Wosniok, Christoph; Lücker, Mathias; Lehfeldt, Rainer; Bensmann, Felix; Roosmann, Rainer

Standardized Web Processing of Hydro-Engineering Operations

ICHE 2014, Hamburg - Lehfeldt & Kopmann (eds) - © 2014 Bundesanstalt für Wasserbau ISBN 978-3-939230-32-8

Standardized Web Processing of Hydro-Engineering Operations

C. Wosniok, M. Lücker & R. Lehfeldt
Federal Waterways Engineering and Research Institute, Hamburg, Germany

F. Bensmann & R. Roosmann
University of Applied Sciences, Osnabrück, Germany

ABSTRACT: Hydro-engineering tasks usually involve high resource requirements. To deal with those adequately in distributed infrastructures, standardization is needed. The OGC Web Processing Service provides an interface for executing remote geospatial processes. Although the service has been published in 2007, there is no tool available to set up complex process chains such as hydro-engineering analysis workflows. We set up an environment called RichWPS to enable the assembling of such workflows based on online resources. WPS processes can be linked on a GUI, the workflow is transferred to a server in a self-developed description language. This RichWPS server interprets the description language script and executes involved local and remote WPS processes. As a typical analysis task in numerical simulations environments, we set up a case study to compare field data with modeled data. Individual, reusable processes have been developed for reading, harmonizing and comparing data, not only to test the environment, but also to identify further requirements.

Keywords: Web processing, RichWPS, OGC Services, Service composition, Hydro-Engineering operations, Spatial data infrastructures, Metadata

1 INTRODUCTION

Spatial Data Infrastructures (SDIs) have been established for visualizing and allocating spatial data in distributed networks for a few years now. Starting with basic geodatasets, SDIs increasingly provide data from other domains. Ocean-specific data is made available in initiatives such as SeaDataNet on the European level or in the Marine Data Infrastructure Germany (MDI-DE). These infrastructures are fostered by European directives such as the Marine Strategy Framework Directive (MSFD), demanding the delivery of status reports and data via technologies defined in the INSPIRE directive (Infrastructure for Spatial Information in the European Community). INSPIRE has a 12-year implementation plan and started in 2007. While the usage of web services and metadata as part of SDIs have been predefined from the start on, the development of tools for data providers is ongoing on different governmental levels.

The Open Geospatial Consortium (OGC) defines visualization and download web services such as Web Mapping Service (WMS) and Web Feature Service (WFS). These and other services are well established with tools for production and publication. Other services such as the Web Processing Service (WPS) (Schut, 2007) still require software tools for an easier integration in SDIs. Web processing is considered to be the next evolutionary step in SDIs (Brauner et al., 2009 and Kiehle et al., 2007). The integration of processing services in distributed service-oriented infrastructures extends the possible uses of available geodata, for example provided by other web services. Within the GIS community, several processes for geospatial data such as buffering and intersecting features or coordinate transformations have been developed (Schäffer et al., 2012). The WPS provides an interface with the basic operations getCapabilities, describeProcess and Execute to manipulate geodata with server-located processes. Given the possible varieties of geospatial processing, these operations leave room for the design of individual processes. This generic nature of the WPS enables its usage for applications outside the scope of fundamental GIS processes (Goodall et al., 2011 and Wössner, 2013).

Numerical simulations belong to the most demanding computing tasks. Extensive algorithms with complex data structure, high requirements for performance and data storage demand efficient solutions. The internet as a communication network offers possibilities to connect different processing resources for geospatial processing (Simonis et al., 2003) for example in cloud or grid computing environments (Kim and Tsou, 2013). Although an execution of a complete numerical simulation only based on internet resources is limited by the usually comprehensive requirements, single operations can still be provided as online resources.

The WPS could serve as a general interface to provide repetitive parts of the overall simulation process in distributed infrastructures (Foerster et al., 2010). Performing individual analysis tasks with simulation data using a WPS in an OGC web service environment shows the advantages of distributed computing and the paradigm of Software as a Service. Users can use software on the web without high own resource requirements or single processes as part of a repository can be reused in different service chains.

The usage of WPS has been proposed and implemented for example by Maué et al. (2011) or Castronova et al. (2013). However, the proposed approaches rarely consider the usability of setting up the process workflows or are focused on proprietary systems. A system, which would let users set up processes and process chains from different sources, is still not available.

