Developing geoprocessing services for a hydrological model application

C. Granell, L. Díaz & M. Gould
Centre for Interactive Visualization, Universitat Jaume I, Castellón, Spain

V. Pascual & J. Guimet
Centre de Soporte IDEC, Institut Castogràfic de Catalunya, Barcelona, Spain

P. Carrara & M. Pepe
Institute for Electromagnetic Sensing of the Environment, National Research Council (IREA-CNR), Milano, Italy

Keywords: geoprocessing services, geoinformation processing, INSPIRE, Spatial Data Infrastructure (SDI), open source components

ABSTRACT: Spatial data infrastructures support the most common requirements of spatial information users –discovery, access, and visual overlay of datasets– however specialized users such as hydrological scientists require more advanced services for capturing, analyzing and processing huge volumes of data. To shift from locally-processed complex tasks to on-line, distributed service-data offers, we present a web-based prototype using distributed geoprocessing services for snowmelt hydrological and environmental models. This prototype is being solely built on open source components and technologies following a SDI-based architecture. We illustrate our web-based prototype in a GMES-funded scenario for predicting runoff in Alpine basins.

1 INTRODUCTION

Current developments of Geographic Information Systems (GIS) applications are based on a distributed model composed of independent and specialized geospatial web services, driven mainly by the need of offering distributed functionality over the Internet. Geospatial web services –web services that serve geospatial information to users– have evolved to be an efficient way of interoperability and become key pieces to build modular and distributed GIS applications. In fact, geospatial web services together with Spatial Data Infrastructures (SDI) support effectively the most common requirements and needs of spatial information users, such as discovering, accessing and visualization datasets. Connecting scientists to their data, models, and services is an important challenge for SDI, which is beginning to facilitate access to distributed, heterogeneous geospatial data through a set of policies, common rules, and standards that together help improve interoperability. Traditional discovery and visualization-based SDI is evolving to a more service-based vision in which geospatial web services are used not only to access geospatial data, but also to transform them and process them, often in service chains (Lemmens *et al.* 2006).

Scientists as hydrologists have specific and concrete requirements that make them to keep carrying out time-consuming tasks for capturing, analyzing and processing huge volumes of hydrological data by using traditional GIS tools rather than taking advantage of the benefits of SDI. Since the INSPIRE Directive, which aims to harmonise spatial information across Europe and to improve geospatial data services according to common principles, has been recently approved by the European Parliament and also officially published in the Official Journal of the European Union (European Union 2007), our goal here is to help hydrologists –and scientists in general interested in geospatial information– in approaching to INSPIRE philosophy to meet more efficiently their requirements. While hydrologists provide their knowledge and expertise, SDI researchers and programmers play an active role to provide proper geospatial and geoprocessing services, models and components to facilitate connecting such hydrologists with their data, models, and services within an interoperable architecture, fostering then the cooperation between data providers and users as proposed by the Global Monitoring for Environment and Security (GMES) initiative.

Furthermore, the focus of our paper is twofold. On one hand, we describe how distributed geoprocessing services can be integrated in current SDI infrastructures illustrated by a scenario for predicting runoff in the Alpine basins, in the framework of the ongoing AWARE Project (A tool for monitoring and forecasting Available WAter REsource in mountain environments, www.aware-eu.info), which may serve to easily extrapolate our experiences to other specialized application fields like environmental systems, making the SDI and the INSPIRE initiative available to a broader audience. On the other hand, since this project requires the development of a web-based prototype, similar to other research works (Caldeweyher *et al.* 2006), we are building such a prototype on the basis of open source components and technology, leading to user-friendly and INSPIRE-compliant tool that allows not only scientists and expert users but also non-expert users such as stakeholders to better interact with the prototype in a more understandable way.

The remainder of this paper is structured as follows. Section 2 introduces briefly the case study. Section 3 presents user requirement and system design. Section 4 overviews the underlying open source components and technologies used for our prototype, and it also describes some implementation issues to the user interfaces of our web-based prototype for snowmelt hydrological and environmental models. Finally, section 5 presents the conclusion and future work.

