Observation analysis tool for the FREEWAT GIS environment for water resources management

Massimiliano Cannata[1], Jakob Neumann[1], Mirko Cardoso[1], Rudy Rossetto[2] , Laura Foglia[3],

[1] Institute of Earth Sciences, University of Applied Sciences and Arts of Southern Switzerland, Cannobio, Switzerland
[2] Scuola Superiore Sant'Anna, Pisa, Italy
[3] Technischen Universität Darmstadt, Darmstadt, Germany

Corresponding Author:
Massimiliano Cannata[1]

Email address: massimiliano.cannata@supsi.ch

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2127v2 | CC BY 4.0 Open Access | rec: 17 Aug 2016, publ: 17 Aug 2016

Abstract:

Time-series are an important aspect of environmental modelling, and are becoming more available through the requirements of the water framework directive as well as more important with the advancement of numerical simulation techniques and increased model complexity. For this reason, within the H2020 FREEWAT project, which aims at facilitating the adoption of modeling for water resource management, the integration of a tool for time-series analysis and processing has been foreseen. As a result the Observation Analysis Tool was developed to enable time-series visualisation, pre-processing of data for model development, and post-processing of model results. Observation Analysis Tool can act as a pre-processor for calibration observations, and will be expanded to incorporate its processing capabilities directly into the calibration process. The tool consists in an expandable Python library and in an interface integrated in the QGIS FREEWAT plug-in which include a large number of modelling capabilities, hydro-chemical data management tools and calibration capacity. The tool has been extensively used and tested in different european institutions, to collect a number of indications to drive the future development.

Introduction:

The Water Framework Directive 2000/60/CE (WFD) requirement for increased monitoring activities on water quantity and quality led to an increased collection and availability of data. Often the exploitation of the data up to now does not reflect information available within the data, which is frequently analysed with simple algorithms that do not provide insight into the dynamic systems (Rossetto et al. 2015). An increased exploitation of information provided by the data would allow implementation of the directives in more efficient water management. In the same paper (Rossetto et al. 2015) it was also suggested that decisions on water management and planning, that take into account the possibilities offered by state-of-the-art ICT, should move beyond the geographically lumped, 20 year average water budget management approaches to consider the fine spatial and temporal variability of hydrological systems and variables. FREEWAT (FREE and open source tools for WATer resource management) is an H2020 project financed by the European Comission, seeking to promote water management and planning by simplifying the application of the Water Framework Directive 2000/60/CE (WFD) and other related directives. The WFD recognises modelling as a relevant activity for a range of applications, of which FREEWAT will aim to facilitate: testing hypothesis on conceptual models, validating scenarios to be included in River Basin Management Plans, water resource evaluation and forecasting, impact assessment of large engineering works, and evaluation of the effectiveness of proposed remediation activities. In the water resource modelling process observations play a crucial role in creating effective models that adhere with reality, or at least are as close as possible to the simulation objective. For these reasons one of the FREEWAT modules is the Observation Analysis Tool (OAT): a time-series pre- and post-processor and visualization tool designed to support modelling.

Background and Motivation:

Any simulation model, and more specifically any hydrologic or groundwater model, aimed at providing Water Resource Managmenet (WRM) relevant information, requires quantitative data which defines the physical properties of the study area and describes its temporal evolution. This to define the system conditions, characteristics and resource exchanges. Identifying data needs, as well as collecting data is an essential part of any WRM related task. This tool for time-series analysis should therefore helps in handling this kind of data, which in spatio-temporal simulation models are: inputs for model conceptualization and implementation, outputs data for validation and observation data for sensitivity analysis, calibration and uncertainty analysis. In the context of FREEWAT and specifically groundwater modelling using MODFLOW-2005 (MODular three-dimensional finite-difference groundwater FLOW model) (Harbaugh, 2005), time-series data can be climate data such as rainfall and temperature, boundary conditions such as surface water stages or discharge measurements, or pumping and irrigation rates among others. The observations, which often take the form of time series data, are used for calibration within FREEWAT using UCODE 2014 (Poeter et al., 2014). UCODE is a model-independent code for sensitivity analysis, calibration and prediction with widespread use in groundwater model calibration (Hill & Tiedeman, 2006). Previously, only the Time Series PROCessor tool developed by Westenbroek et al. (2012) existed to pre-process time series data into observations that can be practically used in model calibration, specifically with the Parameter ESTimation (PEST) code (Doherty, 2016). Similarly to how TSPROC functions for PEST, OAT acts as a data preprocessor for calibration with UCODE, as well as a general purpose tool for time series analysis for model input and from model results. The difference between TSPROC and OAT reside in the software design and implementation language as well as the capacity to dynamically connect with online data services.

