Integrated Environmental Modeling — An SDI-based Framework for Integrated Assessment of Agricultural Information

Muhammad Imran, Raul Zurita-Milla and Rolf A. de By

Department of Geo-Information Processing,
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente,
The Netherlands
{imran19632, zurita-milla, deby}@itc.nl

ABSTRACT

The field of Spatial Data Infrastructure (SDI) has emerged with the spirit of providing 'Data as a Service' to overcome the technical and conceptual barriers in sharing spatial data. In integrated environmental modeling, software components like the GIS, the spatial DBMS, expert rules and analysis tools, and even the associated computational models, are re-used and combined into an environmental decision support system or integrated assessment tool. Our work aims to support the future design of open service platforms for such systems. By an *open service platform*, we mean a loosely coupled system that connects models with data sources with a high level of substitutability of components. Such a platform will improve the wider exploitation of environmental, spatial data offered by current SDI services, and will positively influence the robustness of current environmental models. For this paper, we wanted to deploy an optimization modeling system as a service, essentially with the philosophy of 'Software-as-a-Service' (SaaS). Here, we report on our first prototype, which achieves data harmonization loosely through a OGC-compliant wrapper implementation, and we discuss the prospect for more mature, probably INSPIREd, transformation services, in the context of our services framework for integrated assessment of agricultural information.

1 INTRODUCTION

Integrated environmental modeling, often requires to integrate (spatial) data and computational models from a variety of disciplines (e.g., related to physical, biotic, social, and economic environments) and at different scales, to understand and to solve complex societal problems that arise from the interaction of humans and environment, and to contribute in this way to establishing the foundation of sustainable development, to inform policy and to support decision-making (Rothman, 1997, Parker, 2002). Regarding computational models, two categories of modeling approaches have been identified: predictive and goal-oriented modeling (Jakeman, 2008). Predictive models focus on mechanistic understanding of the underlying physical processes in environmental systems. These models are normally integrated with goal-oriented models, to provide certain inputs that pertain to the underlying environmental systems. Goal-oriented approaches include environmental optimization (multicriteria or single criterion) modeling based on system analysis that is required to explore and access uncertain future states of the system and to explore consequences of a range of resource-use combinations mostly to support integrated assessments and decision making. Under this category, modeling systems have been designed for large-scale and complex problems; specifically to develop models to support environmental-economic research. To this end, the General Algebraic Modeling System[1] (GAMS) provided: separation of the data from the models and separation of the modeling language from the execution system[2]. Since environmental modeling involves spatial data sets collected at different spatial scales, modeling frameworks should be spatially-aware to connect such heterogenous scales. For instance, the recently developed SEAMLESS modeling framework links predictive and optimization models using an integrated database of agricultural (spatial) data across the European scales and disciplines (van Ittersum, 2008). SEAMLESS adopted the OpenMI[3] standard

[1] http://www.gams.com/
[2] http://www.optimizationservices.org/
[3] http://www.openmi.org/

interface for linking GAMS-based agricultural optimization models and spatial datasets. However, this interface requires language-specific implementation for a compliant model component, technically, using a software framework like .NET or J2EE (Knapen, 2009). Consequently, the dependencies may be difficult to resolve when using the model elsewhere.

Using Service Oriented Architecture (SOA) and XML-based standards, the idea of 'Model Web' recently started offering distributed modeling systems as interoperable, loosely-coupled web services using standard interfaces and information models (Geller, 2008). In the same direction, the Optimization Services (OS) research project[4] proposed standards for a distributed optimization modeling system, in which components like language software, solver software, and data that are used to generate a model instance might reside on different machines using different operating systems (Fourer, 2010). However, the proposed standards only address economic models for management science and operations research applications. Therefore, these standards do not have the necessary wider scope for integrated environmental-economic models, in the context of GI services.

Web-based GIS has emerged to offer loose-coupling of spatial data and processing, ideally in an open and flexible SOA, using XML-based standards. Following this, Spatial Data Infrastructures (SDIs) mostly disseminate spatial data sets in an efficient and flexible way. These spatial data sets are offered as distributed web services, based on Open Geospatial Consortium (OGC) standards. To provide environmental models as distributed web services, recent research addresses 'model as a service' (MaaS) to incorporate legacy systems in service chains. Following this direction, as a specialized form with special attention for environmental optimization problems, the provision of an 'optimization model as a service' will potentially enable wider exploitation of environmental, spatial data offered by current SDI services. Therefore, the purpose of this paper is to study the feasibility of MaaS, to make environmental-economic optimization models spatially-aware by loosely coupling them with web-based GIS.

