

# OWS4SWAT: Publishing and Sharing SWAT Outputs with OGC standards

Gregory Giuliani<sup>1,2</sup>, Kazi Rahman<sup>1</sup>, Nicolas Ray<sup>1,2</sup>, Anthony Lehmann<sup>1</sup>

<sup>1</sup>Institute for Environmental Sciences, enviroSPACE University of Geneva 1227 Carouge, Switzerland

<sup>2</sup>United Nations Environment Programme Global Resource Information Database, 1211 Châtelaine, Switzerland  
gregory.giuliani@unige.ch

**Abstract**—The Soil and Water Assessment Tool (SWAT) is a widely used hydrological model that produces several useful outputs (e.g. evapotranspiration, soil moisture, aquifer recharge, river discharge) as text files. Currently, visualizing and publishing SWAT outputs as geospatial data requires a lot of time and repetitive processing steps. Moreover, data used and produced are often not interoperable and restricted to software like ArcGIS or MapWindow. Consequently, integrating SWAT outputs with other datasets and/or models is difficult. To solve these issues, we propose an innovative, scalable and interoperable framework allowing (1) the automatization of post-processing tasks using orchestrated Web Processing Services (WPS) and (2) the publishing of SWAT outputs using interoperable data services (e.g. Web Feature Service, Web Map Service). The proposed framework simplifies map/data production and facilitates exchange/integration of hydrological data with other sources.

**Keywords**—enviroGRIDS; SWAT; Water; Interoperability; WPS; OGC; GEOSS; Black Sea

## I. INTRODUCTION

Humans are exerting significant impacts on the global water system [1] through activities such as the accelerated melting of snow and ice in alpine zones, the removal of trees that lead to increased runoff, reduced transpiration and impacts on the water table and its salinity, the draining of wetlands, the irrigation for agriculture, the alteration of flow through dams, the transfer of water between catchments, and finally the pollutions from industrial, agricultural and domestic sources. To better understand these modifications and impacts, water science research needs to follow a holistic research approach in order to effectively inform policy for sustainable water management about the dynamics of water in the context of global needs.

However, it is recognized that many policy-relevant research areas are still facing the problem of readily and timely access and exchange of data. Access and availability of reliable time-series on environmental, statistical, and socio-economical data is still a major barrier to effective and efficient informed policy-making [1]. Additionally, there are currently also gaps in term of knowledge when analyzing different water cycles: water scarcity (e.g. droughts), water abundance (e.g. floods), water quality (including sediment loads evaluation), water use and renewability, interactions between extremes (e.g. interconnections between drought and flood distribution), and ecosystem services maintenance.

Effective and efficient water management requires coordination of actions, the first one being the access and provision of reliable data and information (e.g., state of the resources, changes, pressures) and second the capacities to interpret correctly and meaningfully these information [2, 3]. Hydrological modeling, being interdisciplinary, complex and dynamic by nature, intrinsically asks for better integration of data, information and models [4-6]. The objective is to bring to policy/decision-makers efficient tools as well as suitable and reliable information, both supported by scientific knowledge and models.

Hydrological models are simplified representations of parts of the water cycle. They are primarily used for predictions and for understanding hydrological processes. The Soil and Water Assessment Tool<sup>1</sup> (SWAT) [7, 8] is a semi-distributed, continuous watershed simulator operating on a daily time step. It is developed by USDA Agricultural Research Service (USDA-ARS) and Texas A&M AgriLife Research, for assessing the impact of management options as well as global changes on water supplies, sediment transportation and agricultural chemical yields in watersheds and larger river basins. This model performs simulations that integrate various processes such as hydrology, climate, chemical transport, soil erosion, pesticide dynamics and agricultural management. SWAT accounts for soil and land cover conditions by subdividing the simulated catchment into homogeneous Hydrological Response Unit (HRU). The model uses a daily to sub-hourly time step, and can perform continuous simulation for a 1- to 100-year period. SWAT simulations are typically prepared from ArcGIS<sup>2</sup> (ArcSWAT) or MapWindow<sup>3</sup> interfaces (MWSWAT). For the simulation, SWAT requires data on elevation, soil, land cover, reservoirs, agricultural practices and weather for model setup. River discharges, water quality and crop yield (as available) are needed for calibration and uncertainty analyses. Once this data is gathered and formatted, SWAT is able to model the water cycle inland and in-stream components.

SWAT was used to simulate the continent of Africa [9] and in the "Hydrologic Unit Model for the United States" (HUMUS) [7], where the entire U.S. was simulated with good results for river discharges at around 6000 gauging stations.

<sup>1</sup> <http://swat.tamu.edu>

<sup>2</sup> <http://swat.tamu.edu/software/arcsbat/>

<sup>3</sup> <http://swat.tamu.edu/software/mwswat/>

This study is now extended within the national assessment of the USDA Conservation Effects Assessment Project (CEAP 21). Other large scale SWAT application included the work of Gosain et al. [10] where twelve large river catchments in India were modeled with the purpose of quantifying the climate change impact on hydrology. SWAT is recognized by the U.S. Environmental Protection Agency (EPA) and has been incorporated into the EPA's BASINS system (Better Assessment Science Integrating Point and Non-point Sources) [11].

In every SWAT simulation, several important hydrological variables are estimated (e.g., precipitation, snow or ice melting, evapotranspiration, water yield, groundwater recharge/transfer, suspended/dissolved load, pollutants) and stored in different text files<sup>4</sup>. They provide temporal resolution of yearly, monthly and daily time steps based on user interest for subbasin (output.sub), main river reach (output.rch) and HRU (output.hru). Potentially, SWAT provides many useful outputs for both scientists and decision-makers, but in a not very accessible format.

Preparing, calibrating, executing and publishing outputs of a SWAT model are often time-consuming and repetitive tasks. In particular, while gathering the required data to set up a SWAT model, users are regularly facing the problem of data accessibility and data heterogeneity (e.g., different data formats). Additionally, results of a SWAT simulation can not be visualized directly on a map. Different processing steps in various software are required for generating geospatial data from the output text files. Finally, these results are often prepared using closed/proprietary formats and therefore limit their usability and making them difficult to integrate with other data sources and/or models.

