hddtools: Hydrological Data Discovery Tools

Claudia Vitolo
Brunel University London
5 September 2016

Paper DOI: http://dx.doi.org/10.21105/joss.00056
Software Repository: https://github.com/cvitolo/hddtools/
Software Archive: PENDING

Summary

The hddtools (Vitolo 2016) (hydrological data discovery tools) is an R package (R Core Team 2016) designed to facilitate non-programmatic access to a variety of online open data sources relevant for hydrologists and, more 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.

References


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