The Observation Analysis Tool: a free and open source tool for time series analysis for groundwater modelling

M. Cannata*
J. Neumann*

* Istituto scienze della Terra, Scuola universitaria professionale della Svizzera italiana, Canobbio

OAT is a module of the FREEWAT plugin for QGIS, developed to handle time-series data from various sources and formats and to ease the incorporation of environmental monitoring data into GIS applications, specifically water resources management and groundwater modelling. OAT has been successfully applied to the creation of a hydrogeological model of the Lake Lugano watershed, in order to pre-process data from various sources, including online monitoring services, and to post-process model results for calibration and validation. The OAT represents a user-friendly method to incorporate time-series data into GIS and to process that data further within the GIS environment. By incorporating an easily expandable process library, OAT is very adaptable to modellers needs and will grow to incorporate new methods as the FREEWAT community grows. OAT may be used beyond its current application to any spatio-temporal GIS applications.
Keywords*: time-series; QGIS; environmental modelling; temporal GIS*

Observation Analysis Tool: uno strumento free ed open source per l'analisi di serie temporali a supporto della modellazione idrogeologica. *OAT è un modulo del plugin di QGIS denominato FREEWAT. Tale modulo è stato sviluppato per consentire la gestione e l'analisi di serie temporali in ambiente GIS così da agevolare l'utilizzo di dati provenienti da sistemi di monitoraggio ambientali nella gestione delle risorse naturali ed in particolare nella modellazione idrogeologica. OAT è stato utilizzato con successo nella creazione di un modello idrogeologico del bacino del Lago di Lugano, specificatamente per il pre-processamento di dati provenienti da diversi formati e diverse sorgenti inclusi servizi Web ed il post-processamento dei risultati del modello e della calibrazione. Nell'ambito di quest'applicazione, OAT ha notevolmente facilitato il compito del modellista non solo abilitando la gestione di serie-temporali insieme alle altre informazioni su cui il modello si basa, ma anche offrendo funzioni di processamento che consentono l'integrazione diretta di questa tipologia di dato nella modellazione integrata in ambiente QGIS. Basandosi su una libreria facilmente espandibile, OAT si adatta agevolmente ai bisogni del modellista così da includere, al crescere della comunità, nuovi metodi di analisi. Sebbene sviluppata ed incorporata in FREEWAT, OAT è utilizzabile in applicazioni spazio-temporali di diversi contesti.*
Parole chiave*: serie temporali; QGIS; modellazione ambientale; GIS temporali.*

1. Introduction

1.1. Background

The Observation Analysis Tool (OAT) was developed as one of the modules of the FREEWAT (FREE and open source tools for WATer resource management) plugin for QGIS, as part of the H2020 project of the same name, seeking to promote water management and planning by simplifying the application of the Water Framework Directive 2000/60/CE (WFD) and other related directives. OAT aims to simplify the processing and manipulation of time-series data to ease its exploitation for water quality and quantity management. The current exploitation of the increased data available through the mandated monitoring of the WFD does not sufficiently reflect the information available within said data. An increased exploitation of information provided by the data would allow implementation of the directives in more efficient water management. Rossetto *et al.*, 2015) suggest 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 aims to facilitate several of the range of modelling activities that the WFD recognises as relevant within water resource modelling. Observations from environmental monitoring play a crucial role in creating effective models that adhere to reality. Thus the OAT is a time-series pre- and post-processor and visualization tool designed to support modelling within FREEWAT. Within the organisation of the FREEWAT plugin, OAT falls outside the core Hydrological module and fills the function of a tool for general operations to prepare input data and post-processing functionalities; for time series data generally, as well as online sensors specifically.

1.2. Aim

The groundwater models produced through FREEWAT using MODFLOW-2005 (MODular three-dimensional finite-difference groundwater FLOW model) (Harbaugh, 2005), aimed at providing Water Resource Management (WRM) relevant information, require quantitative data which define the physical properties of the study area and describe its temporal evolution. Time-series data may include climate data such as rainfall and temperature, boundary conditions such as surface water stages or discharge measurements, pumping and irrigation rates, or groundwater observations. 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. Therefore, there is a need for 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. Within FREEWAT OAT handles and simplifies the application this kind of data for the pre-processing of time-series data to drive the implementation and analysis of conceptual and numerical models. Additionally, it enables the visualisation and post-processing of model results, may support management decisions, as well as providing a direct link to the observation creation portion of the sensitivity analysis/calibration FREEWAT module.