Recognizing the need for efficient and effective data accessibility and considering that SWAT outputs can be valuable for different community of users (e.g. hydrologists, environmentalists, biologists, decision-makers), a scalable and interoperable framework simplifying and automatizing repetitive tasks like data gathering and map production can be beneficial. Based on these considerations the aim of this paper is to present a proof of concept (1) to facilitate gathering and harmonization of SWAT data inputs, (2) to facilitate the publishing of SWAT simulation outputs, (3) to expose these results in a standardized way using interoperable services, and (4) to facilitate the exchange of data, and integration with other resources.

## II. BACKGROUND

### A. SWAT models preparation and outputs visualization

SWAT models preparation and outputs visualization are generally realized in ArcGIS [12] using ArcSWAT and VizSWAT<sup>5</sup> extensions for building the model and respectively visualizing results as dynamic graphs or maps. There is also an open source alternative to ArcSWAT providing the same functionalities for model preparation based on the

MapWindow<sup>6</sup> GIS system and the MWSWAT plugin [13]. In term of output visualization two other solutions are currently available. Field\_SWAT<sup>7</sup> [14] is a graphical user interface developed in MatLab for preparing maps and SWATShare<sup>8</sup> is a web-based tool for uploading, executing and visualizing SWAT simulations.

The first task required to users while setting up a SWAT model is to gather the necessary data. Traditionally, users must:

- 1) Identify the relevant data sources,
- 2) Download these data on their computer,
- 3) Harmonize data formats, resolution, and projections.

The data needed to prepare a SWAT model are:

- *Geospatial data (raster and vector)*: Digital Elevation Model (DEM), Land Use (LU), Soils, and river network.
- *Weather data (tables)*: (daily) precipitation and temperature.

Nutrient and sediment loads can be used to predict water quality. Additionally, climatic data from global or regional climate models can be used to predict the impacts of climate changes on the hydrological model. Once all these data have been gathered, downloaded and harmonized, then users can prepare, calibrate and execute their SWAT models.

Once a model has been executed then users want to visualize the results of their simulations. A typical workflow for processing outputs to prepare maps involves the following steps:

- 1) Open a text editor to filter and remove unnecessary columns in the output file (e.g., output.sub, output.rch)
- 2) In a spreadsheet editor open the cleaned output file, separate independent variables using tab delimited option, use pivot table to calculate average the values, and finally save independent value (each variable) in a text file for each variable (in csv format).
- 3) In a GIS software, join the shapefile of the watershed delineation or the river reaches with the table values, classify data according to the values of the variable, and finally save the map.

These tasks involve different proprietary software that use closed formats [12]. Consequently, a standardized approach for a rapid collection of required datasets for a given geographical area is needed. It should automatically structure the data into the input format of SWAT. An automatic procedure to publish maps-results based on non-proprietary formats can be very convenient as well to improve the interoperability of SWAT outputs.

<sup>4</sup> [http://swat.tamu.edu/media/69395/ch32\\_output.pdf](http://swat.tamu.edu/media/69395/ch32_output.pdf)

<sup>5</sup> <http://swat.tamu.edu/software/vizswat/>

<sup>6</sup> <http://www.mapwindow.org>

<sup>7</sup> [http://baegrisk.ddns.uark.edu/SWAT\\_Model\\_Tools/Field\\_SWAT/](http://baegrisk.ddns.uark.edu/SWAT_Model_Tools/Field_SWAT/)

<sup>8</sup> <https://water-hub.org/swat-tool>

### B. Interoperability in the water domain

Presently, hydrological and meteorological data still remain difficult to find, access, and integrate because of various incompatibilities (e.g., data formats, models specifications, quality needs), missing documentation (e.g., metadata), data fragmentation and replication, data policies, and these systems are operating in isolation [15]. Interoperability, the ability to exchange and use information between two or more systems/components, is therefore an essential condition to enable efficient data publishing, discovery, evaluation and access to environmental data [16].

Current technologies are suitable to match these requirements only if open software interfaces and standards are developed allowing these technologies to interoperate at a large scale [17]. The Open Geospatial Consortium (OGC) aims to develop and provide such standards enabling communication and exchange of information between systems of different types operated with distinctive software. Indeed, a non-interoperable system cannot share data and computing resources, inducing scientists to spend much more time than necessary on data discovery and transformations. One of the major benefits of interoperability is to enable locally managed and distributed heterogeneous systems (e.g., different operating systems, databases, data formats) to exchange data and provide a service [18]. A good example is the distributed information system developed by the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI). It is a web-based system for storing, publishing, sharing, processing, and analyzing hydrological data through a full suite of software and standardized/interoperable services [19].

The OGC, completed by ISO standards, is providing a suite of standard specifications to search, discover, and access of heterogeneous geospatial resources. These resources can be maps served with Web Map Service (WMS) [20], vectors and raster data published respectively as Web Feature Service (WFS) [21] and Web Coverage Service (WCS) [22], or processing algorithms exposed as Web Processing Service (WPS) [23]. Data and services can be documented through International Organization for Standardization (ISO) 19115 (resource metadata), 19139 (metadata encoding) and 19119 (service metadata). ISO standards are complemented by the OGC Catalog Service for the Web (CSW) specification [24] defining an interoperable interface to publish, discover, search and query metadata.

Currently, the OGC has several projects underway related to water resources<sup>9</sup>. In particular, it has a Hydrology Domain Working Group<sup>10</sup> that is seeking to develop and provide solutions for describing and exchanging data related to water resources. For example, WaterML 2.0 has been recently accepted as a standard<sup>11</sup>. WaterML is an XML-based specification used to formally describe hydrological data and act as an interchange format via the Internet through web services. It contains specifications for both point and spatial

coverage data (via XML elements) as well as a set of generic vocabularies. Additionally, the OGC is also conducting Interoperability Experiments on Surface Water, Ground Water, and Hydrologic Forecasting as well as developing pilot studies on Hydro-climatology Information Sharing.

