Karlsruhe Institute of Technology

Steinbuch Centre for Computing (SCC),
Institute of Water and River Basin Management (IWG)

V-FOR-WaTer – a virtual research environment to access and process environmental data

Marcus Strobl*, Elnaz Azmi, Sibylle K. Hassler, Mirko Mälicke, Jörg Meyer, Erwin Zehe
*email address: marcus.strobl@kit.edu

Motivation

Extent and diversity of environmental data continuously increase due to more and new sensors with higher spatial and temporal resolution and due to the growth and automation of observational networks. These data form the basis for a better understanding of ecological systems either by data driven methods or by comparisons of data with model predictions.

Still, a considerable amount of data are difficult to access or even stored on local data storage devices making it difficult or yet impossible to find, access and re-use it. In addition the data often lack a proper description (metadata) required for an interoperable analysis, hence they are barely useful for science. This results in very time consuming preparation and pre-processing of data, especially when datasets from different sources are combined.

Objectives

- **Quick and simple access** to hydrological data and tools
- **Quick pre-processing** of data from diverse data sources
- Shared tools for **reproducible data analysis**
- Opportunity to **easily upload data** to established data repositories **for publication**
- **Centralize hydrological data** from universities and state offices for a coordinated long-term monitoring
- **Security layer** to ensure that users can access only data for which they have access rights

Portal

- Login via B2ACCESS
- Filter menu to select data
- Map tools
- Tools for hydrological analyses
- Tools and data available through web portal and interfaces for common computing environments (*Matlab, Python and R*)
- Data export (cvs, shapefile)
- Metadata export (XML, ***ISO19115***)

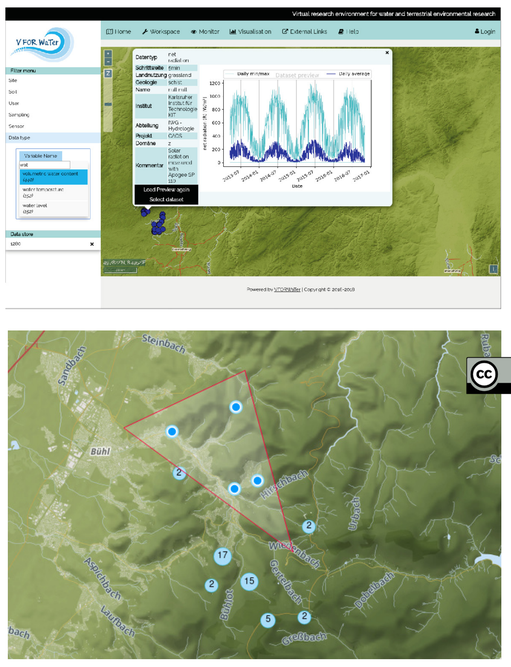

Fig. 1: Screenshots of V-FOR-WaTer web portal (upper part) and detail with the active 'filter on map' drawing tool.

Data

- Extensive metadata model to ensure usability of stored datasets
- Initial data in the portal come from the KIT Hydrology group (IWG) and LUBW (Landesanstalt für Umwelt, Baden-Württemberg)
- Development and testing of the portal is based on datasets from the CAOS research unit (Catchments as organised systems)
- Extensible database for user data and interface to open data repositories

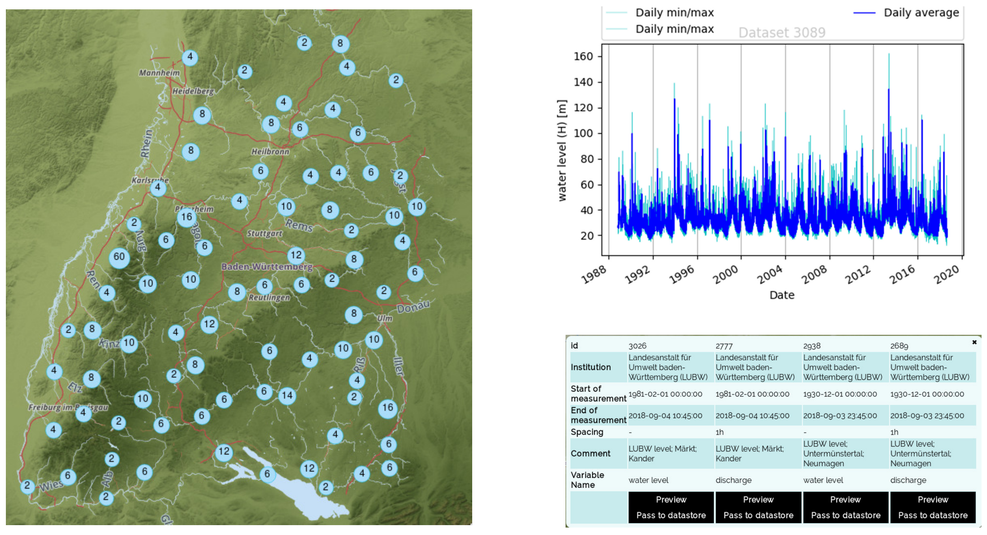

Fig. 3: Screenshots of map with clusters of data with similar sample location (left), data preview (upper right) and metadata preview (lower right).

Tools

- Geostatistics and analysis tools integrated as Web Processing Services, WPS (see scikit-gstat in Software).
- Based on the standalone python toolbox hydrobox (see Software)
- ***Results will come with references to enable users to cite data owners appropriately***
- *Graphical workflow tool will be integrated*

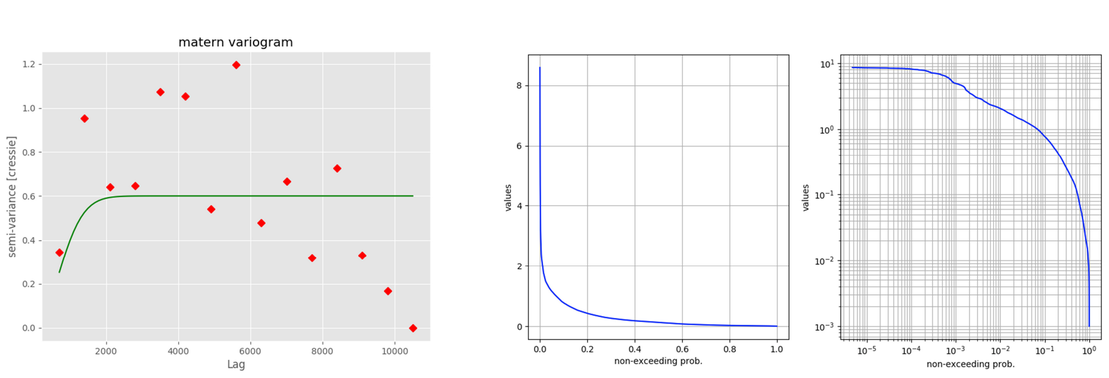

Fig. 2: Preliminary graphs from WPS toolbox for variogram analysis (left) and flow duration curves (right).

Software

- vforwater-portal: Portal of the virtual research environment **https://github.com/VForWaTer/vforwater-portal**
- pleasant: Django based skeleton of a web portal application with maps **https://github.com/VForWaTer/pleasant**
- hydrobox: hydroligical preprosessing and analysis toolbox **https://github.com/mmaelicke/hydrobox**
- scikit-gstat: geostatistics tools **https://github.com/mmaelicke/scikit-gstat**

Steinbuch Centre for Computing

Website:
http://www.vforwater.de/
Stay tuned!

Baden-Württemberg
MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST

HELMHOLTZ
RESEARCH FOR GRAND CHALLENGES