2. Method

OAT is a Python package which is integrated in the FREEWAT environment through a user interface. 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 (Cannata *et al.* 2016). The library applies the behavioural "visitor pattern" (Palsberg and Jay, 1998) which permits to separate the algorithm (Method object) from the element it applies to (Sensor object). This approach greatly facilitates the expansion of the available algorithms for time series analysis by just adding a new child object of the Method class.

The existing packages OAT relies on are: PANDAS (McKinney, 2010), NUMPY and SCIPY (Van der Walt *et al.* 2011). The Sensor class took inspiration from the Sensor Observation Service standard, as it is re-

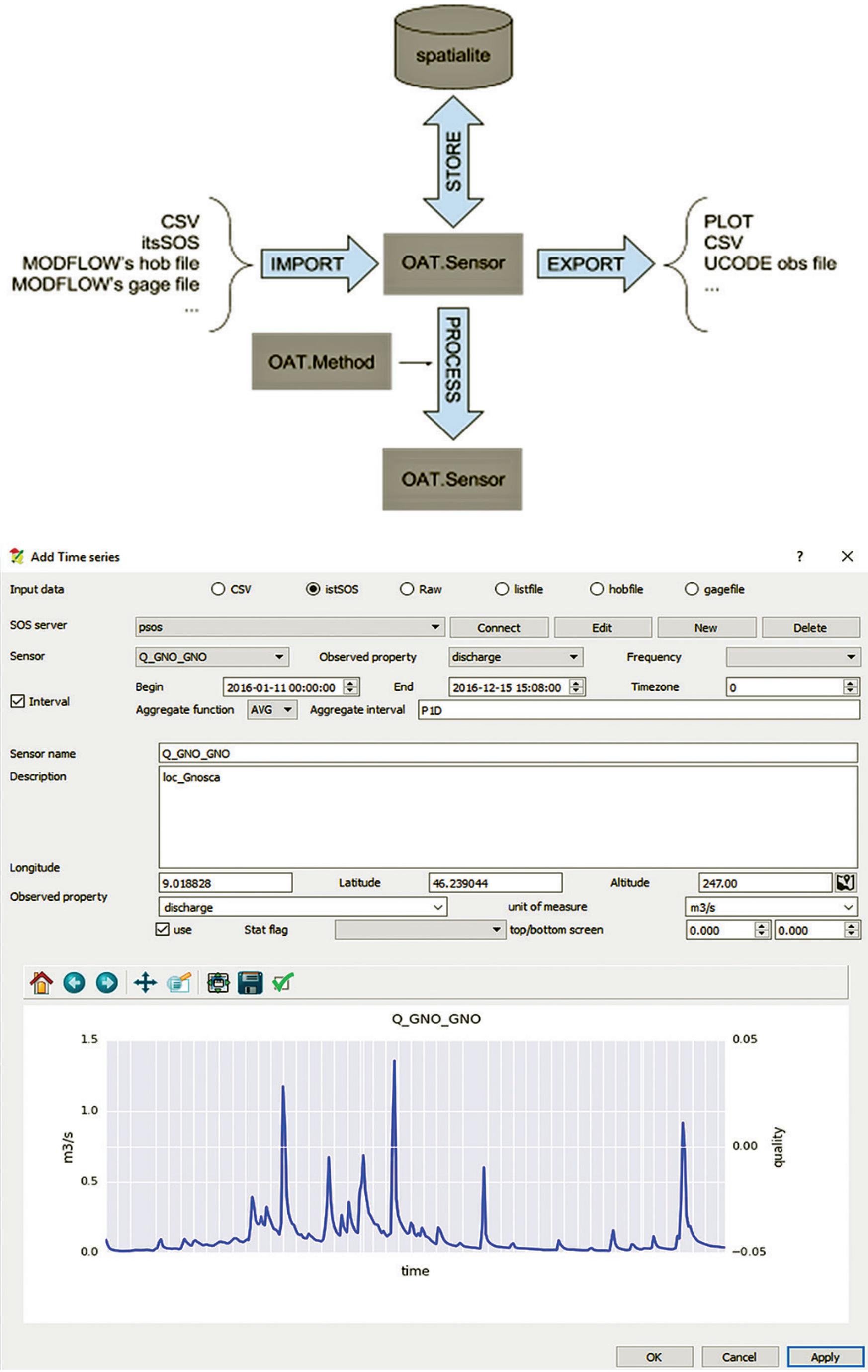

Fig. 1. Above: The OAT data retrieval, processing, storage and export workflow. Below: the OAT interface to add a sensor from an istSOS server. The first row, "Input Data" shows the other possible sources of sensor data.
dida italiano

ferred as "SOS" hereafter, and its istSOS implementation (Cannata *et al.*, 2015), with a specific simplification to ease its usage. istSOS (Istituto scienze della Terra Sensor Observation Service) has been used since 2009 as a simple implementation of the SOS for the management, provision and integration of hydro-meteorological data collected in Canton Ticino. It has also been used with the Verbano Lake Early Warning System for lake level forecasting and flooding hazard assessment (Cannata *et al.*, 2015). istSOS has also been used to develop a smart system to sustain optimal water usage in irrigation within the European project ENORASIS (Cannata & Antonovic, 2015).