Finally, several initiatives are catalyzing data sharing by promoting interoperability to maximize the (re)use of data and supporting easy access to and utilization of geospatial data. At the global level, the Global Earth Observation System of Systems (GEOSS) [25] has a dedicated Societal Benefit Area on Water<sup>12</sup> and related activities like the Water Cycle Integrator or Interoperability Experiments on Weather, Ocean, and Water. Similarly, Eye on Earth has recently launched a special initiative on Water Security<sup>13</sup>. At the European level the Infrastructure for Spatial Information in the European Community (INSPIRE) [26] has a "Cross-Border Water Management" Initiative<sup>14</sup> to contribute to the implementation of the Water Framework Directive.

Accordingly, making SWAT outputs interoperable will simplify their sharing/exchange, expend their potential applicability and facilitate their contribution initiatives like GEOSS or INSPIRE.

### III. OWS4SWAT FRAMEWORK

The proposed framework called OGC Web Services for SWAT (OWS4SWAT) is entirely based on OGC standards to help: (1) data gathering and harmonization, while setting up a model, and (2) map preparation and publication, when results of simulations are available. Currently, OWS4SWAT is not influencing model preparation, calibration, and execution. These operations are still manually executed on desktop computers (Table 1).

TABLE I. COMPARISON BETWEEN "TRADITIONAL" AND OWS4SWAT WORKFLOWS

	Traditional	OWS4SWAT
<b>Data gathering and harmonization</b>	Manual download, processing/harmonization, repetitive tasks, heterogeneous and not interoperable data.	Automatic download and processing accomplished by webservices on the server, based on OGC standards
<b>Model preparation, calibration, and execution</b>	Manual on desktop computer	Manual on desktop computer
<b>Outputs preparation and publication</b>	Manual preparation and upload on server, repetitive tasks, results often not interoperable	Automatic preparation and publication done by webservices, OGC-compliant

Thanks to the use of interoperable services, this will not be dedicated to specific software, consent the use of different clients (e.g., desktop, web), and facilitate data access, exchange and integration.

<sup>9</sup> <http://www.opengeospatial.org/node/1535>

<sup>10</sup> <http://www.opengeospatial.org/projects/groups/hydrologydwg>

<sup>11</sup> <http://www.opengeospatial.org/standards/waterml>

<sup>12</sup> [http://www.earthobservations.org/geoss\\_wa.shtml](http://www.earthobservations.org/geoss_wa.shtml)

<sup>13</sup>

<http://www.ogcnetwork.net/system/files/EoE%20SI%20Water%20Security%20-12-for%20Summit.pdf>

<sup>14</sup> <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/2688>

### A. Architecture and design

The two main functions of OWS4SWAT are built with different software packages:

- **Data processing/handling:** is programmed using PyWPS<sup>15</sup>, an OGC WPS 1.0.0 implementation written in Python. PyWPS does not process data by itself but wraps different backends to both access geospatial (GRASS) and statistical (R) functionalities.
- **Data publishing:** is based on the OpenGeo Suite Community Edition<sup>16</sup>. It is an integrated software package made of different components to store (e.g., PostgreSQL/PostGIS), publish (e.g., GeoServer), and develop web-mapping applications (e.g., OpenLayers, GeoExt) based on WMS, WFS, and WCS OGC standards.

The advantages of using these software solutions is that they fully implement OGC standards, are Free and Open Source, have an important community of both users and developers, and can be installed on different Operating Systems (e.g., Linux, Windows, Mac).

Within this architecture, different WPS services are developed to answer the requirements of data download/harmonization and SWAT results publishing. Outputs data are made available using WMS for visualization and WFS for data access. Once published data can be accessed, manipulated, styled, and integrated in desktop (e.g., QGIS) or web-based (e.g., GeoExplorer) clients (fig.1).

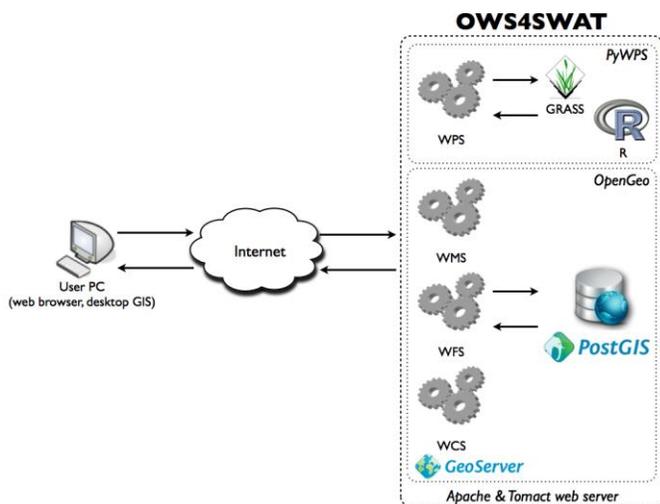


Fig. 1. OWS4SWAT architecture for managing SWAT data inputs and publishing its outputs as OGC web services.

### B. Web Processing Service (WPS)

To enable interoperability and automatization of tasks, OWS4SWAT framework mostly relies OGC Web Processing Service (WPS) specification. WPS processes are flexible and remotely accessible algorithms available through web services and can be reused in different workflows [27, 28]. The core element of a WPS is the process, a calculation with defined inputs and outputs [29, 30]. It allows users to know which processes are available, to select the required input data and their formats, to create a model and run it, to manage processes

(status, storage for the output ...) and to return the output once computation is completed.