2 CASE STUDY

River discharge is an important issue to be monitored because of its significant influence on environmental systems and on human lives for water resource exploitation and hazards related to floods and landslides.. The first task for supporting these monitoring applications is to identify the people who would use the software. Often data providers, who own the data necessary, and scientists, experts in hydrological models, differ in goals and objectives, leading to a lack of collaboration among them. Then, the main goal of the AWARE project, a multidisciplinary project that is carried out by a team of hydrologists, remote sensing specialists, and information system researchers (Rampini *et al*. 2006), is to put together data providers, experts and scientists by developing a user-friendly web-based prototype for running snowmelt hydrological models, permitting not only expert users (hydrologists and other scientists) to run such models but also other types of end users and data providers (water policy makers, water supply and hydropower companies, irrigation consortia, public authorities). This implies usability requirements in our system design leading to a flexible, easy-to-use prototype that features intuitive interfaces and wizards assisting non-experts with the complex tasks of runoff forecast modelling and interpreting the results. Expert users often are more comfortable directly handling and analyzing data, and feel that the meaning and quality of data provided should be accurately investigated by them. In summary, the design of our web-based prototype should take into account usability, utility and flexibility. Such a tool is being developed according to these criteria as described in the following section.

3 SYSTEM DESIGN

As mentioned in the case study section, many end users own, locally, data necessary to run the model and so the first choice should be to allow them to feed the model with the local data they possess. Still, a goal of the AWARE project is to be compliant with the INSPIRE initiative so that other users might also discover and access geospatial data via common SDI services such as catalogues. For example any user might be interested in searching catalogues for appropriate satellite imagery for the study basin (geographic constraint) and during the snowmelt season (temporal constraint).

Apart from catalogue services, other kinds of services are used for our prototype. Most services implement interfaces defined by the Open Geospatial Consortium (OGC, www.opengeospatial.org/), and a recent resulting demonstration within the OWS Web Services Phase 4 (www.opengeospatial.org/pub/www/ows4/index.html) shows how those services may work collaboratively to offer geospatial interoperability in real-world scenarios. Examples of those services are Web Map Service (WMS) and Web Feature Service (WFS), which provide standards interfaces for querying and accessing map layers and geospatial data respectively. However, these are actually basic services and they seem to be insufficient to suit the processing requirements necessary for hydrological models.

The ongoing OGC Web Processing Service (WPS) specification (Schut & Whiteside 2005) provides interfaces for accessing more complex services and also for wrapping existing off-line services as web services. This recent specification allows not only to discover or present spatial information but to provide geoprocessing functionality, which are also independent of data and context, thereby providing high reusability. OGC WPS-based services or web geoprocessing services expose three methods to provide the geoprocessing functionality of a certain geoprocessing services by first using the *getCapabilities* method, common in other OpenGIS Web Services (OWS) (e.g., WMS and WFS), in order to know the methods offered. The geoprocessing services define input and output parameters in a very detailed way by providing a *describeProcess* method. Finally, the *execute* method actually invokes the geoprocessing service. In our opinion, OGC WPS services go beyond providing unique geoprocessing services because it can be an interesting specification with which to wrap both spatial and non-spatial processing services, leading to increased interoperability between OWS services and general purpose web services.

We are currently testing the OGC WPS specification for implementing complex tasks of hydrological models as distributed geoprocessing services (Díaz *et al.* 2007), by implementing a middleware platform made up of several basic and complex geoprocessing services that encapsulate all required tasks to access, analyze, and transform hydrological data into useful information, in order to then forward the results to the presentation tier. Next section focuses precisely on the presentation tier, showing the set of open source components and technologies necessary both to build the user interface and to interact with the geoprocessing services in the middleware.

4 IMPLEMENTATION

This section covers basic open source components and technologies used for our web-based prototype and also some implementation issues related to user interface aspects.

The web-based prototype is being developed entirely on open source software and components, selecting Java as base development platform on which other components rely on. Figure 1 illustrates how the open source components are related among them. Some of these components in a top-down order (see Figure 1) are the following:

- Struts Framework (struts.apache.org/) is an open source framework for building Servlet/JSP based web applications.
- OpenLayers (www.openlayers.org/) serves as a web mapping client compliant with OGC WMS and WFS services.
- Apache server (httpd.apache.org/) is the most popular web server.
- Apache Tomcat (tomcat.apache.org/) acts as a web container for java servlets execution and Java Server Pages (JSP). In our prototype, Apache server and Tomcat work together to offer java servlet capabilities through traditional HTTP queries.
- UMN MapServer (mapserver.gis.umn.edu/) provides web mapping capabilities. In this case, UMN plays the roles of WMS and WFS.
- PostgreSQL (www.postgresql.org) and PostGIS extension (postgis.refractions.net) as the spatial database.