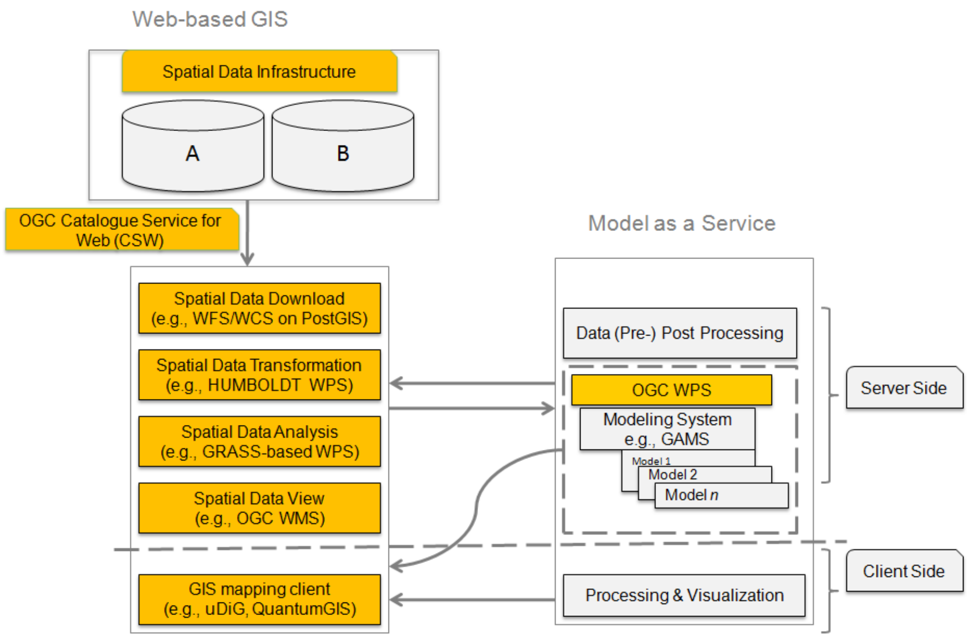

Figure 1: Interacting MaaS with web-based GIS

. To offer a model as a reusable service and to allow chaining it with other environmental services, various standardized approaches have been proposed in the ongoing ENVISION[5] research project, including the OGC WPS standard. This standard interface has recently been implemented for

[4] https://projects.coin-or.org/OS/
[5] http://www.envision-project.eu/

predictive models in the hydrological domain. For instance, Fitch et al. (2009) and Ferencika et al. (2010) implemented the interface to make available predictive hydrological models to a wider range of clients, and to make them spatially-aware. In this paper, we report on a WPS interface implemented to offer a MaaS. To this end, we selected GAMS to provide its model as a service, at an SDI node. This MaaS is then tested to run a bio-economic optimization model as a service, to conduct agricultural integrated assessment for a study area in Burkina Faso. This service allows pre- and post-processing of the spatial data sets offered by third-party agricultural services, to initialize a number of inputs/parameters of the optimization model service at a household farm location. Wrapping optimization modeling systems with an OGC standard interface provides new opportunities of incorporating SDI services, which are beyond those found in the conventional integrated frameworks. For instance, a WPS interface can encapsulate existing GIS system functionality like that of GRASS, SEXTANTE, HUMBOLDT, etc., and legacy modeling systems like GAMS, which can then be orchestrated to operate over spatially referenced data offered by other services (Figure 1). The descriptions of realistic available data and services in the region might be published in SDI catalogues (OGC CSW). OGC WFSs are used to download the available spatial datasets. The data models of these WFSs act as source schema for the transformation services that are provided in the development of an OGC WPS wrapper for the model.

2 IMPLEMENTING A MAAS

For this work, we selected the GAMS model to be offered as a MaaS, using the WPS interface. This interface facilitates discovery of and binding to the published processes. The standards provide wrapping for simple geospatial processes (e.g., a buffer calculation) as well as complex computational models that operate on spatially referenced data (OGC, 2007). Since there are no restrictions on format, network location, platform, and number of data inputs/outputs to/from a geospatial process, an environmental model can be provided as a WPS process, and then be linked with other spatial data services (Figure 1). (Figure 2) shows the deployment of WPS processes on the server side and the interaction pattern for the GAMS modeling system using the uDig client.