As any web service, a WPS instance must expose different operations that are accessible through standardized web communication (e.g., XML). In particular, descriptions of algorithms through metadata that are usable and understandable both by humans and other web services [31] are essential elements to develop chains of services [32]. WPS specification includes a set of three operations that can be called using HTTP-GET, HTTP-POST or SOAP/WSDL:

- **GetCapabilities:** answer to a client describing its capabilities in an XML document. It tells the client which kinds of process are available.

```
http://localhost/cgi-bin/wps?
service=WPS&request=getcapabilities
```

- **DescribeProcess:** describe the parameters of a selected process also through an XML document (e.g., input and output).

```
http://localhost/cgi-bin/wps?service=WPS&version=1.0.0&
identifier=buffer&request=describeprocess
```

- **Execute:** execute a selected process.

```
http://localhost/cgi-bin/wps?service=WPS&version=1.0.0&
identifier=buffer&request=execute&datainputs=[
data=http://foo.bar/cities.gml; width=3]
```

The output of a process can be obtained either by a direct download (e.g., result sent immediately to the client after the end of the execution) or as resource stored on the server and accessible through the web using URLs [33]. In such a case, the Execute response will be an XML document providing the URLs to access each stored output.

In a web service environment, it is possible to seamlessly couple and reuse services to perform various (complex) tasks organized in sequences or chains of processes. Service chaining can be defined as a mechanism to build flexible, coherent, and efficient workflows by combining individual web services to create customized web applications based on distributed services [34-36]. This offer significant potential to modularize, reuse, and share software components [37]. The International Organization for Standardization (ISO) through its ISO 19119 standard defines three types of chaining mechanisms: (1) *transparent* where the workflow is defined and managed by users, (2) *translucent* where users are familiar with atomic services that compose the chain and invoke a service to manage the chain, (3) *opaque* where users invoke an aggregated service that execute the chain but do not have knowledge about the atomic service constituting the chain. These chains of services can be coordinated by orchestration engines generally based on Simple Object Access Protocol (SOAP) and Web Service Description Language (WSDL) to exchange structured information [38].

### C. Implementation

The general workflow when working with SWAT is the following:

- 1) A user needs to download and harmonize required data to build a SWAT model.
- 2) Once he has all data he prepares, calibrates and executes the model on a desktop computer.
- 3) The user retrieves the results of a simulation from output.sub, output.rch, output.hru files and prepare maps and/or graphs to visualize and share its results.

This general workflow has been transposed in a web service environment (fig.2) assuming that the step 2 (e.g., model preparation, calibration, execution) is done on a Desktop computer and therefore can be considered as a standalone task that do not require interactions with web services.

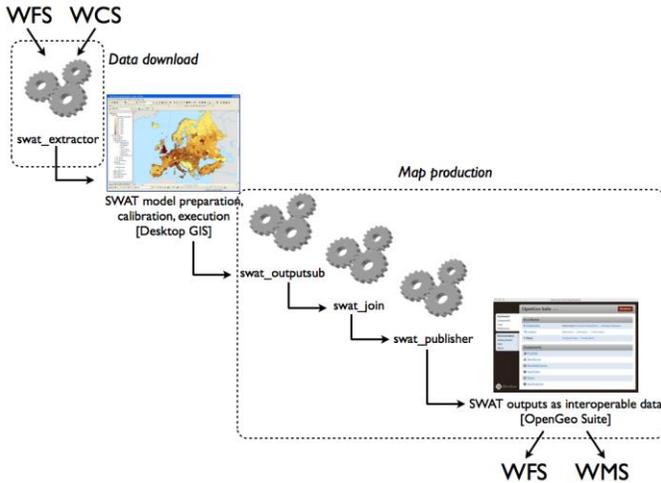


Fig. 2. General workflow in a web service environment for downloading SWAT input data (step1), modeling (step2) and publishing of the outputs (step3)..

To differentiate between data download (step 1) and map production tasks (step 3), the general workflow has been subdivided into two independent workflows. This also reflects the fact that step 1 (fig.3) is accomplished before the modeling exercise, while step 3 (fig.4) is completed after successful execution of the model.

The tasks mentioned in section II.A have been disaggregated in atomic functionalities and implemented as dedicated WPS processes that can be chained to achieve the objective of automating of processing tasks for data download and maps production. The different processes are written as PyWPS scripts that, depending on the required functionalities, will interact with GRASS for geoprocessing algorithms and R for statistical functions

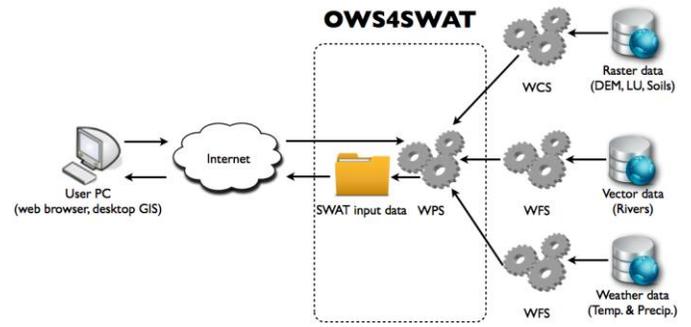


Fig. 3. Download workflow implemented in OWS4SWAT with the swat\_extractor WPS process (step1).

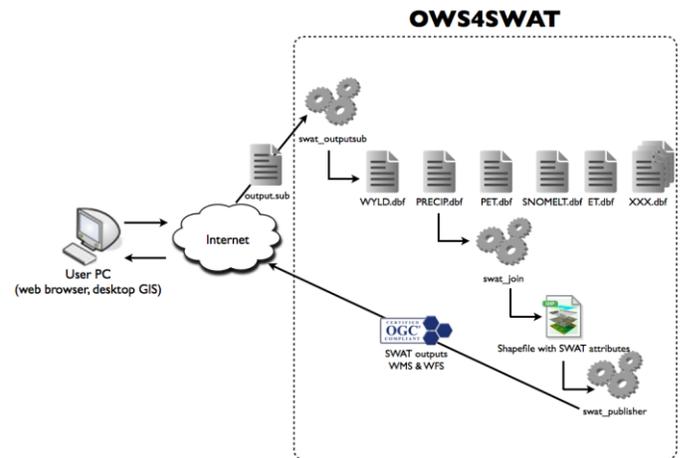


Fig. 4. Maps production workflow in OWS4SWAT with a chain of three WPS processes (swat\_outputsub, swat\_join, swat\_publisher) (step3).