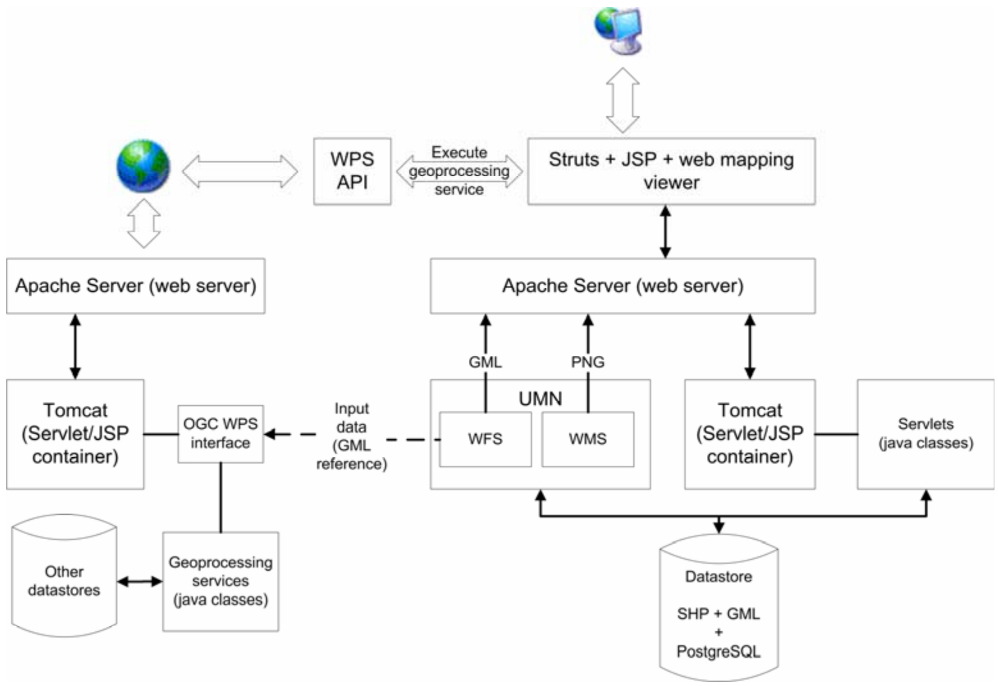

Figure 1. Open source components for the web-based prototype.

Users interact with the web-based geoportal for data discovery, quality control and input of parameters into selected hydrological models. This implies at least uploading local data necessary to feed the models, visualizing results, and executing geoprocessing services when necessary. The first two tasks are closely related with the geoportal user interface, which directly involves the components on the right-side of figure 1. Web forms and pages are implemented by means of the Struts framework and JSP pages. According to the application business logic, it is sometimes necessary to visualize local data and intermediate results in a dynamic map to provide feedback to users. This is done by embedding a web mapping viewer component such as OpenLayers in the web pages. Figure 2 and 3 show this component running on the right side. As usual in any web application, web pages are served by the web server (Apache server), yet in this case we have integrated Tomcat and UMN MapServer with Apache server to offer a single entry-point for these three services, taking advantage of the robustness and performance features provided by Apache server. The UMN MapServer allows us to perform two key aspects. First, the WMS provides images from well known geodata formats (e.g. shapefiles) stored in data repositories. Then, the web mapping viewer queries appropriate layers to the UMN WMS to visualize them. For example, Figure 2 shows a pair of layers retrieved by OpenLayers, one a remotely-fetched mosaic layer and the other (in green colour) the basin boundary for the study area in question served by the UMN WMS. So, as users introduce local data via the geoportal, they are dynamically loaded in the web mapping client in a distributed way. Also, the UMN Map server provides GML (Geographic Markup Language) data via WFS queries. Again, as in the case of WMS, the web mapping viewer permits the user to load GML data performing appropriate queries to the WFS service. Figure 3 shows a cloud of points representing location of temperature, precipitation and stream gauge sensors which are relevant for calibrating hydrological models. All these data are stored as GML files.