OAT is a module of the FREEWAT modelling environment, which is implemented as a plugin of the open source software QGIS (2016), aimed at simplifying the processing and manipulation of time-series data to ease its exploitation for water quality and quantity management. The idea is to achieves this in the FREEWAT framework specifically through: (i) the pre-processing of time-series data to drive the implementation and analysis of conceptual and numerical models for Water Resource Management (WRM) and (ii) the preparation of numerical model input data. Furthermore, enabling the visualisation and post-processing of model results may support management decisions. Raw digital time-series data available through monitoring networks, meteorological services or environmental authorities are always subject to pre-processing: specific to the temporal discretisation of the model, which in itself is subject to the end goal of modelling, and the possible spatial relationship to the area of interest. There's therefore a need of a tool designed to use time-series data processing to help in the preparation of model input data and in the statistical analysis of observations and model results, all within the same GIS platform.

Method:

OAT is a Python package which is integrated in the FREEWAT environment through an interface exposing its features to modellers and non programmer users. OAT library implements two main classes: the Sensor class that is designated to handle time-series data and metadata and the Method class which is designated to represent a processing method. The library applies the behavioral visitor pattern which allows the separation of an algorithm from the object on which it operates: thanks to this design pattern it is possible to add a new processing capability by simply extending the Method class without the need to modify the Sensor class. From a dependency point of view, OAT takes advantage of the PANDAS (McKinney, 2010), NUMPY and SCIPY (Van der Walt et. al. 2011) packages.

The Sensor class took inspiration from the Sensor Observation Service standard and it's istSOS implementation (Cannata et. al., 2015), with a specific simplification to ease its usage. Each Sensor object is characterized by a single time-series that is represented by a data section consisting in a PANDAS time-series and a location/metadata section.

The general structure and implemented usage of the library in FREEWAT is depicted in Figure 1.

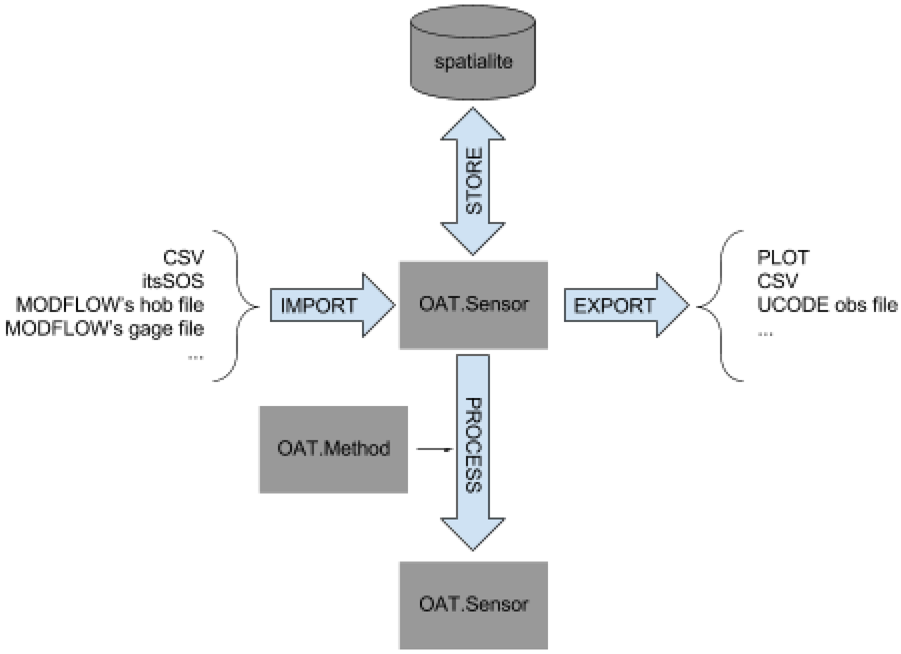

Figure 1 The OAT data retrieval, processing, storage and export workflow

Every OAT.Sensor object can be stored in a spatialite DB and re-loaded back in OAT.Sensor with its own data and metadata. For each sensor, the "metadata section" includes name, description, location (lat, lon, elev), unit of measure, observed property, coordinate system, time-zone, frequency, weight statistic and data availability (time interval). The "data section" contains time, data, name, and quality index, as well as a tag marking whether or not an individual observation in the series will be used. In addition to requesting sensor data from an istSOS server, OAT can retrieve data stored in local files or databases into the FREEWAT GIS environment for further use. Further input accepted format may be implemented in the future.