- *swat\_extractor*: extract required SWAT input data from different WFS/WCS endpoints for a dedicated area (e.g, bounding box) specified by the user and store them in zip file.
- *swat\_outputsub*: generates from the output.sub (e.g., .rch and .hru files are not handled) file a set of DBF files, one for each SWAT simulated variables.
- *swat\_join*: join all the DBF files with the a geospatial file of the computed subbasin.
- *swat\_publisher*: store in a PostGIS database the geospatial file with all SWAT variables as attributes and publish this file with GeoServer using OGC WMS and WFS standards.

Processes with their identifier, computing backends, data input and output variables are summarized in Table 2.

TABLE II. OWS4SWAT PROCESSES DESCRIPTION

identifier	backend	inputs	outputs
swat_extractor	GRASS	riversInUrl demInUrl soilsInUrl landuseInUrl tempInUrl precipInUrl northBound southBound eastBound westBound	swatOut
swat_outputsub	R	outputSub	swatDbf
swat_join	GRASS	dbfFile shpFile	shpOut
swat_publisher	GRASS	shpIn wkspIn dstoreIn pgIn	Shapefile stored in PostGIS and published in GeoServer with WMS and WFS

All these WPS processes are individually available and they can be executed with various desktop or web-based WPS clients. For example, ArcGIS 10.1 or QGIS 1.8 (with WPS plugin) are able to consume the proposed services and to execute the different tasks of data download and map production/publication. More specifically, with the map production/publication workflow, swat\_outputsub, swat\_join, and swat\_publisher processes can be chained because output of a process can be used as input for the following process [39]. Since version 3.2, PyWPS implements a SOAP/WSDL interface enabling chaining WPS services in WSDL-based Workflow Management Systems. [39].

The WSDL file is generated dynamically by making a WSDL request to the WPS instance: <http://localhost/cgi-bin/wps?wsdl>. The description is created applying a XSLT template to a *DescribeProcess* operation output and contains all the processes request/responses provided by the WPS instance allowing exposing them as WSDL-based services. These different processes have been successfully chained with Taverna<sup>15</sup> and SEXTANTE<sup>16</sup> modeler. The former is a scientific workflow management system while the latter is a geospatial data processing framework available for different Desktop GIS clients like QGIS and providing different tools such as a graphical modeler. Interestingly, Taverna, that is not meant to be a geospatial workflow manager, has been able to efficiently handle geospatial data.

Consequently, executing this workflow enables users to automatically process and publish the different simulated SWAT variables with OGC standards using both specialized and non-specialized software. Users only have to provide their output.sub file and the chain of processing services will execute all the tasks and greatly simplified the publication of their results.

D. Use-case: publishing results of the enviroGRIDS Black Sea catchment SWAT model.

The prototype OWS4SWAT framework has been developed in the context of the enviroGRIDS project, funded by the European Commission (EC) Seventh Framework Program. This project concentrates on the unsustainable development and the inadequate resource management that is affecting the Black Sea catchment region. A large catalog of environmental data sets (e.g., land use, hydrology, and climate) has been gathered, published and used to carry out distributed spatially explicit simulations to build scenarios of key environmental changes.

Advances in distributed computing in conjunction with data availability from interoperable web services have made high-resolution modeling of distributed hydrologic processes possible [40, 41]. In the frame of enviroGRIDS, a high-resolution SWAT (sub-catchment spatial and daily temporal resolution) model of the entire Black Sea catchment has been developed. This model, divided into 12982 subbasins, which were further divided into 89202 HRUs, will be used to predict water quality and quantity according to the different scenarios in the region (e.g. Land Use, Climate, and Demographic changes). Subsequent analyses of land use change, agricultural management change, and/or climate change can then predict the consequence of various scenarios. Finally, all results of the Black Sea SWAT model and scenarios will be registered and made available as OGC services to feed the Global Earth Observation System of Systems (GEOSS) and contribute to the Water Societal Benefit Area (SBA). Consequently, the OWS4SWAT framework has been used to test the validity of the proposed approach for facilitating the publication of the Black Sea SWAT model results. This first experiment have shown that it facilitates and accelerates the publication of SWAT outputs, allow to process larger files that where difficult to handle on desktop computer, and simplify the access and integration with other data sources in different clients (fig.5&6).

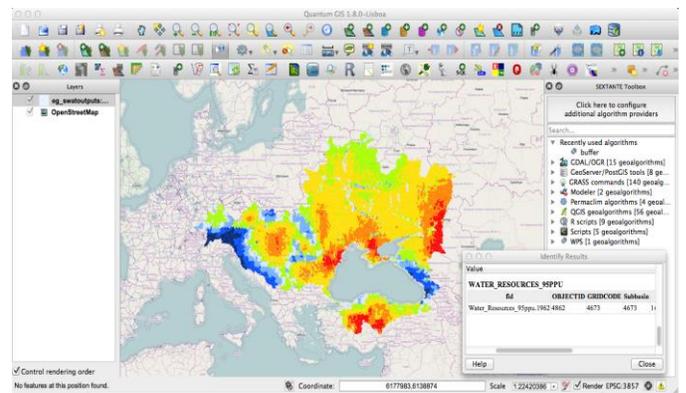


Fig. 5. Soil Water content (e.g., SW variable) modeled with SWAT, accessed in WFS in QGIS Desktop GIS, and integrated with OpenStreetMap background.

<sup>15</sup> <http://www.taverna.org.uk>

<sup>16</sup> <http://www.sextantegis.com>

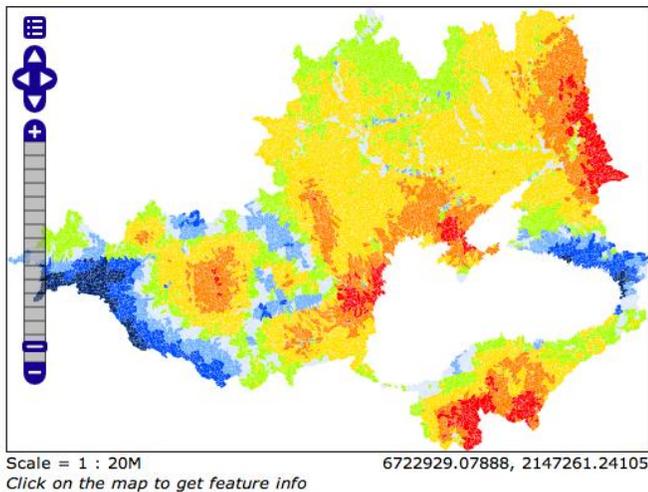


Fig. 6. Soil Water content (e.g., SW variable) in the Black Sea catchment visualized with WMS in a web-based OpenLayers client.