As part of the research and development project RichWPS, we developed an environment to handle such complex WPS applications. A graphical user interface enables the design of process workflows and the RichWPS server interprets the model description in specifically developed description language. The RichWPS server executes the process chain and allocates process results for process chain users. A case study comparing measured field data and modeled data from a simulation system exemplarily shows the abilities of the environment: the processes read data, harmonize, compare and format data are set up in the environment and used in an interactive user interface.

2 THE RICHWPS ENVIRONMENT

Developing such complex applications within a WPS is still an ambitious task, requiring programming skills and knowledge of software libraries. The effort for integration in existing software frameworks and INSPIRE-compliant data infrastructures is high. To extend the options for dealing with composite WPS applications further tools are required. The RichWPS compositional environment developed by University of Applied Sciences Osnabrück and Disy information systems provides a user-friendly toolset for defining composite workflows based on distributed existing WPS processes. These workflows can be deployed on an adapted WPS server and thereby made available as common WPS processes. The toolset is able to orchestrate processing steps defined in the scenario and do the subsequent calls.

A directory service called SemanticProxy provides a search interface for processes and services available for modelling geospatial workflows. Users can include them into workflows by using the ModelBuilder, the central client-side application of the environment. The ModelBuilder is equipped with a diagramming user interface and connects to a WPS server that is extended with an orchestration engine. For transfer and execution, the static model is transformed into a sequential form represented through the specifically developed RichWPS Orchestration Language (ROLA). Figure 1 shows the components of the environment.

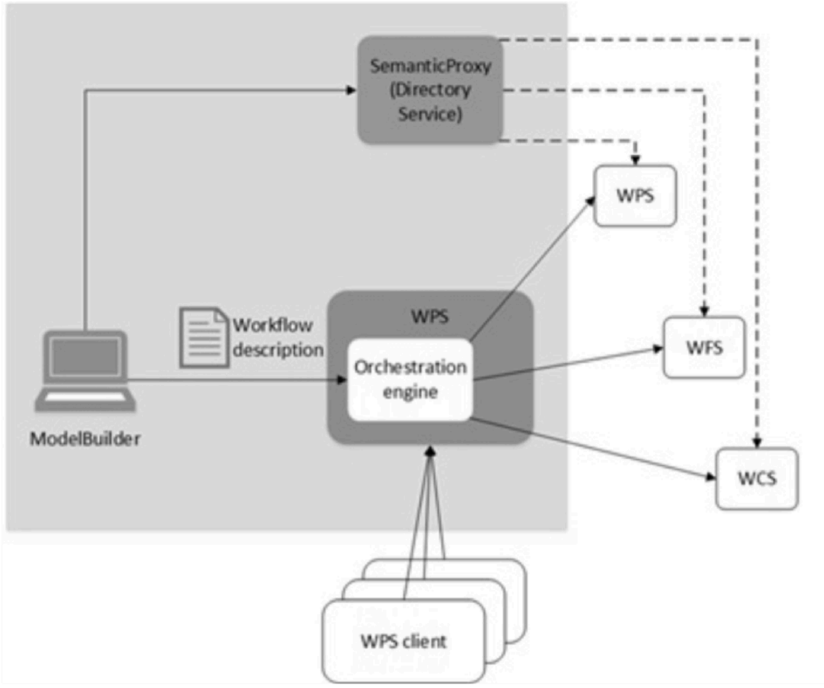

Figure 1. Setup of the RichWPS orchestration environment.

The RichWPS orchestration environment was developed to overcome shortcomings of existing solutions and to facilitate the application of WPS for domain experts and make it more attractive. One key concept for this is the orchestration of WPS and further OGC Web Services (OWS). Web service orchestration refers to the coordination of multiple services from a central point of view, different web services (WS) are executed in a certain order and form a business process (Peltz, 2003). WS and WPS orchestration has been subject of research for quite a long time (Foerster et al., 2010; Schäffer, 2009 and Jesus et al., 2012).

WPS and other OWS differ from common W3C Web Services (WS) in various ways e.g. their use of remote procedure calls or interface description. Orchestration techniques that came up together with Service Oriented Architectures (SOA) and WS can only be adopted with additional effort (Kiehle, 2006). Also, they are hard to operate, especially for non-experts (Jesus et al., 2012).

