supported and funded through the CSIRO. We would like to acknowledge our partners in the development of Basin Futures, including, Flowmatters Pty Ltd, Catchment Research Pty Ltd, Council on Energy, Environment and Water, the Australian National University and International Commission on Irrigation and Drainage.

## REFERENCES

- Al-Sabhan, W, M Mulligan, and G. A Blackburn. (2003). “A Real-Time Hydrological Model for Flood Prediction Using GIS and the WWW.” *Computers, Environment and Urban Systems* 27 (1): 9–32. doi:10.1016/S0198-9715(01)00010-2.
- Beven, K. (2007). “Towards Integrated Environmental Models of Everywhere: Uncertainty, Data and Modelling as a Learning Process.” *Hydrol. Earth Syst. Sci.* 11 (1): 460–67. doi:10.5194/hess-11-460-2007.
- Mac Coombe, Paul Neumeyer, Joe Pasanen, Chris Peters, Chris Sharman, Peter Taylor. (2017). “Senaps: A Platform for Integrating Time-Series with Modelling Systems.” To appear at MODSIM 2017, Hobart Australia.
- Blaschek, Michael, Daniel Gerken, Ralf Ludwig, and Rainer Duttmann. (2015). “The CLIMB Geoportal - A Web-Based Dissemination and Documentation Platform for Hydrological Modelling Data.” In , 17:6483. <http://adsabs.harvard.edu/abs/2015EGUGA..17.6483B>.
- Castrogiovanni, E. M., G. La Loggia, and L. V. Noto. (2005). “Design Storm Prediction and Hydrologic Modeling Using a Web-GIS Approach on a Free-Software Platform.” *Atmospheric Research* 77 (September): 367–77. doi:10.1016/j.atmosres.2004.12.015.
- Center For International Earth Science Information Network-CIESIN- Columbia (last). (2016). “Gridded Population of the World, Version 4 (GPWv4): Population Count.” *Gridded Population of the World, Version 4 (GPWv4): Population Count*. doi:<https://doi.org/10.7927/h4x63jvc>.
- Dasgupta, Partha, and Karl-Göran Mäler. (2000). *The Environment and Emerging Development Issues*. Oxford University Press.
- Deoliveira, Justin. (2008). “GeoServer: Uniting the GeoWeb and Spatial Data Infrastructures.” In *Proceedings of the 10th International Conference for Spatial Data Infrastructure, St. Augustine, Trinidad*, 25–29.
- Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, et al. (2015). “The Climate Hazards Infrared Precipitation with Stations—a New Environmental Record for Monitoring Extremes.” *Scientific Data* 2 (December). doi:10.1038/sdata.2015.66.
- Heller, Thomas. (2017). *Ctypes: Create and Manipulate C Data Types in Python, Call Functions in Shared Libraries* (version 1.0.2). Windows, Linux, MacOS X, Solaris, FreeBSD. C, Python. Accessed August 2. <http://starship.python.net/crew/theller/ctypes/>.
- Lehner, Bernhard, and Günther Grill. (2013). “Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World’s Large River Systems.” *Hydrological Processes* 27 (15): 2171–86. doi:10.1002/hyp.9740.
- Messenger, Mathis Loïc, Bernhard Lehner, Günther Grill, Irena Nedeva, and Oliver Schmitt. (2016). “Estimating the Volume and Age of Water Stored in Global Lakes Using a Geo-Statistical Approach.” *Nature Communications* 7 (December): 13603. doi:10.1038/ncomms13603.
- Palomino, Jenny, Oliver C. Muellerklein, and Maggi Kelly. (2017). “A Review of the Emergent Ecosystem of Collaborative Geospatial Tools for Addressing Environmental Challenges.” *Computers, Environment and Urban Systems* 65 (September): 79–92. doi:10.1016/j.compenvurbsys.2017.05.003.
- Pendergrass, Angeline, and National Center for Atmospheric Research Staff. (2016). *The Climate Data Guide: GPCP (Daily): Global Precipitation Climatology Project*. <https://climatedataguide.ucar.edu/climate-data/gpcp-daily-global-precipitation-climatology-project>.
- Perrin, Charles, Claude Michel, and Vazken Andréassian. (2003). “Improvement of a Parsimonious Model for Streamflow Simulation.” *Journal of Hydrology* 279 (1): 275–89. doi:10.1016/S0022-1694(03)00225-7.
- Perry, Matthew. (2017). *Rasterstats: Summarize Geospatial Raster Datasets Based on Vector Geometries* (version 0.12.0). OS Independent. Python. Accessed July 31. <https://github.com/perrygeo/python-raster-stats>.
- Princeton University. (2015). “Princeton/NASA Earth System Data Record of Atmospheric Evaporative Demand.” [http://hydrology.princeton.edu/data.measures\\_pet.php](http://hydrology.princeton.edu/data.measures_pet.php).
- Rahman, Joel. (2017). *Climate-Utils*. <https://github.com/flowmatters/climate-utils>.
- Robert Hijmans. (2017). “Global Administrative Areas.” Accessed July 31. <http://www.gadm.org/>.
- Svatoň, Vaclav, Michal Podhoranyi, Radim Vavřík, Patrik Veteška, Daniela Szturcová, David Vojtek, Jan Martinovič, and Vít Vondrák. 2017. “Floreon+: A Web-Based Platform for Flood Prediction, Hydrologic Modelling and Dynamic Data Analysis.” In *Dynamics in GIScience*, 409–22. Lecture Notes in Geoinformation and Cartography. Springer, Cham. doi:10.1007/978-3-319-61297-3\_30.
- The Global Runoff Data Centre. (2017). *Global River Discharge*. 56068 Koblenz, Germany.