#### IV. DISCUSSION

The OWS4SWAT framework is, to our knowledge, among the first attempt to bring interoperability based on OGC specifications around SWAT software. The only study we have found concerns the derivation of HRUs based on a WPS implementation [42]. The proposed approach was developed as a proof-of-concept and the implementation was successful. The first results have highlighted both benefits and limitations but we are convinced that such approach can bring relevant benefits for the numerous SWAT users (e.g., hydrologists) as well as users of SWAT results (e.g., scientists of other communities).

##### A. Benefits

The major benefit of this approach is to enable interoperability around SWAT applications. Indeed, using OGC standards helps to solve the problem of seamlessly integrating multiple heterogeneous data source [43]. From a scientific perspective, having data published in a comparable form considerably facilitate data acquisition, interpretation and comprehension [43]. Several studies have already demonstrated the benefits of interoperability and web services in the water domain for data visualization [44], data publication [45], data distribution [46], data discovery and retrieval [47] and modeling [48]. All these authors stress the fact that interoperability offers new and promising opportunities to the water research community for systematic data management, publication and analysis [49].

Our approach helps overcoming the problems of poor data accessibility and of different formats for specialized scientific models. It enables a standardized approach for rapid data collection on a given geographic area and automatically formats data into the input format of the targeted model. It makes environmental modeling much more efficient by making better use of existing data sources and by reducing the time for finding, gathering and preparing environmental input data. It also simplifies and accelerates the publication of model results by automatizing repetitive tasks through a chain of processing services. Moreover, due to the fact that the processing is

executed on a server and not on a desktop computer, it allows faster processing of data of a given size and also allows processing larger data sets (i.e., higher resolution). Finally, it increases model results availability and discovery by publishing them according to OGC standards. This facilitates sharing, exchange, and integration of SWAT output data with other data sources. Data can be interactively visualized with WMS, accessed with WFS for subsequent geospatial or graph analysis [50], and processed with WPS. Consequently, data are no more restricted to dedicated software or formats but instead can be consumed by a wide variety of clients. It can be light web-based client like the enviroGRIDS portal<sup>17</sup>, desktop GIS clients like QGIS, or specialized hydrological software like HydroDesktop [51].

The scalability of the OWS4SWAT is another advantage thanks to the use of interoperable services. Indeed, it is not difficult to incorporate different (data and/or processing) services implemented by other providers or create new services to build new and more complex workflows. Therefore, composability, extension, and integration are further enhanced and allow envisioning interaction with different scientific disciplines and coupling other models. Integration through web service of heterogeneous data and modeling resources have been demonstrated in hydrology and climatology [52], surface dynamics [53], biodiversity [54], and ecosystem services [55]. Despite the fact that many challenges need to be overcome (e.g., discovery, semantic, ontologies, performances) [56], all authors recognize that there is a need to improve collaboration among scientific disciplines and the diversity of environmental models requires effective and efficient solutions of using and reusing functionalities or services provided by others [57]. Promising solutions have been developed to expose models either using WPS [58] or Open Modeling Interface (OpenMI) [59] standards. Model integration based on interoperable services can significantly reduce time, efforts, and technical challenges for scientists who only have to concentrate on their expertise while developing components [4]. In the coming years, the Group on Earth Observation (GEO) Model Web initiative will certainly catalyze and improve models access, sharing, and integration [56] and ultimately enable an holistic vision/understanding of the Earth system to better address decision-making processes.

##### B. Limitations.

The current implementation of OSW4SWAT framework has different technical limitations that may be overcome with further developments.

At the moment, this is still a prototype and has only been tested for publishing some intermediate SWAT outputs. Moreover, only output.sub files at daily time step can be handled and two other processes need to be developed to process .rch and .hru files. The next step is therefore to move to a production scenario ingesting and publishing all results of a SWAT model.

SWAT outputs visualization not only concerns maps but should also graph generation. Two solutions exist to tackle this issue: (1) develop a dedicated process to generate graphs

<sup>17</sup> <http://portal.envirogrids.net>

directly from raw results (e.g., output.sub file) or (2) due to the fact that SWAT results can already be published in WFS, it is possible to access the attribute table and consequently developing a process to access directly data in WFS and generate graphs.

The use of interoperable services can also be perceived as a limitation because as of today not all data providers are publishing their data using standards. Moreover, even if data providers use standards, they may differ from one scientific community to another (e.g., netCDF in climatology, WaterML in hydrology). Consequently, the current implementation of OWS4SWAT is limited to resources exposed as WMS, WFS, WCS, and WPS. A possible solution to access resources based on standard and non-standard capacities is to rely on the brokering approach to access heterogeneous resources in a consistent and uniform manner [60].

### C. Perspectives

Thanks to its scalability the OWS4SWAT framework will be improved with forthcoming developments:

- Process to publish SWAT results using the recently adopted OGC WaterML2.0. This will increase interoperability between SWAT and other hydrological systems/applications.
- Implementation of a process to visualize SWAT outputs as graphs.
- Develop a new workflow to compute vulnerability maps of water resources based on SWAT outputs.

### V. CONCLUSIONS

The proposed OWS4SWAT framework enables SWAT users to share more easily and efficiently their model results using OGC standards. By making these results interoperable, it facilitates the exchange of data, and integration with other resources. SWAT model outputs accessibility and visualization are consequently no more restricted to dedicated software but can be available for various types of (desktop or web-based) clients. This can greatly expand the use of SWAT results and their applications.