Possible redundancies in the interface description are a remarkable characteristic in the OWS concept making the concept variable. On the one hand OWS are designed with a self-describing interface, on the other hand they can, like common Web Services, be described using the Web Service Description Language (WSDL). WSDL is applied in some of the widespread forms of WS orchestration, for instance in BPEL (Ivanova, 2006). WPS makes WSDL documents available through their WS interface; individually managed processes are not regulated by standards in sufficient detail. This is especially the case for WPS with its process concept.

In WS the Simple Object Access Protocol (SOAP) is used to add numerous features, like they are described in the WS-* standards, for instance WS-Security or WS-Addressing. WPS adopts SOAP as optional to make these standards usable. In practice this increases the technology stack. Among others this affects the technical realization of orchestration, since possibly additional infrastructure may be required and through the increasing complexity new sources of errors emerge (Brauner et al., 2009).

These various standards lead to a technical overhead for domain experts. The RichWPS orchestration environment aims to avoid these problems by initially focusing on the chaining of a selected subset of OWS. At first WPS is taken into account, later on WFS, WCS and other services shall be considered. OWS are orchestrated by using their self-description interface instead of WSDL descriptions. This interface is, in contrast to WSDL, mandatory in every OWS which is a step towards interoperability.

Through limitation to OWS/WPS a new and simpler form of workflow description can be used that does not rely on WSDL. The key concept for this is the use of a domain specific language (DSL) called ROLA. ROLA describes the workflow as a sequence of calls of processes, variable assignments and so on. The language implements the properties of the services or WPS processes by its type and achieves a reduction of complexity. This means by knowledge of the type of an OWS, the handling of the service can be determined automatically. Workflows can be graphically modelled using the ModelBuilder and automatically be translated into a ROLA document. The document is then sent to the orchestration engine within the RichWPS server, where the described workflow is published as a common WPS process.

The ModelBuilder enables abstraction from the orchestration environment with its technical aspects and thereby offers a domain view of the workflow to the user. It follows proven principles of graphic editors such as Taverna. The graphical notation of the workflow model is based on a data and control flow diagram. Figure 2 shows a screenshot of the ModelBuilder with a workflow model in the central area, on the left hand side area are the available services and processes that can be retrieved via the SemanticProxy. On the right hand side detailed information about the modelling components are displayed and at the bottom, information regarding deployment, testing, profiling etc. is shown.

