The Journal of Open Source Software

hddtools: Hydrological Data Discovery Tools

Claudia Vitolo[1]

1 European Centre for Medium-range Weather Forecasts

DOI: 10.21105/joss.00056

Software

- Review
- Repository
- Archive

Summary

The hddtools (Vitolo 2016) (**h**ydrological **d**ata **d**iscovery **tools**) is an R package (R Core Team 2016) designed to facilitate access to a variety of online open data sources relevant for hydrologists and, in general, environmental scientists and practitioners. This typically implies the download of a metadata catalogue, selection of information needed, formal request for dataset(s), de-compression, conversion, manual filtering and parsing. All those operation are made more efficient by re-usable functions.

Depending on the data license, functions can provide offline and/or online modes. When redistribution is allowed, for instance, a copy of the dataset is cached within the package and updated twice a year. This is the fastest option and also allows offline use of package's functions. When re-distribution is not allowed, only online mode is provided.

Datasets for which functions are provided include: the Global Runoff Data Center (GRDC), the Scottish Environment Protection Agency (SEPA), the Top-Down modelling Working Group (Data60UK and MOPEX), Met Office Hadley Centre Observation Data (HadUKP Data) and NASA's Tropical Rainfall Measuring Mission (TRMM).

This package follows a logic similar to other packages such as rdefra (Vitolo, Russell, and Tucker 2016) and rnrfa (Vitolo, Fry, and Buytaert 2015): sites are first identified through a catalogue (if available), data are imported via the station identification number, then data are visualised and/or used in analyses. The metadata related to the monitoring stations are accessible through the functions: `catalogueGRDC()`, `catalogueSEPA()`, `catalogueData60UK()` and `catalogueMOPEX()`. Time series data can be obtained using the functions: `tsGRDC()`, `tsSEPA()`, `tsData60UK()`, `tsMOPEX()` and `HadDAILY()`. Geospatial information can be retrieved using the functions: `KGClimateClass()` returning the Koppen-Greiger climate zone and `TRMM()` which retrieves global historical rainfall estimations.

The retrieved hydrological time series (e.g. using `tsData60UK()`) can be used to feed hydrological models such as fuse (Vitolo et al. 2012; Vitolo et al. 2016), topmodel (Buytaert 2011) and hydromad (F.T. Andrews, B.F.W. Croke, and Jakeman 2011; Andrews and Guillaume 2016).

For more details and examples, please refer to the help pages and vignette.

References

Andrews, Felix, and Joseph Guillaume. 2016. *Hydromad: Hydrological Model Assessment and Development*. http://hydromad.catchment.org/.

Buytaert, Wouter. 2011. *Topmodel: Implementation of the Hydrological Model Topmodel in R*. https://CRAN.R-project.org/package=topmodel.

F.T. Andrews, B.F.W. Croke, and A.J. Jakeman. 2011. "An Open Software Environ-

ment for Hydrological Model Assessment and Development." *Environmental Modelling & Software* 26 (10): 1171–85. doi:http://dx.doi.org/10.1016/j.envsoft.2011.04.006.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Vitolo, Claudia. 2016. *Hddtools: Hydrological Data Discovery Tools*. doi:10.5281/zenodo.61570.

Vitolo, Claudia, Matthew Fry, and Wouter Buytaert. 2015. *Rnrfa: UK National River Flow Archive Data from R*. https://CRAN.R-project.org/package=rnrfa.

Vitolo, Claudia, Andrew Russell, and Allan Tucker. 2016. "Rdefra: Interact with the UK AIR Pollution Database from DEFRA." *JOSS* 1 (4). The Open Journal. doi:10.21105/joss.00051.

Vitolo, Claudia, Peter Wells, Martin Dobias, and Wouter Buytaert. 2012. *Fuse: Framework for Understanding Structural Errors*. doi:http://dx.doi.org/10.5281/zenodo.14005.

———. 2016. "Fuse: An R Package for Ensemble Hydrological Modelling." *The Journal of Open Source Software* 1 (8). The Open Journal. doi:10.21105/joss.00052.