The use of chained WPS services has simplified the processing of SWAT results and the automating of repetitive tasks of data handling, map preparation, and data publication. Typically, these tasks are executed more rapidly compared to the "traditional" way of manipulating SWAT model outputs for preparing maps. Furthermore, it enhances the scalability of the framework allowing one to easily and seamlessly incorporate new services, to extend existing workflows or create new ones.

Interoperable data and processing services not only help scientists to share their data or computational algorithms but also enhance their reusability and therefore can facilitate the development of complex scientific workflows to solve complex problems. Indeed, by linking distributed and heterogeneous data and processing services, it offers new opportunities to scientists for processing data and for communicating scientific results and hence contributing of better resource management and help decision-makers. It makes possible to benefit from the abundant scientific resources to more efficiently explore and

better understand complex interactions between the different components of the Earth system. Ultimately, it offers a promising potential for coupling different models (e.g., the output of one model serves as input in the other, the communication is based on interoperable services) and therefore facilitate the development of integrated models. Finally, sharing data, processes and models, is also part of the elementary scientific approach and thus enhance scientific accountability, credibility, and facilitate the replication and comparison of workflows and methodologies.

### ACKNOWLEDGMENT

The authors would like to acknowledge the European Commission "Seventh Framework Program" that funded the enviroGRIDS (Grant Agreement no. 227640) and ACQWA projects (Grant Agreement no. 212250). We thank Jorge Mendes de Jesus, Jachmy Cepicky, and PyWPS community for their valuable help, support and developments. The views expressed in the paper are those of the authors and do not necessarily reflect the views of the institutions they belong to.

### REFERENCES

- [1] UNEP, Global Environment Outlook (GEO) - 5: Environment for the future we want, 2012. p. 550.
- [2] Gerlak, A.K., J. Lautze, and M. Giordano, Water resources data and information exchange in transboundary water treaties. International Environmental Agreements-Politics Law and Economics, 2011. **11**(2): p. 179-199.
- [3] Roehring, J., Information Interoperability for River Basin Management, in Technology Resource Management and Development2002. p. 127-134.
- [4] Argent, R.M., An overview of model integration for environmental applications. Components, frameworks and semantics. Environmental Modelling & Software, 2004. **19**(3): p. 219-234.
- [5] Buytaert, W., et al., Web-Based Environmental Simulation: Bridging the Gap between Scientific Modeling and Decision-Making. Environmental Science & Technology, 2012. **46**(4): p. 1971-1976.
- [6] Papajorgji, P., H.W. Beck, and J.L. Braga, An architecture for developing service-oriented and component-based environmental models. Ecological Modelling, 2004. **179**(1): p. 61-76.
- [7] Arnold, J., et al., Large area hydrologic modeling and assessment - Part I: Model development. Water resources bulletin, 1998. **34**(1): p. 73-89.
- [8] Srinivasan, R., et al., Large area hydrologic modeling and assessment - Part II: Model application. Water resources bulletin, 1998. **34**(1): p. 91-101.
- [9] Schuol, J., et al., Modelling Blue and Green Water Availability in Africa at Monthly Intervals and Subbasin Level. Water Resources Research, 2008. **44**: p. 1-18.
- [10] Gosain, A.K., S. Rao, and D. Basuray, Climate change impact assessment on hydrology of Indian river basins. Current Science, 2006. **90**(3): p. 346-353.
- [11] Gassman, P.W., et al., THE SOIL AND WATER ASSESSMENT TOOL: HISTORICAL DEVELOPMENT, APPLICATIONS, AND FUTURE RESEARCH DIRECTIONS. Transactions of the ASABE, 2007. **50**(4): p. 1211-1250.
- [12] Winchell, M., et al., AcSWAT interface for SWAT2005 - User's guide, 2007. p. 436.
- [13] George, C. and L.F. Leon, WaterBase: SWAT in an Open Source GIS. The Open Hydrology Journal, 2008. **2**: p. 1-6.
- [14] Pai, N., D. Saraswat, and R. Srinivasan, Field\_SWAT: A tool for mapping SWAT output to field boundaries. Computers & Geosciences, 2012. **40**(0): p. 175-184.
- [15] Beniston, M., et al., Obstacles to data access for research related to climate and water: Implications for science and EU policy-making. Environmental Science & Policy, 2012. **17**(0): p. 41-48.