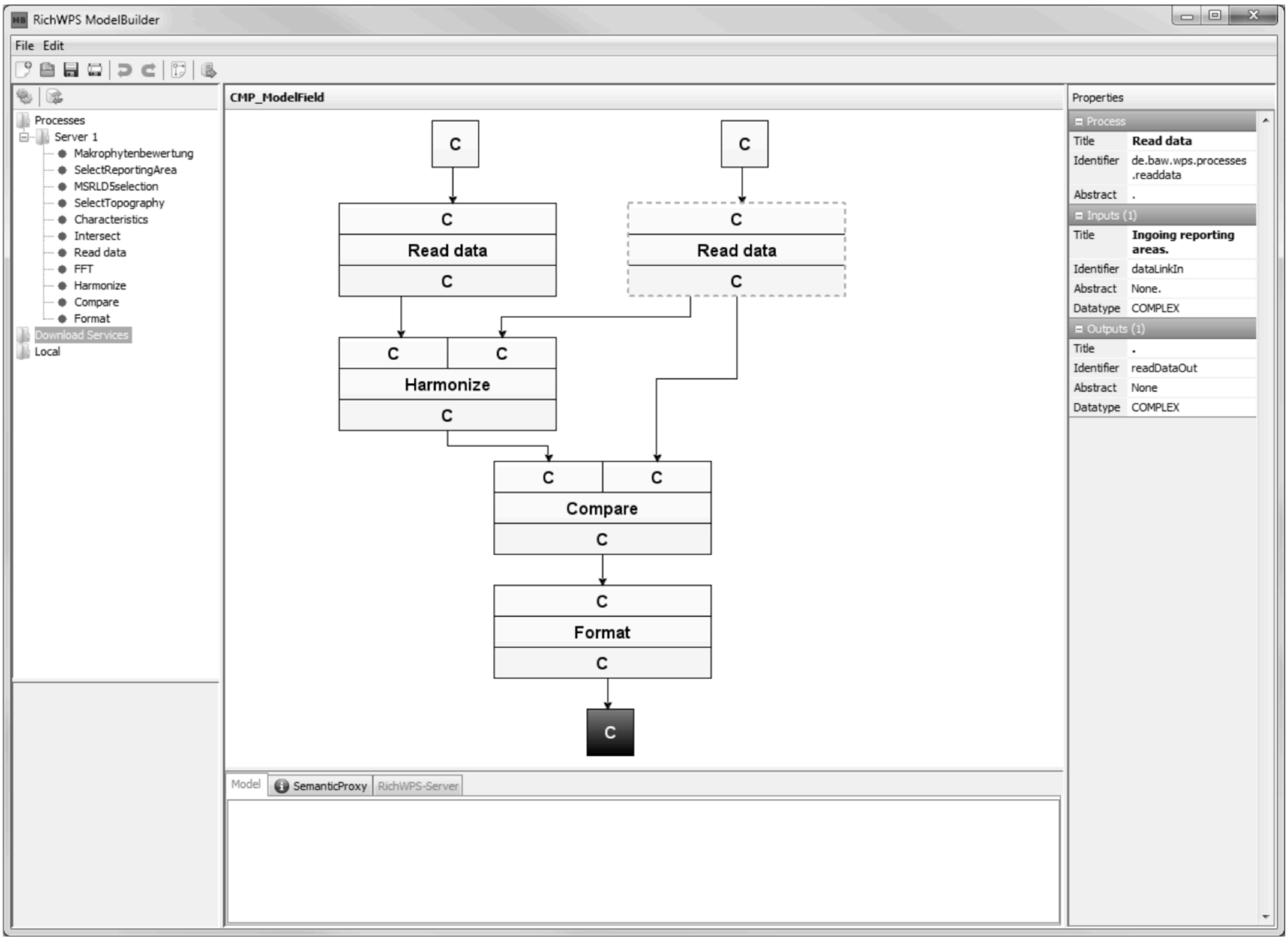

Figure 2. Screenshot of the ModelBuilder.

Besides modelling and workflow deployment the ModelBuilder can be used to test workflows in their later execution environment and to profile their execution time. The addition of further functionality like a debug function is conceptually possible but might not be able to be realized within the near future. This also applies for full or semi automatically optimization of workflows by replacing single processes in the workflows with better performing processes. The fundamental work for this, namely the realization of a WPS monitor, is already in progress. The next steps will be the extension with control structures as well as the support for further OWS like WFS and WCS and sometime in the future common WS.

Finally, the environment has been tested with a use case scenario and it is proved to be a sustainable concept. However, depending on the area of application efficient use is dependent on a coherent set of data types and basic processes.

3 COMPARING MODEL- AND FIELD DATA

To evaluate the newly developed components under near-productive conditions, we set up a two part case study based on a common simulation analysis task. We compare simulation results data sets with measured data at the same position from numerical models of water level time series (1-dimensional) and vertical profiles of water flow rates (2-dimensional). Three different kinds of processes have been developed for these scenarios and are provided as part of a WPS by the Federal Waterways Engineering and Research Institute (BAW): “Harmonize data sets”, ”Compute difference of two data sets” and “Analyze data, e.g. single side amplitude spectrum”. The modelled and measured data sets are provided by OGC web services that are called by the first process. The usage of OGC web services and standardized data exchange formats is important for maintaining interoperability when integrating a WPS process chain in a spatial data infrastructure (SDI).

In the 1D-scenario data was made accessible by a Sensor Observation Service (SOS), a specialized OGC web service to transmit time series data. Data sets retrieved by a SOS are coded in the OGC standard Observations & Measurements (O&M) XML format. Data sets from measurements and from modelling often come in different time steps, it is therefore important to make them comparable. To achieve this, the data sets are temporally harmonized in equidistant time steps by a spline interpolation in the first process. The next process computes the actual difference between the two harmonized time series. Finally, a single sided amplitude spectrum of both time series are calculated by a fast fourier transformation. Based on these treatments, a qualitative assessment of the modeled data is now possible. The elements of the BAW case studies are shown in figure 3.

As two dimensional data sets are compared in the second scenario, an OGC Web Coverage Service (WCS) is used to transmit the profile data sets. This service was developed to provide access to raster data (coverage). Like in the first scenario, the data sets must be harmonized. In this case it has to be an area harmonization due to different resolutions in x and y direction of the modelled and measured data sets (fig. 4). In a next process the difference of both data sets will be computed. Finally, statistical analysis can be applied in a last process, which is currently under development.

