

# V-FOR-WaTer – A virtual research environment to discover and analyse environmental data

Jörg Meyer\*, Elnaz Azmi, Sibylle Hassler, Mirko Mälicke, Marcus Strobl, Erwin Zehe, Achim Streit  
\*email address: joerg.meyer2@kit.edu

## Motivation

The 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. The observed data form the basis for a better understanding of ecological systems either by data driven methods or by comparisons of the data with model predictions. However, a considerable amount of these data are difficult to access or even still stored on local data storage devices making it difficult or even impossible to find, access and re-use the data. In addition the data often lack a proper metadata description 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.

## 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

## 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

## Data

- Extensive metadata model to ensure usability of stored datasets
- Initial data in the portal come from the KIT Hydrology group 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

## Tools

- Geostatistics and analysis tools integrated as Web Processing Services (WPS)
  - Simple Variogram analysis (see Fig. 2 left)
  - Flow duration curve (see Fig. 2 right)
- Based on standalone python toolbox (see Software)
- Results will come with references to enable users to cite data owners appropriately
- Graphical workflow tool will be integrated

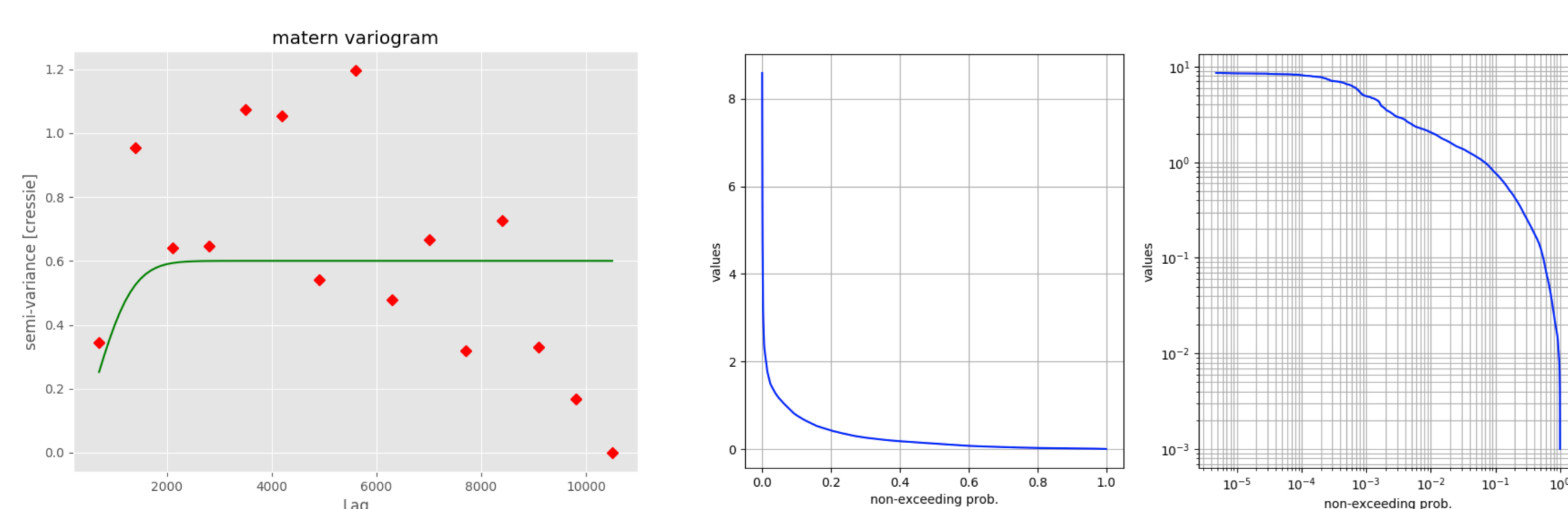


Fig. 2: Output graphs of WPS (variogram, flow duration curve)