- [16] Giuliani, G., et al., Sharing Environmental Data through GEOSS. International Journal of Applied Geospatial Research, 2011. **2**(1): p. 1-17.
- [17] McKee, L., 18 reasons for open publication of geoscience data, 2010, Earthzine. p. 1-8.
- [18] Open Geospatial Consortium, The Havoc of Non-Interoperability, 2004. p. 7.
- [19] Ames, D.P., et al., Introducing the Open Source CUAHSI Hydrologic Information System Desktop Application (HIS Desktop). 18th World Imacs Congress and Modsim09 International Congress on Modelling and Simulation, 2009: p. 4353-4359.
- [20] Open Geospatial Consortium, OpenGIS Web Map Server Implementation Specification, 2006. p. 85.
- [21] Open Geospatial Consortium, Web Feature Service Implementation Specification, 2005. p. 131.
- [22] Open Geospatial Consortium, Web Coverage Service (WCS) Implementation Specification, 2006, 143.
- [23] Open Geospatial Consortium, OpenGIS Web Processing Service, 2007. p. 87.
- [24] Open Geospatial Consortium, OpenGIS Catalogue Services Specification, 2007. p. 218.
- [25] GEO secretariat, Global Earth Observation System of Systems 10-Year Implementation Plan Reference Document, 2005. p. 209.
- [26] European Commission, Directive 2007/2/EC of the European Parliament and the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007: Brussels. p. 14.
- [27] Kiehle, C., K. Greve, and C. Heier, Standardized geoprocessing - Taking Spatial Data Infrastructures one Step Further, in 9th AGILE Conference on Geographic Information Science 2006: Visegrad. p. 273-282.
- [28] Michaelis, C.D. and D.P. Ames, Evaluation and Implementation of the OGC Web Processing Service for Use in Client-Side GIS. Geoinformatica, 2009. **13**(1): p. 109-120.
- [29] Granell, C., et al., Interlinking Geoprocessing Services. Second International Conference on Advanced Geographic Information Systems, Applications, and Services: Geoprocessing 2010, Proceedings, 2010: p. 99-104.
- [30] Granell, C., M. Gould, and M.A. Esbri, Geospatial Web Service Chaining, in Handbook of Research on Geoinformatics, H.A. Karimi, Editor 2009. p. 9.
- [31] Kiehle, C., Business logic for geoprocessing of distributed geodata. Computers & Geosciences, 2006. **32**(10): p. 1746-1757.
- [32] Schaffer, B. and T. Foerster, A client for distributed geo-processing and workflow design. Journal of Location Based Services, 2008: p. 194-210.
- [33] Schaeffer, B., Towards a Transactional Web Processing Service (WPS-T), in GI Days 2008: Münster, Germany. p. 27.
- [34] Diaz, L., C. Granell, and M. Gould, Case study: Geospatial processing services for web-based hydrological application, in Geospatial Services and Applications for the Internet 2008. p. 31-47.
- [35] Friis-Christensen, A., et al., Designing Service Architectures for Distributed Geoprocessing: Challenges and Future Directions. Transactions in GIS, 2007. **11**(6): p. 799-818.
- [36] Granell, C., L. Diaz, and M. Gould, Distributed Geospatial Processing Services, in Encyclopedia of Information Science and Technology 2009, Information Science Reference. p. 1186-1193.
- [37] Guru, S., M., et al., Challenges in using scientific workflow tools in the hydrology domain, in World IMACS/MODSIM Confess 2009: Cairns, Australia.
- [38] Fleuren, T. and P. Muller. BPEL workflows combining standard OGC Web services and grid-enabled OGC Web services. in 2008 34th Euromicro Conference Software Engineering and Advanced Applications (SEAA). 2008. Parma, Italy: Ieee.
- [39] de Jesus, J., et al., WPS orchestration using the Taverna workbench: The eScience approach. Computers & Geosciences, 2012. **47**(0): p. 75-86.
- [40] Lecca, G., et al., Grid computing technology for hydrological applications. Journal of Hydrology, 2011. **403**(1-2): p. 186-199.
- [41] Ray, N., et al., Distributed Geocomputation for Modeling the Hydrology of the Black Sea Watershed, in Environmental Security in Watersheds: The Sea of Azov, V. Lagutov, Editor 2012, Springer. p. 141-157.
- [42] Cepek, A., Proceedings of the Workshop Geoinformatics FCE CTU. 2008. **3**.
- [43] Horsburgh, J.S., et al., Components of an environmental observatory information system. Computers & Geosciences, 2011. **37**(2): p. 207-218.
- [44] Kao, S., et al., Visualisation of hydrological observations in the water data transfer format. Environmental Modelling & Software, 2011. **26**: p. 1767-1769.
- [45] Horsburgh, J.S., et al., An integrated system for publishing environmental observations data. Environmental Modelling & Software, 2009. **24**(8): p. 879-888.
- [46] Kanwar, R., U. Narayan, and V. Lakshmi, Web service based hydrologic data distribution system. Computers & Geosciences, 2010. **36**(7): p. 819-826.
- [47] Huang, M.T., D.R. Maidment, and Y. Tian, Using SOA and RIAs for water data discovery and retrieval. Environmental Modelling & Software, 2011. **26**(11): p. 1309-1324.
- [48] Goodall, J.L., B.F. Robinson, and A.M. Castronova, Modeling water resource systems using a service-oriented computing paradigm. Environmental Modelling & Software, 2011. **26**(5): p. 573-582.
- [49] Tarboton, D.G., et al., Development of a Community Hydrologic Information System. 18th World Imacs Congress and Modsim09 International Congress on Modelling and Simulation, 2009: p. 988-994.
- [50] Jiang, J., et al., The research and application of WFS based GML, in XXIst ISPRS Congress 2008: Beijing, China.
- [51] Ames, D.P., et al., HydroDesktop: Web services-based software for hydrologic data discovery, download, visualization, and analysis. Environmental Modelling & Software, 2012. **37**(0): p. 146-156.
- [52] Salas, F.R., et al., Crossing the digital divide: an interoperable solution for sharing time series and coverages in Earth sciences. Nat. Hazards Earth Syst. Sci., 2012. **12**(10): p. 3013-3029.
- [53] Peckham, S.D. and J.L. Goodall, Driving plug-and-play models with data from web services: A demonstration of interoperability between CSDMS and CUAHSI-HIS. Computers & Geosciences, (0).
- [54] Nativi, S., et al., Biodiversity and climate change use scenarios framework for the GEOSS interoperability pilot process. Ecological Informatics, 2009. **4**(1): p. 23-33.
- [55] Feng, M., et al., Prototyping an online wetland ecosystem services model using open model sharing standards. Environmental Modelling & Software, 2011. **26**(4): p. 458-468.
- [56] Nativi, S., P. Mazzetti, and G.N. Geller, Environmental model access and interoperability: The GEO Model Web initiative. Environmental Modelling and Software, 2013. **39**: p. 214-228.
- [57] Papajorgji, P., A plug and play approach for developing environmental models. Environmental Modelling & Software, 2005. **20**(10): p. 1353-1357.
- [58] Castronova, A.M., J.L. Goodall, and M.M. Elag, Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. Environmental Modelling & Software, 2013. **41**(0): p. 72-83.
- [59] Castronova, A.M., J.L. Goodall, and M.B. Ercan, Integrated modeling within a Hydrologic Information System: An OpenMI based approach. Environmental Modelling and Software, 2013. **39**: p. 263-273.
- [60] Nativi, S., M. Craglia, and J. Pearlman, The brokering approach for multidisciplinary interoperability: a position paper. International Journal of Spatial Data Infrastructures Research, 2012. **7**: p. 1-15