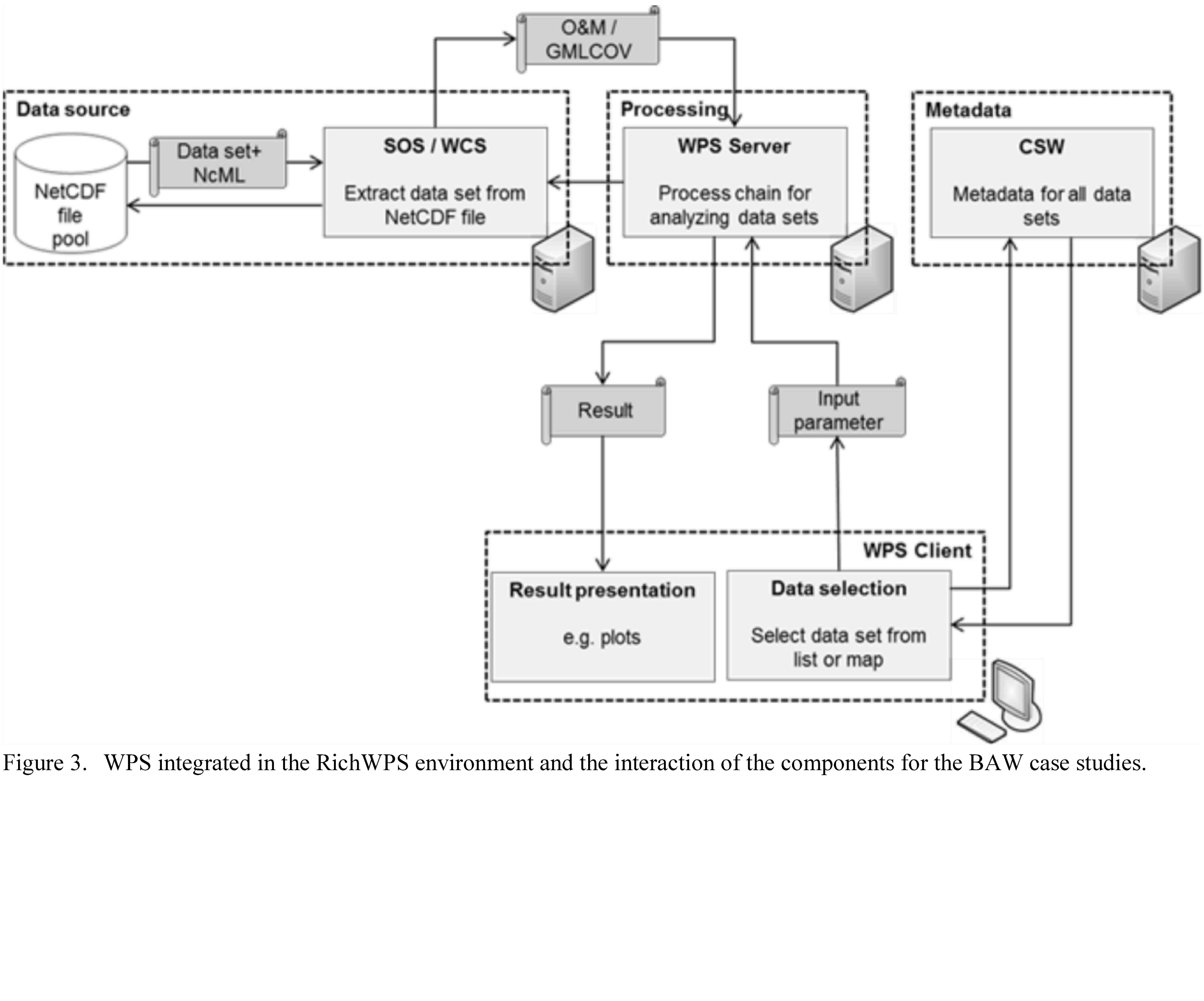

Figure 3. WPS integrated in the RichWPS environment and the interaction of the components for the BAW case studies.