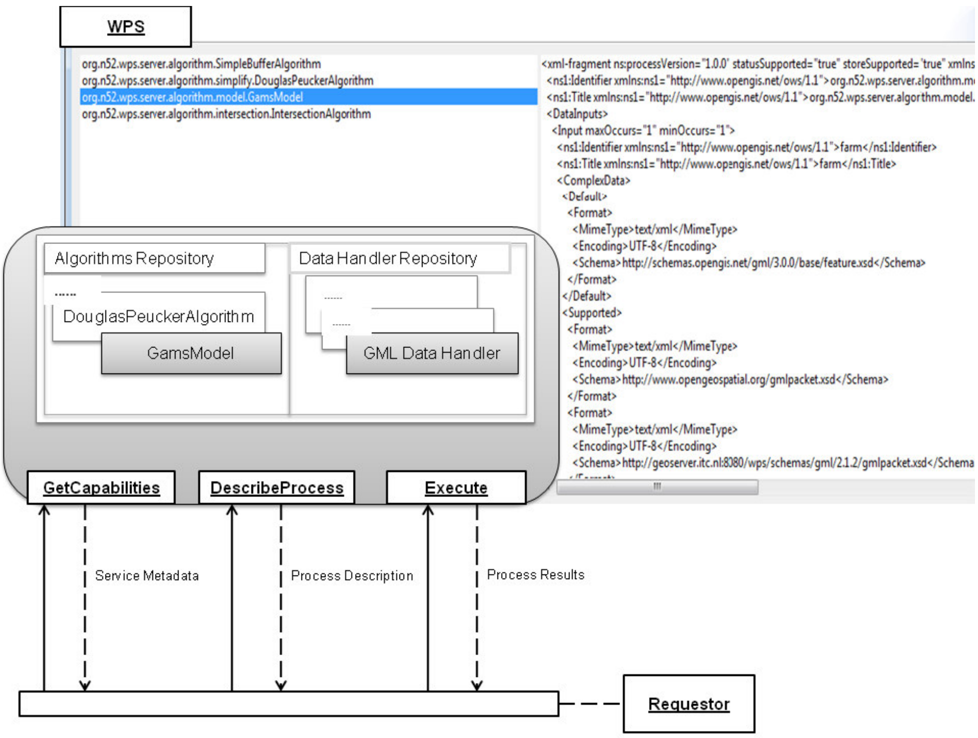

Figure 2: Accessing GAMS modeling system as a WPS process from a uDig client

Following the standard web service style operation, a list of available processes, is included in the capabilities document whilst the details of each process are accessed through a `DescribeProcess` operation. Furthermore, domain-specific processes like spatio-agricultural models require quite detailed process description, for instance, to achieve semantic interoperability of various model inputs/parameters and to describe the model quality like model assumptions and corresponding spatio-temporal scales, and so forth. For this, a WPS application profile may define, for instance, the accepted values and formats of spatial data sets involved in the modeling problem in hand. This may be utilized by web service registries for efficiently interacting with other services in SDI.

To interact with the modeling system, GAMS provides Java Native Interface (JNI) and GDX API to be implemented by Java-based external applications. These interfaces are used in various components of web-based implementation of the optimization modeling framework (Figure 3). In this framework, WPS and WFS standards are used to offer the MaaS. Implementation of the process scenario described in this research is based on the GeoServer application server (GeoServer, 2010) and the 52°North WPS framework (52°North, 2007). Interactions among various modeling components are explained in the next paragraphs.

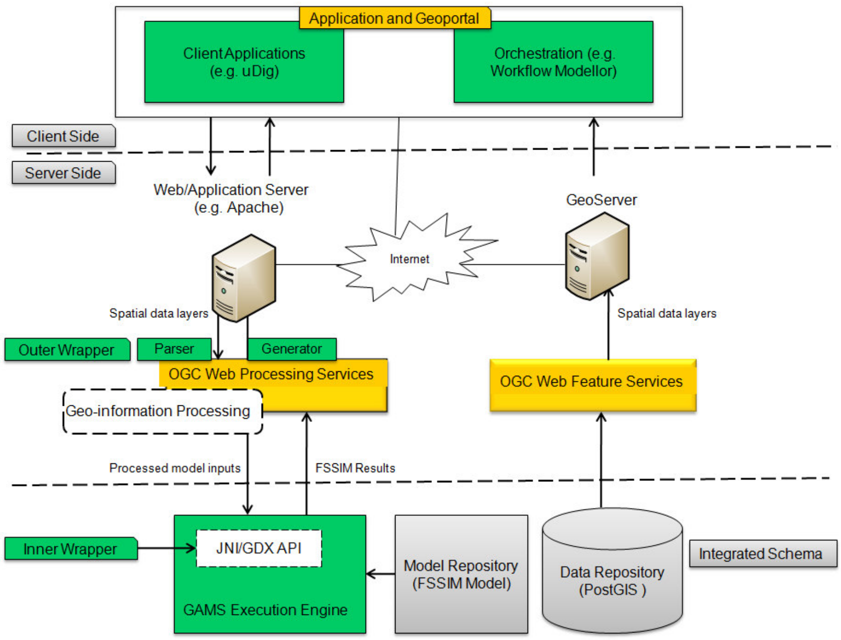

Figure 3: Wrapping GAMS modeling system with a standard interface to interact it with other SDI services.