Each *OAT.Sensor* object is characterized by a data section which contains time, data, name, and quality index; and a metadata section, which includes sensor name, description, location (latitude, longitude, elevation), unit of measure, observed property, coordinate system, time-zone, frequency, weight statistic and data availability.

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.

Most of the currently available *OAT.Method* objects are based on Time Series PROCessor (TSPROC) tool developed by Westenbroek *et al.*, 2012). TSPROC enables the analysis of raw data decomposition, filtering of time-series, aggregation, exceedance-time calculation, summary and period statistics, integration and hydrograph separation among others. The difference between TSPROC and OAT resides in the software design and implementation language as well as the capacity to dynamically connect with online data services. Additionally, the *OAT.Method* library is easily expandable and additional methods, such as the calculation of evaporation from temperature using the Hargreaves-Samani equation (Allen *et al.* 1998) among others, have already been implemented.

3. Application

3.1. Lake Lugano Case study

OAT has been heavily used and demonstrated through the hydrogeological modelling of the Lugano Lake watershed and aquifers as a FREEWAT case study by SUPSI-IST (paper forthcoming). Temperature, precipitation, stream discharge, stream height, humidity, atmospheric pressure, and lake level, among others, are monitored through istSOS in the model area. Figure 3 of the Lake and study area shows the location of a selection of monitoring stations and measurement points used for model creation as OAT sensors. The data used for the OAT sensors originate from monitoring points administered by SUPSI-IST as well as the state administered monitoring stations.

Raw data processing capability of OAT are particular important for the preparation of model input. Since available data for most modelling applications do not have the temporal discretisation chosen for the model, various filter and resampling or data fill methods are available to create new, but statistically sound, time-series to be used as model input. In the case of the Lake Lugano case study model, a temporal discretisation of 7-day stress periods was chosen, with daily time steps, as the system is greatly dependant on highly dynamic surface waters. In total 52 stress periods were simulated, just under a year, using data from the first 364 days of 2012. The time series used to generate input for model stresses and boundary conditions were all resampled to model stress periods and aggregated using either the mean or the summation, e.g. for stream stages or precipitation respectively.

3.2. Application of OAT capabilities for model creation

Precipitation data from the sensors was imported from istSOS and

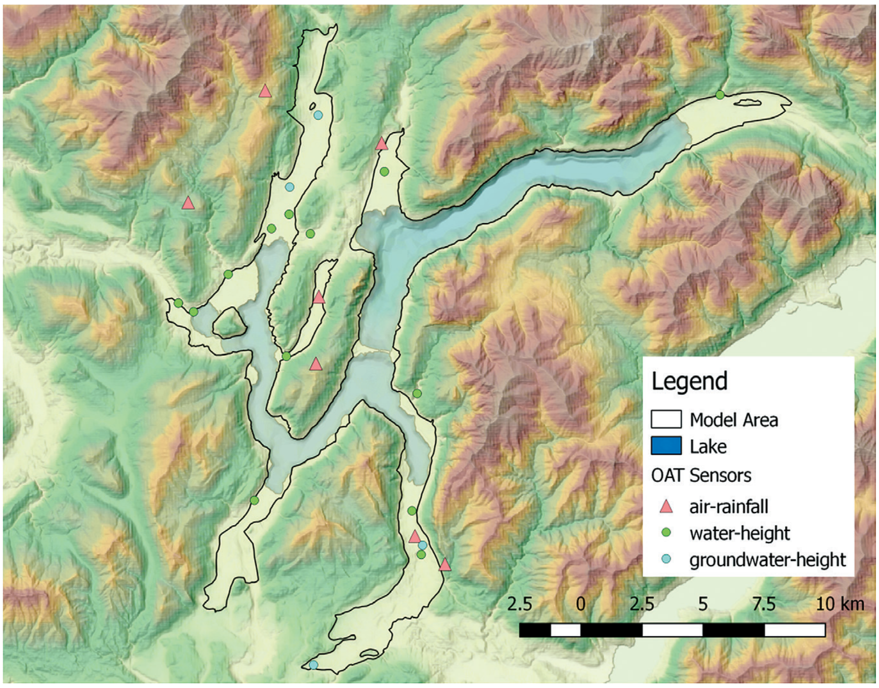

Fig. 2. The Lake Lugano case study area, as well as a selection of the OAT sensors used. *dida italiano*

used to calculate the recharge applied to the model cells using the MODFLOW recharge package, mountain front recharge using the MODFLOW well package, as well as direct precipitation to the lake surface. The data was aggregated and resampled, then used in conjunction with multipliers based on CORINE land use polygons (EEA, 2012) and flow accumulation rasters calculated from the digital elevation model to determine the rate of recharge each model cell.