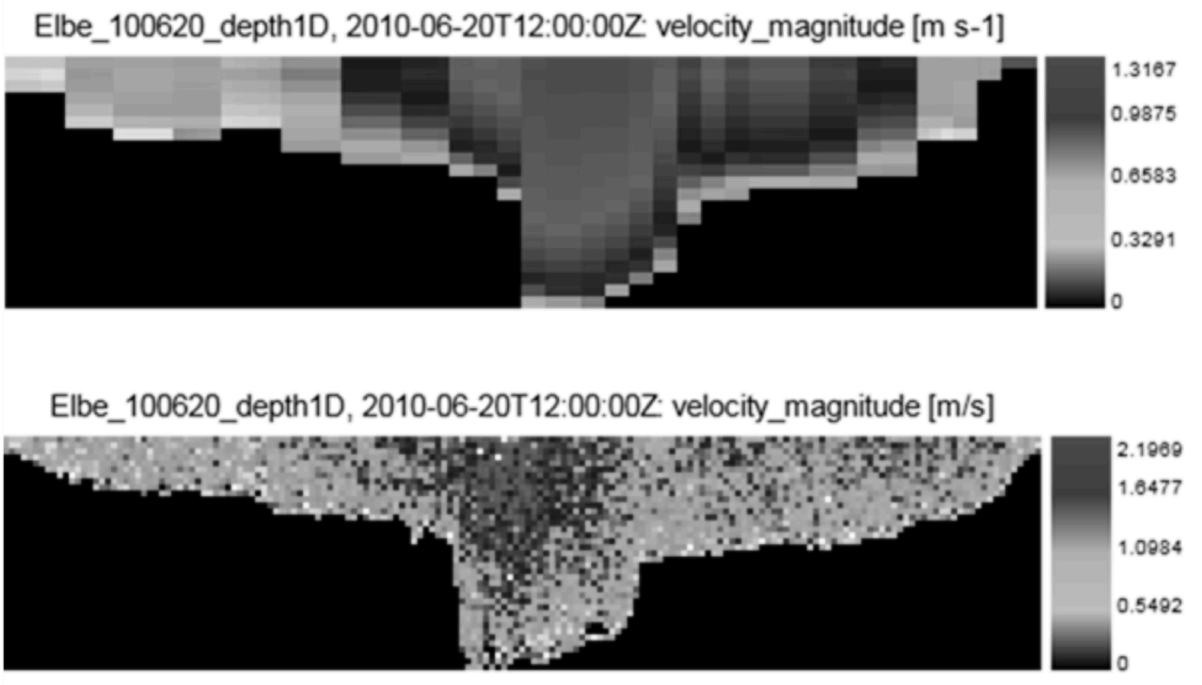

Figure 4. Top: modelled profile data set, bottom: measured profile data set.

For our case study, we deploy the above described processes in the RichWPS environment. In the ModelBuilder, processes can be chosen and connected without the need of programming skills and deployed on the RichWPS server (compare fig. 2).

To show the usability of the developed and deployed analysis processes, a testing environment has been set up as part of the case study at BAW. The interactive user interface enables the user-friendly choice of two sample data sets of modeled and measured water levels, the comparison and visualization of the results in different plots and the description of the involved elements along with relevant metadata (fig. 5). This hides the complexity of the processes modelled in the RichWPS ModelBuilder from the user and enables the choice of various input data sources like OGC WFS services in the future.