Regarding the execution of geoprocessing services, the WFS is a crucial component to better manage huge quantities of data (Granell *et al*. 2007). When a geoprocessing service execution is required, the prototype builds a XML-based query with input and output description on-the-fly. Such inputs can be either embedded in the query itself or passed by reference specifying a valid URL to remotely fetch such data. The latter is performed by the WFS in our geoportal.

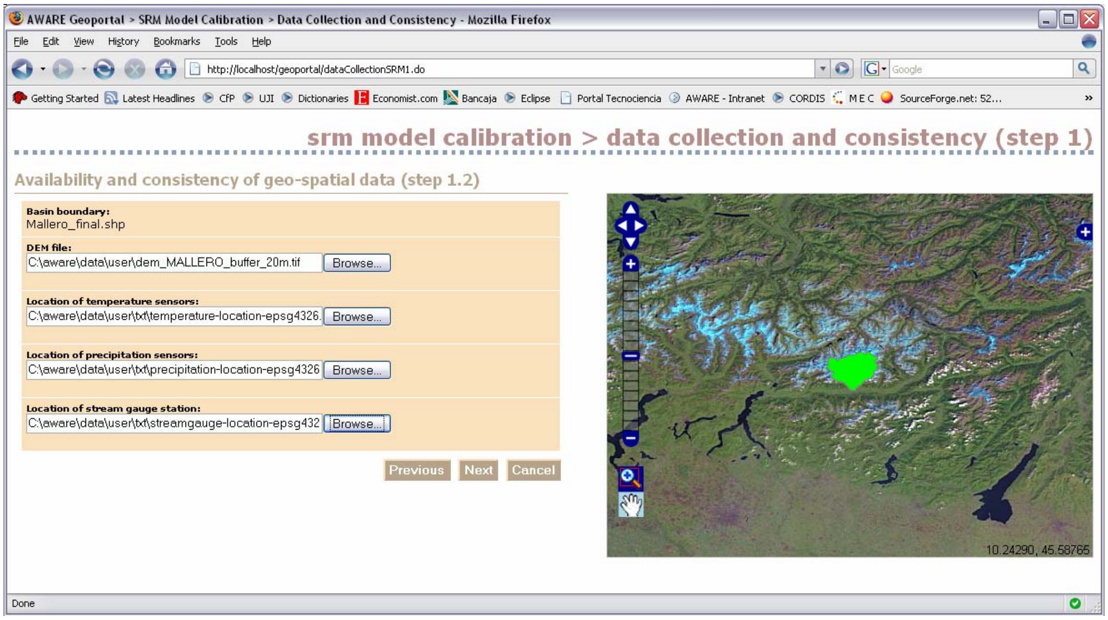

Figure 2. A web-based prototype interface to check the availability and consistency of geo-spatial data for running the SRM model.