Temperature data was used with a newly added process to calculate evaporation through the Hargreaves-Samani equation and was used for the evaluation of direct evaporation from the lakes. Stream and river stage gages were used to determine the weekly variations of water height. These weekly variations were applied to the known water heights at the start of each river segment, where it enters the model domain. Since no river gauge is located at the end of the rivers, the lake stage has been used as representative of this value. The stage of each individual river cell was then automatically determined by FREEWAT through linear interpolation along the length of each river segment.

3.3. Application of OAT for model calibration and validation

FREEWAT supports model calibration using UCODE (Poeter, 1999) and the MODFLOW observation packages which are used to specify observations of head or river flow. Additionally, other sensor data such as simulated lake stages can be used for model validation and calibration at a later step. The case study uses OAT for the automatic conversion of sensor time series data from groundwater observation wells stored as sensors, monitored by SUPSI-IST, to MODFLOW head observations layers.

Within FREEWAT, head and river observations can be created by the user by importing a specifically predefined CSV file in QGIS for each well that need to be associated with name of a spatial feature. OAT greatly improves this feature by being able to convert raw time series data to model time resolution from sensor objects. Within the interface to create head observation layers, a simple checkbox was implemented, which can be used to activate the link to a selected OAT sensor. Provided the sensor has the correct attributes, and metadata, it will be used for the head observations. The observations in the OAT sensor time-series are applied to the correct stress period and time step and model layer automatically, applying weights specified in the sensor metadata. The same can be done to create flow observation layers. In this case, the difference between two sensors is used, i.e. the difference between two streamflow gages representing the water gained or lost by a river segment between the two gages.

Specifically, for transient models with longer and possibly non-continuous observation periods, OAT offers improved visualisation capabilities because the observations are separated by observation well and correctly plotted along the time axis This is done 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 groundwater stresses. In this way, the influence of boundary conditions to observation wells can be investigated graphically.

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

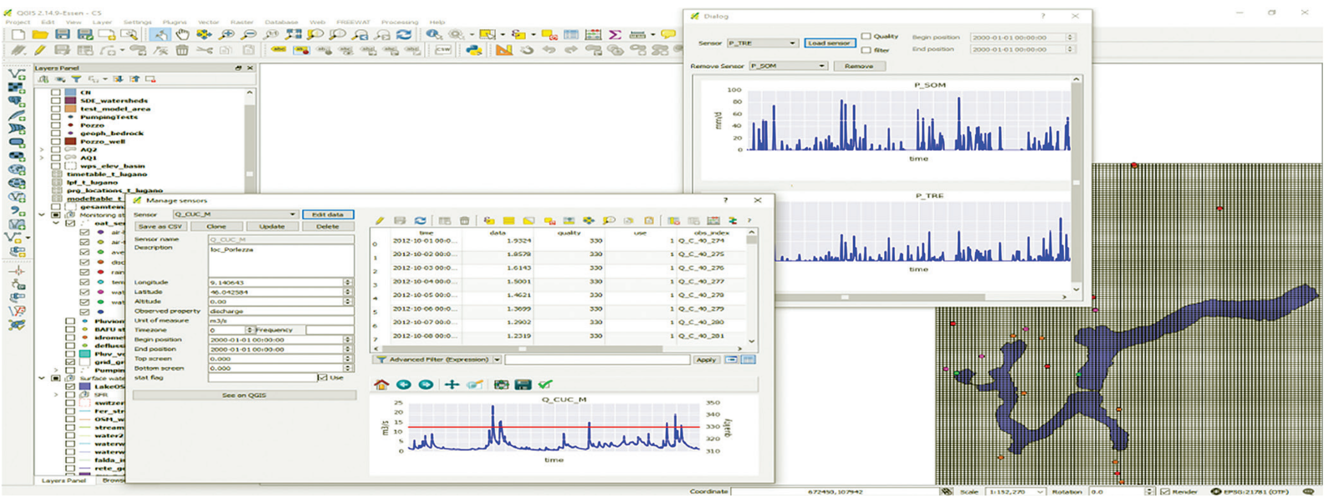

Fig. 3. Screenshot of the OAT menu in a sample simulation of the Lake Lugano case study within QGIS showing the Manage Sensor and Compare Sensor interfaces, with the case study grid in the background.
dida italiano

file volumetric budget components (either as cumulative budgets or rates), and gage file components are currently supported formats. 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.