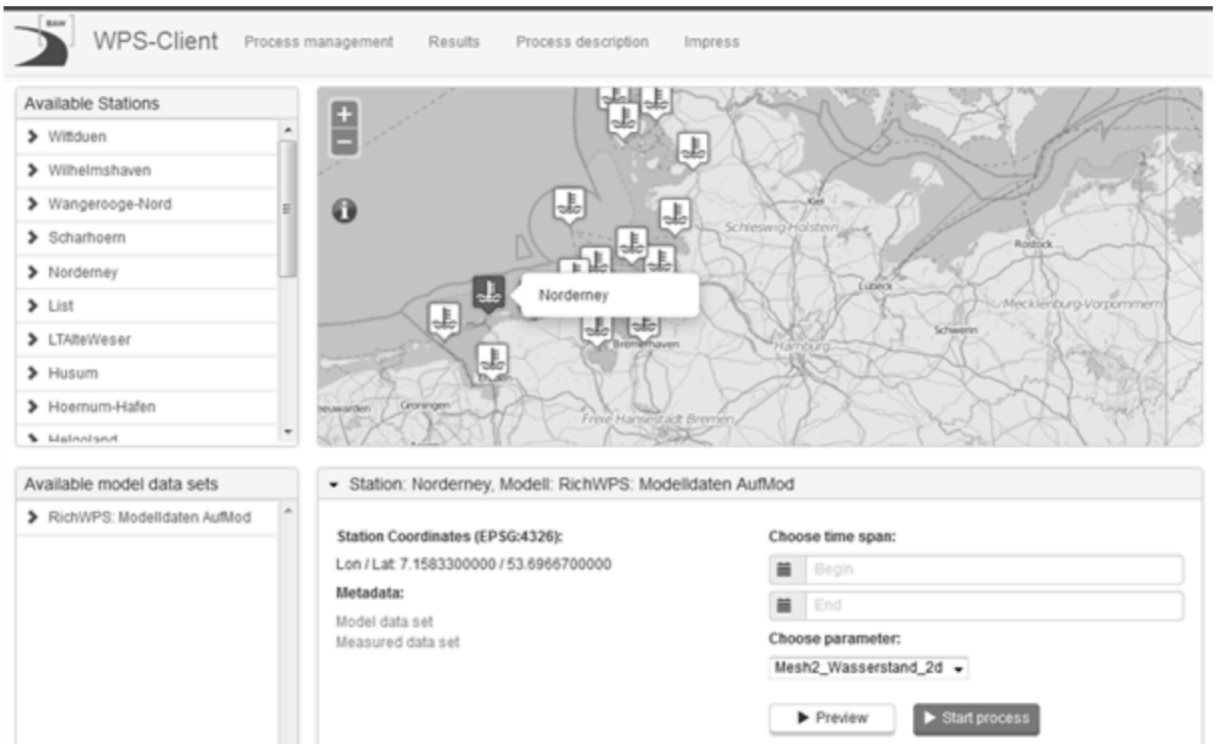

Figure 5. The interactive user interface for comparing measured data (stations) and model data

The final outcome of the WPS service chain should include information on the input data sets and the processing steps performed to reach the result. Input data metadata can be extracted when reading NetCDF files, which include standardized descriptions in NetCDF Metadata Language (NcML) (Nativi et al., 2005). To describe the process chain a custom web service was developed specifically for the case study, which generates a description of the used process chain modeled in SensorML (Botts and Robin, 2007), an XML-encoding. SensorML already provides a schema for the description of process chains. The metadata of a result dataset therefore includes, besides the actual result, links to the origin of the input datasets and a description of the process chain.

4 CONCLUSIONS AND FUTURE WORK

The RichWPS system provides a user-friendly interface for composing complex WPS chains from a pool of singular processes and applicable data. The more processes are available as part of an SDI to complement to and thus build more complex applications, the more the system is able to unfold its potential. The case study of comparing simulation results from numerical models with measured data shows the possibilities of analysis tools as orchestrated processes in distributed networks. The RichWPS compositional environment offers an essential benefit for the development of complex tasks in the interoperable environment of OGC web services.

WPS as part of SDIs, however, are still not very widely available (Lopez-Pellicer et al., 2012). The discovery of WPS requires effective mechanisms which are not available yet (Fitzner et al., 2011), the other, more important reason might be the generic nature of the WPS specification. While simple GIS-processes have been implemented with little effort, specialized process chains require custom clients for and effective operational use. Generic WPS clients can be used but cannot provide user friendly input selection and output presentation for any kind of data. One approach to handle this difficulty is to use OGC standardized XML formats for data in- and output, like we did in our use cases, for example GML (in WFS), GMLCOV (in WCS) or O&M (in SOS). With the use of standardized data formats the “degrees of freedom” of WPS can be reduced. Larger and mightier applications can be constructed and an integration of WPS in spatial data infrastructures based on OGC web services becomes easier.

The usage of common data formats simplifies the integration of online available WPS processes into data and program structures of institutions. But while the WPS technology is extensively researched, the belonging, and as part of SDIs essential, metadata for process chains is often neglected. To include the case study results and its metadata seamlessly in SDIs, we aim to develop ISO-compliant metadata sets in the future. While single data sets can be transformed from the NcML metadata, and ISO-compliant metadata profile for generic WPS Process compositions are yet to be developed. The provision of the process chain description could be a feature in a future version of the RichWPS-Server.

REFERENCES

Botts, M., Robin, A. (eds.) (2007). OpenGIS ® Sensor Model Language (SensorML) Implementation Specification. Open Geospatial Consortium (OGC® 07-000).

Brauner, J.; Foerster, T.; Schäffer, B., Baranski, B. (2009). Towards a Research Agenda for Geoprocessing Services. Sester, M., Bernard, L., Paelke, V. (eds.), Advances in GIScience. Proceedings of the 12th AGILE Conference. Berlin.