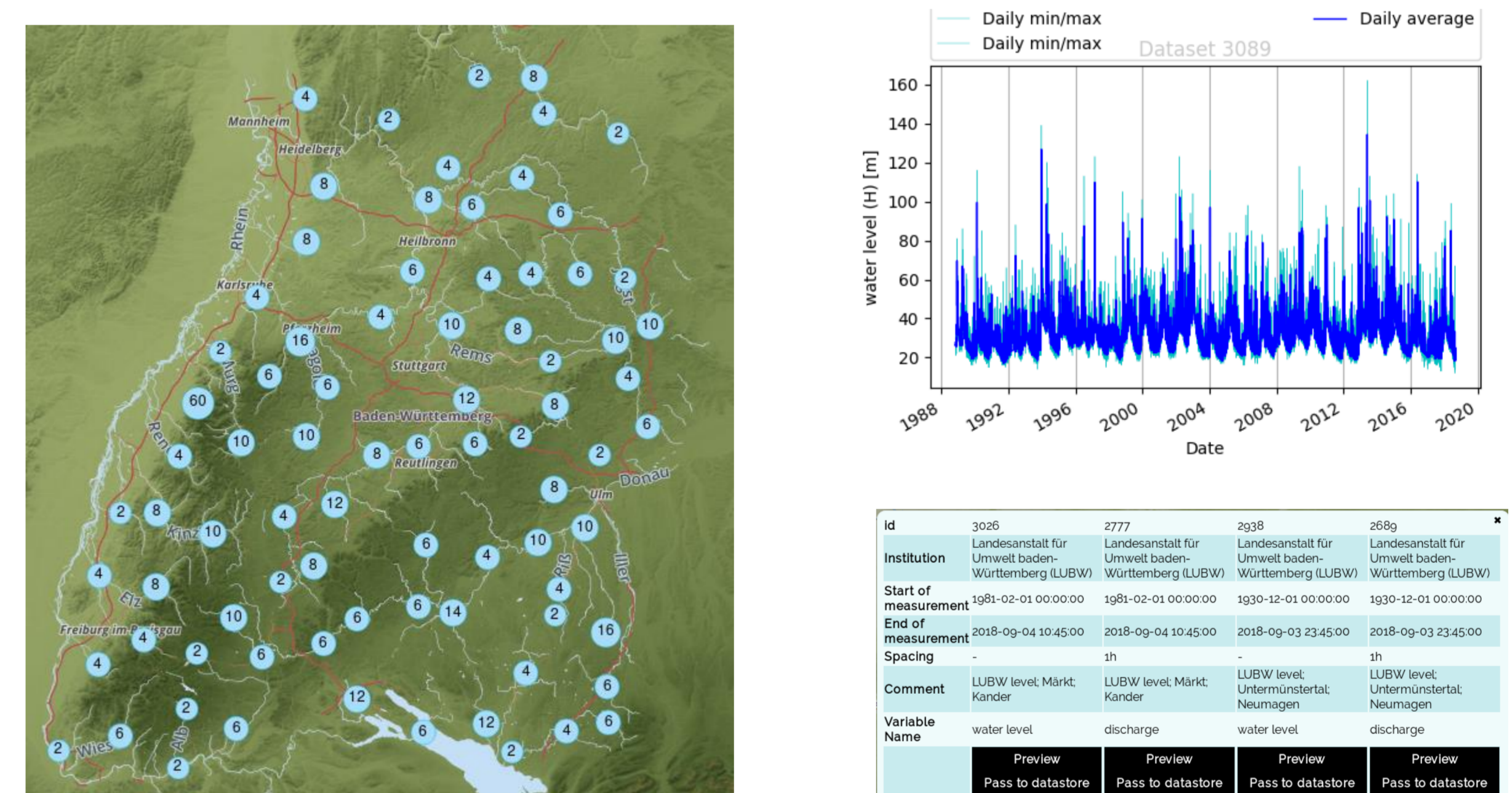


Fig. 3: Screenshot of map, metadata and data preview

## 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: hydrological preprocessing and analysis toolbox <https://github.com/mmaelicke/hydrobox>
- scikit-gstat: geostatistics tools <https://github.com/mmaelicke/scikit-gstat>

