Providing a user-friendly outlier analysis service implemented as open REST API

Doron Goldfarb*, Johannes Kobler, Johannes Peterseil

Umweltbundesamt GmbH, Vienna, Austria

* Corresponding author: doron.goldfarb@umweltbundesamt.at

EGU 2020, Virtual Meeting

This work was done in the context of the EOSC-HUB (grant agreement no 777536) project which has received funding from the European Union's Horizon2020 research and innovation programme.

Outline

- Outlier detection is crucial to reliable research
- Definition of „outlier“ varies by application
- Increasing online availability of environmental time-series data

- Approach: Provide public Web service for outlier analysis
  - Based on ensemble of outlier detection methods from different R-packages
  - Operating on data sourced from Sensor Observation Service (SOS) or from files from the cloud
  - Accessible via REST API
- Outlook: Refine initial EOSC-HUB prototype throughout eLTER+ project

Encapsulation behind REST API - Using OpenAPI/Swagger

Timeseries Outlier Analysis 0.1 OAS3

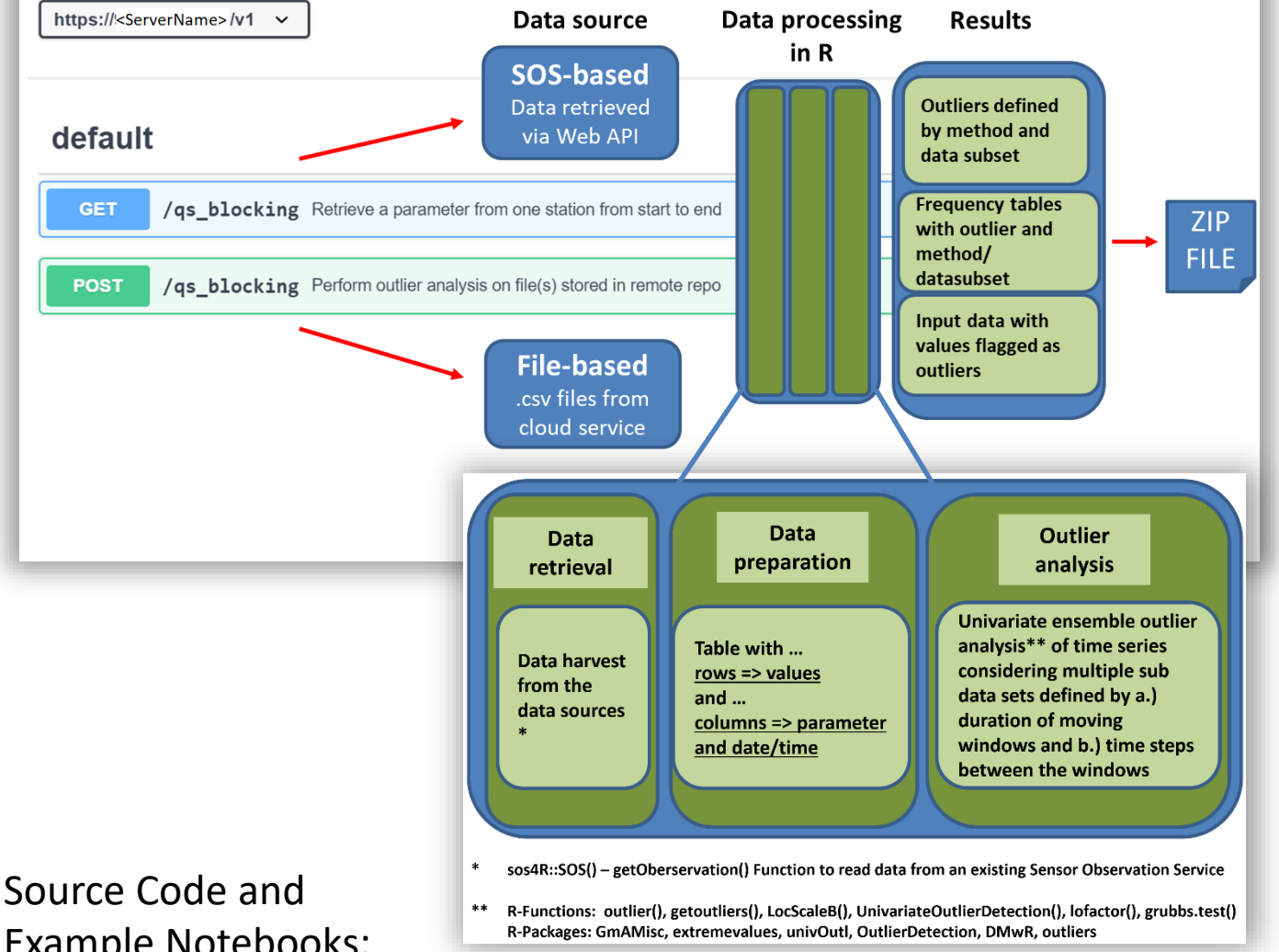
Servers

https://<ServerName>/v1

default

GET /qs_blocking Retrieve a parameter from one station from start to end

POST /qs_blocking Perform outlier analysis on file(s) stored in remote repo



* sos4R::SOS() – getOberservation() Function to read data from an existing Sensor Observation Service

** R-Functions: outlier(), getoutliers(), LocScaleB(), UnivariateOutlierDetection(), lofactor(), grubbs.test()
R-Packages: GmAMisc, extremevalues, univOutl, OutlierDetection, DMwR, outliers

Source Code and Example Notebooks:
https://github.com/d0rg0ld/OutlierDetection4EOSC