Castronova, A. M.; Goodall, J. L., Elag, M. M. (2013). Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. Environmental Modelling & Software, Vol. 41, pp. 72-83, doi: 10.1016/j.envsoft.2012.11.010.

Fitzner, D. (2011), Formalizing Cross-Parameter Conditions for Geoprocessing Service Chain Validation. Int. J. of Applied Geospatial Research, Vol. 2, No. 1, pp. 18-35, doi: 10.4018/jagr.2011010102.

Foerster, T.; Schäffer, B.; Baranski, B., Brauner, J. (2010, Geospatial Web Services for Distributed Processing. Zhao, P., Di, L. (eds.), Geospatial Web Services, pp. 245-286, doi: 10.4018/978-1-60960-192-8.ch011.

Goodall, J. L.; Robinson, B. F., Castronova, A. M. (2011), Modeling water resource systems using a service-oriented computing paradigm. Environmental Modelling & Software, Vol. 26, No. 5, pp. 573-582, doi: 10.1016/j.envsoft.2010.11.013.

Ivanova, E. (2006), Orchestrating Web Services – Standards and Solutions. Proceedings of Scientific Conference Mathematics, Informatics and Computer Science, V. Tarnovo, pp. 137-142.

Jesus, J. de; Walker, P.; Grant, M., Groom, S. (2012), WPS orchestration using the Taverna workbench: The eScience approach. Computers & Geosciences, Vol. 47, pp. 75-86, doi: 10.1016/j.cageo.2011.11.011.

Kiehle, C. (2006), Business logic for geoprocessing of distributed geodata. Computers & Geosciences, Vol. 32, No. 10, pp. 1746-1757, doi: 10.1016/j.cageo.2006.04.002.

Kiehle, C.; Greve, K., Heier, C. (2007), Requirements for Next Generation Spatial Data Infrastructures-Standardized Web Based Geoprocessing and Web Service Orchestration. Transactions in GIS, Vol. 11, No. 6, pp. 819-834, doi: 10.1111/j.1467-9671.2007.01076.x.

Kim, I.-H., Tsou, M.-H. (2013), Enabling Digital Earth simulation models using cloud computing or grid computing – two approaches supporting high-performance GIS simulation frameworks. Int. J. of Digital Earth, Vol. 6, No. 4, pp. 383-403, doi: 10.1080/17538947.2013.783125.

Lopez-Pellicer, F. J.; Rentería-Agualimpia, W.; Béjar, R.; Muro-Medrano, P. R., Zarazaga-Soria, F. J. (2012), Availability of the OGC geoprocessing standard: March 2011 reality check. Computers & Geosciences, Vol. 47, pp. 13-19, doi: 10.1016/j.cageo.2011.10.023.

Maué, P.; Stasch, C.; Athanasopoulos, G., Gerharz, L. (2011), Geospatial Standards for Web-enabled Environmental Models. Int. J. of Spatial Data Infrastructures Research, Vol. 6, pp. 145-167, doi: 10.2902/1725-0463.2011.06.art.

Nativi, S.; Caron, J.; Davis, E., Domenico, B. (2005), Design and implementation of netCDF markup language (NcML) and its GML-based extension (NcML-GML). Computers & Geosciences, Vol. 31, No. 9, pp. 1104-1118, doi: 10.1016/j.cageo.2004.12.006.

Peltz, C. (2003), Web services orchestration and choreography. Computer, Vol. 36, No. 10, pp. 46-52, doi: 10.1109/MC.2003.1236471.

Schäffer, B. (ed.) (2009), OGC OWS-6 Geoprocessing Workflow Architecture Engineering Report. Open Geospatial Consortium (OGC 09-053r5).

Schäffer, B.; Müller, M., Kadner, D. (2012), Bringing the Process to the Data - Introducing the WPS AppStore for Geoprocessing Functionality. Presentation, Prague, The Czech Republic, 21.05.2012.

Schut, P. (ed.) (2007), OpenGIS Web Processing Service. Open Geospatial Consortium (OGC 05-007r7).

Simonis, I.; Wytzisk, A., Streit, U. (2003), Integrating Simulation Models Into SDIs. Proceedings of The 6th AGILE International Conference on Geographic Information Science, Lyon.

Wössner, R. (2013), Untersuchungen zur praktischen Nutzbarkeit des OGC Web Processing Service - Standards. Master Thesis. Hochschule Karlsruhe - Technik und Wirtschaft, Karlsruhe.