4. 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 water resources management conceptual and numerical models within the current modelling environment. 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. Currently OAT can be used to plot and process MODFLOW inputs as well as results from the observation, listing and gage files. The usage of OAT to facilitate the modelling process has been demonstrated in the Lake Lugano case study model, and will further be demonstrated through the case studies of various partner institutions within the FREEWAT consortium. This testing phase allows the collection of a number of indications on the user interface as well as on the tool's, fostering at the same time the creation of an open community of users. The feedback from the community, based on their experiences applying OAT in their case studies, will be used to drive the future development and to add further processes to the OAT library. As its community of users grows, OAT capabilities may also incorporate other processes, beyond those used for groundwater modelling, to facilitate the usage and application of temporal-spatial GIS data in any field.

While not being innovative in term of time series analysis, OAT brings new capabilities for QGIS and FREEWAT users which were not possible before. Uploading data from the web or other sources, elaborating and using them within an integrated modelling environment simplifies and speeds-up the modeller works. As a consequence, it facilitates the evaluation of new hypothesis and scenarios by quickly adding, removing or changing observations used to define stress and boundary conditions. Also automatic calibration is now easier. In fact, using OAT, the tedious, error prone and time consuming procedure of manually creating observations files is now automatic. Finally, the applied design permits the easy implementation of new algorithms that could be used also outside FREEWAT.

5. Bibliography

Allen, R.G., Pereira, L.S., Raes D., & Smith, M., 1998. *Crop evapotranspiration-Guidelines for computing crop water requirements*. FAO Irrigation and drainage paper 56. FAO, Rome.

Cannata, M., Antonovic, M., 2010. *istSOS: investigation of the sensor observation service*. WebMGS 1st international workshop on pervasive web mapping, geoprocessing and services. Como, Italy. pp. 26-27.

Cannata, M., Antonovic, M., 2015. *Sensor Observations Service for Environmentally Optimizing Irrigation: istSOS within the ENORASIS Example*. International Journal of Geoinformatics. 11(3).

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), pp. 635-650.

Cannata, M., Neumann, J., Cardoso, M., Rossetto, R., Foglia, L., 2016. *Observation analysis tool for the FREEWAT GIS environment for water resources management*. PeerJ Preprints, 4, e2127v2.

Cocca, P.A., Doherty, J., Kittle Jr, J.L., 2003. *Hydrologic Calibration Strategies for the HSPF Watershed Model: Identifying Effective Objective Functions for Use with the Parameter Estimation (PEST) Program*. Proceedings World Water & Environmental Resources Congress 2003.

Doherty, J., 2004, *PEST: Model-independent parameter estimation, User manual: 5th ed.* Watermark Numer. Comput, Brisbane, Queensland, Australia.

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), pp. 1-82.

EEA, 2012. *CORINE Land Cover Project*. European Commission, Copenhagen.

Hunter, J.D., 2007. *Matplotlib: A 2D Graphics Environment*. Computing in Science & Engineering, 9, pp. 90-95.

Kim, Y.J., Kim, H.D., & Jeon, J.H., 2014. *Characteristics of water budget components in paddy rice field under the asian monsoon climate: Application of hspf-paddy model*. Water, 6(7), pp. 2041-2055.

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

Palsberg, J. & Jay, C.B., 1998. *The essence of the visitor pattern. In Computer Software and Applications Conference, 1998*. COMPSAC'98. Proceedings. The Twenty-Second

Annual International (pp. 9-15). IEEE.

Poeter, E.P., Hill, M.C., 1999. *UCODE, a computer code for universal inverse modeling*. Computers & Geosciences 25.4,pp. 457-462.

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, pp. 252-255.

Searcy J.K., 1959. *Flow-duration curves, Manual of hydrology, pt. 2, Low-flow techniques*. U.S. Geological Survey Water-Supply Paper 1542-A.

Van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. *The NumPy Array: A Structure for Efficient Numerical Computation*. Computing in Science & Engineering, 13, pp. 22-30.

Walker, J.F., Hunt, R.J., & Doherty, J., 2007. *Effective Use of Time-Series Data to Calibrate a Coupled Ground Water/Surface Water Model in a Small Headwater Watershed*, Northern Wisconsin. In AGU Fall Meeting Abstracts (Vol. 1, p. 0700).

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).