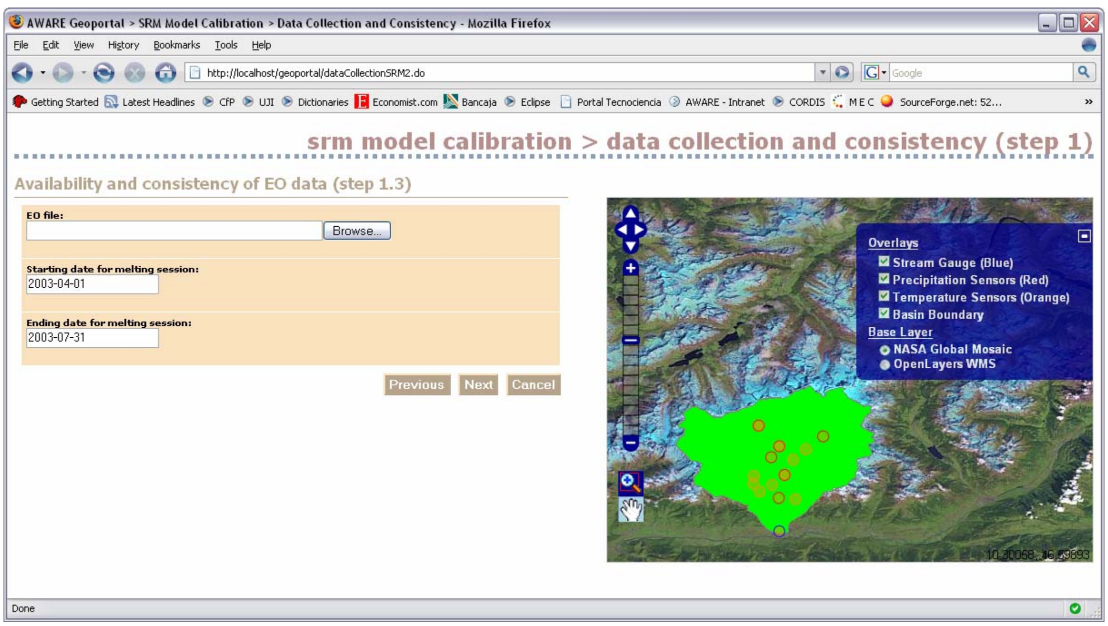

Figure 3. A web-based prototype interface integrating an open source component (web map viewer) to visualize data previously introduced.

Figures 2 and 3 depict geoportal user interfaces for data collection in the calibration phase of the hydrological model Snowmelt Runoff Model - SRM (Martinec *et al.* 1994). In particular, figure 2 shows an overview of the Alps as acquired by the MODIS (or Moderate Resolution Imaging Spectroradiometer) sensor on-board the TERRA (EOS-AM) platform. The basin overlaid is the Mallero river basin (319 km^2), in the Italian Alps, which is one of the AWARE project test site. (This basin has been studied for 3 melting seasons 2002, 2003 and 2004 using ground measurements collected from ARPA Lombardia and snow cover products resulting from the project processing of satellite MODIS data).

5 CONCLUSIONS

This paper presented the first steps towards the implementation of a web-base prototype for hydrological models using distributed geoprocessing services. While there are some stand-alone and web-based applications for hydrological modelling, this work presents a web-based prototype using entirely distributed data and processing services built on SDI-based architectures. Although our web-based Geoportal still offers limited functionality only, it is completely based on open source components and technologies, leading to a feasible and low-cost solution.

Future plans are concentrated on completing the implementation for the SRM Model, including linking and wrapping IDL (Interface Description Language) routines with geoprocessing services, providing IDL's data analysis tasks in a distributed and interoperable way.

ACKNOWLEDGEMENTS

This work has been partially funded by the EC in the framework of AWARE (A tool for monitoring and fore-casting Available WAter REsource in mountain environment), a STREP project in the FP6 GMES Initiative.

REFERENCES

Caldeweyher, D., Zhang, J. & Pham, B. 2006. OpenCIS—Open Source GIS-based web community information system. *International Journal of Geographical Information Science* 20 (8): 885-898.

Díaz, L., Costa, S., Granell, C. & Gould, M. 2007. Migrating geoprocessing routines to web services for water resource management applications. *Proceedings of the 10th AGILE Conference on Geographic Information Science. Aalborg, Denmark, May, 2007*: 1-9.

European Union 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)(2007-04-25) <http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_108/l_10820070425en00010014.pdf> [Visited: 2007-06-01].

Granell, C., Díaz, L. & Gould, M. 2007. Managing Earth Observation Data with Distributed Geoprocessing Services. *Proceedings of the IGARSS Conference*. Barcelona, Spain, July 2007 (in press).

Lemmens, R., Wytzisk, A., de By, R., Granell, C., Gould, M. & van Oosterom, P. 2006. Integrating Semantic and Syntactic Descriptions to Chain Geographic Services. *IEEE Internet Computing* 10(5): 42-52.

Martinec, J., Rango, A. & Roberts, R. 1994. *The Snowmelt Runoff Model (SRM) User's Manual*. Bern: University of Bern.

Rampini, A., de Michele, C:, Lehning, M., Blöschl, G., Brilly, M., Llados, A., Sapio, F. & Gould, M. 2006. AWARE: A tool for monitoring and forecasting Available Water REsource in mountain environment. *Geophysical Research Abstracts* 8 (10780).

Shut, P. & Whiteside, A. (eds.). 2005. OpenGIS Web Processing Service 0.4.0. *Open Geospatial Consortium*. <http://portal.opengeospatial.org/files/?artifact_id=13149&version=1&format=pdf> [visited: 2007-05-31]