Most of the currently available OAT.Method objects are based on TSPROC processing capabilities. They enable the analysis of raw data (resampling, regularization, data interpolation, fitting, filling, validation and data quality assessment); decomposition or filtering of time-series (low, medium, high frequencies); aggregation and exceedance-time calculation, summary statistics and period statistics, integration and hydrograph separation among others. The results of a process is generally a new OAT.Sensor, so that processes can be concatenated, and the final resulting time-series can be saved in the FREEWAT model database or exported. As discussed in the package design, this library of processes is open and expandable to incorporate users needs, specifically as an aspect of the participatory approach adopted in FREEWAT development.

In preparation for model input the raw data processing capability of OAT are particular important. In fact, since often available data will not have the same temporal discretisation as the chosen modelling time, various filter and resampling or data fill methods are available to create new, but statistically sound, time-series to be used as model input.

OAT incorporates the post-processing of model results by creating sensors with appropriate temporal discretisation for the visualisation and further processing of model results as time-series. MODFLOW head and flow observations, listing file volumetric budget components (either as cumulative budgets or rates), and gage file components are currently supported formats. Time-series visualisation is achieved through the Python Matplotlib (Hunter, 2007) module. Various elements of the volumetric water budget for a MODFLOW model can be loaded as sensors and visualised and compared, to other elements of the budget or to other relevant time-series or observations.

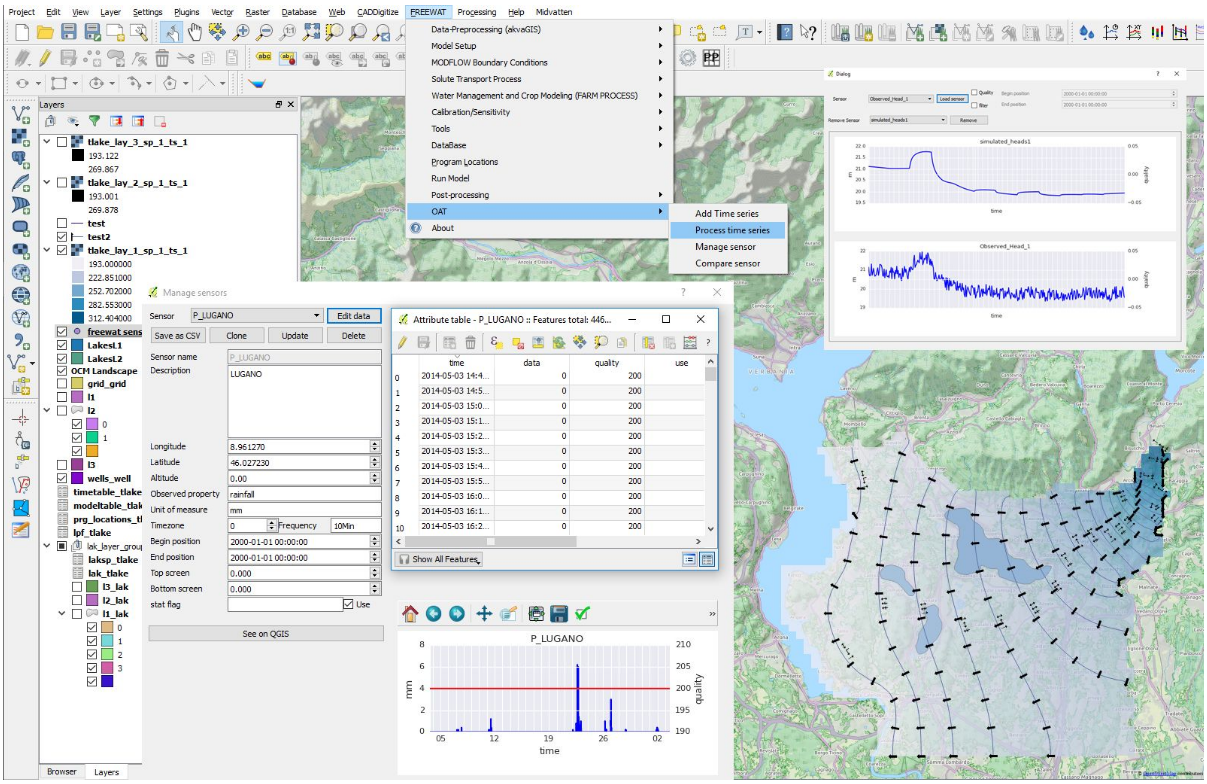

Figure 2 Screenshot of the OAT menu in a sample simulation within QGIS showing the “Manage sensor” and “Compare sensor” interfaces.

Groundwater head observations can be incorporated into the standard MODFLOW head observation package using OAT and simulated heads can be uploaded from the model results. Specifically for transient models with longer and possibly non continuous observation periods, OAT offers improved visualisation capabilities because the observations are separated by well and correctly plotted along the time axis so that is possible to true to time with variable spacing along the x-axis, rather than just sequentially, which allows their comparison with other time-series, e.g. nearby GW stresses.

Discussion & Conclusion:

OAT offers an important supplement to the FREEWAT environment through the dedicated analysis of time-series data, simplifying the use of the increased data available due to the WFD. Pre- and post-processing capabilities are designed specifically for WRM conceptual and numerical models within the current modelling environment, with easily expandable libraries to always offer relevant processing capabilities as FREEWAT incorporates additional modelling capabilities. Currently OAT

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2127v2 | CC BY 4.0 Open Access | rec: 17 Aug 2016, publ: 17 Aug 2016

can be used to plot and process MODFLOW inputs as well as results from the observation, listing and gage files.

Ultimately OAT will allow new calibration possibilities by acting as an observation pre-processor for optimisation. Model optimisation for processed time-series, such as baseflow or exceedance probabilities, will allow the modeller to focus the optimisation on a particular aspect of model performance based on its intended use, or on extreme situations like flood risk or drought situations. As FREEWAT incorporates additional capabilities for WRM simulation, OAT features will be able to act as a toolbox for many aspects of model optimisation and analysis.

The tool, at its current stage of development, has been extensively used and tested in thirteen courses in different european institutions. This testing phase allows the collection of a number of indications on the user interface as well as on the tool's capabilities as well as the creation of an open community. The outcome of the testing will be used to drive the future development. The detailed results and discussion of this testing will be presented in a forthcoming paper.

Acknowledgements

This paper is presented within the framework of the project FREEWAT, which has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement n. 642224.

References:

Cannata, M., Antonovic, M., Molinari, M., & Pozzoni, M. (2015). istSOS, a new sensor observation management system: software architecture and a real-case application for flood protection. *Geomatics, Natural Hazards and Risk*, *6*(8), 635-650.

Doherty, J. (2016). PEST Model-Independent Parameter Estimation User Manual Part I: PEST, SENSAN and Global Optimisers. *Watermark Numerical Computing*, 279.

European Community. (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Parliament*, *L327*(September 1996), 1–82. http://doi.org/10.1039/ap9842100196

Harbaugh, A. W. (2005). MODFLOW-2005 , The U . S . Geological Survey Modular Ground-Water Model — the Ground-Water Flow Process. *U.S. Geological Survey Techniques and Methods*, 253.

Hill, M. C., & Tiedeman, C. R. (2006). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.

Hunter, J.D. (2007) Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering*, 9, 90-95, DOI:10.1109/MCSE.2007.55

McKinney, W. (2010). Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 51-56

Poeter, E.P., Hill, M.C., Lu, D., Tiedeman, C.R., & Mehl, S. (2014). UCODE_2014, with new capabilities to define parameters unique to predictions, calculate weights using simulated values, estimate parameters with SVD, evaluate uncertainty with MCMC, and More: *Integrated Groundwater Modeling Center Report Number: GWMI 2014-02*.

QGIS Development Team, (2016). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Rossetto, R., Borsi, I., & Foglia, L. (2015). FREEWAT: FREE and open source software tools for WATer resource management. *Rendiconti Online Della Società Geologica Italiana*, *35*, 252–255. http://doi.org/10.3301/ROL.2015.113

Van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation, *Computing in Science & Engineering*, 13, 22-30, DOI:10.1109/MCSE.2011.37

Westenbroek, S. M., Doherty, J. E., Walker, J. F., Kelson, V. A., Hunt, R. J., & Cera, T. B. (2012). Approaches in highly parameterized inversion: TSPROC, a general time-series processor to assist in model calibration and result summarization. *U.S. Geological Survey Techniques and Methods*, (7-C7), 112.