The GAMS modeling system allows to separate input data from the models. To provide model inputs/parameters at run-time, the diverse spatial data sets obtained from the SDI can be initialized with the MaaS using WFS services. In this way, the adopted loose-coupling of GI services and environmental computational models has tangible benefits to both GIS users and modelers, however, their difference in data models is the main difficulty in moving towards a more scientifically rigorous approach to their integrated use (Bian, 2007, Brimicombe, 2009). The data models in GIS represent the geographical space, the objects contained therein, and their spatial relationships, whereas, environmental computational models are more concerned with spatial processes, their states, and throughput of quantities. Consequently, it requires, for instance, to parse, process, and spatially analyze, one or many spatial data services to initialize a model input/parameter. Current SDI initiatives, like INSPIRE[6], do not enforce the organizations to change and store existing data sets, instead, the data sets may be transformed or mapped on-the-fly via transformation services. For the

[6] http://inspire.jrc.ec.europa.eu

integrated use of diverse tools and data sets, the transformation services may comply with an integrated schema, in which GIS and environmental models stand in relation to each other. Such integrated schema may be maintained in the database, for example in case of the SEAMLESS framework (Janssen, 2009). In an SDI framework, the integrated schema may be provided as a WPS application profile for a modeling system service. In this work, the mapping is provided in the development of inner and outer wrappers inside the standard interface implementation. However, future work should migrate such mapping to the external loosely-coupled transformation service, probably INSPIREd, in an *open service platform* for environmental integrated modeling.

Following the standard WPS web service style operation, the GAMS MaaS can be invoked by another service or a client application (Geoportal, Workflow Modellor, Udig, and so forth). Conversely, for the model initialization or delivering the model output, the GAMS MaaS can invoke other SDI services. For this, the outer wrapper handles requests and responses of the MaaS. The outer wrapper provides geospatial data manipulation (e.g., using GeoTools) to parse the spatial data with which the services work. The outer wrapper uses various interfaces to parse XML/GML data input to a feature (collection) object and to generate XML/GML outputs from a feature (Feature Collection) object. Various parsers are provided with the 52°North WPS implementation framework that are used in outer wrapper to handel the input data (e.g., with GML) offered by spatial data services. Similarly, various output transformers (generators) are used for turning the result into appropriate responses. WFSs are handled by the outer wrapper for parsing feature objects that can then be manipulated for model inputs. For a location, the outer wrapper obtains iterators for traversing all location-specific features provided by various WFSs. Finally, it passes all geospatial feature objects to the inner wrapper.

The inner wrapper is closer to the optimization system and it changes the model states at simulation run-time. Moreover, inner wrapper provides a communication stack for coordination between inner and outer wrappers. Its main tasks are to analyze and decompose the geospatial feature objects into model parameters. GDX API together with the JNI interface for the GAMS are extensively modified to make them compatible with 52º North WPS process. The modified versions are then employed to develop the inner wrapper for handling model simulation. For a given location, feature objects provided by a WFS are decomposed into model parameters which can then be initialized to run the model. Finally, the simulation results are delivered to a generator in the outer MaaS wrapper to generate a response to the client application.

3 TESTING THE IMPLEMENTED MAAS

To test the implemented MaaS, the FSSIM (Farm System Simulator) model was selected. It is an optimization model developed using GAMS in the SEAMLESS Integrated Framework (van Ittersum, 2008). FSSIM belongs to a family of bio-economic models, typically developed to find an optimal allocation of agricultural resources to satisfy a given set of objectives under certain environmental policy constraints (Louhichi, 2010). FSSIM generates farm responses based on a farming situation, which is represented as a linear combination of production activities subject to constraints over available farm resources. Various model inputs pertaining to a farm location are vital, including bio-physical data (weather, soil, and topography), management aspects (irrigation, fertilization, and pest control), government policies (amount of fertilizer allocation, time frame of distribution, subsidy provision or access to credit), and economic data (market price and demand for a commodity). At this moment, we investigate (1) how to calibrate the FSSIM model for subsistence farming in Burkina Faso, and (2) how to offer the FSSIM model as a web processing service that can be employed in defining geoprocessing workflows, either locally or through orchestrating further services. In this paper, we report on the latter, that is, an offering of the FSSIM model as a web processing service along with other spatial data services. The spatial data services provide model inputs pertaining to a farm location to find out an optimal and feasible solution for an objective function. One such objective could be to maximize the production of a household farm in Burkina Faso, after securing enough food for all household members.

To offer location-based data inputs from diverse disciplines to the FSSIM model, various agricultural services are implemented, including those representing the farm, labor, parcel, price, and production inputs (Table 1). For a given farm location, these spatial data services feed various inputs/parameters to the FSSIM model WPS. For instance, FSSIM model inputs pertain to labor and production e.g., `labreq(i,aez)`, `yield(f,I,aez)`, `inputs(f,I,aez)` depend on the agri-ecological zone (soil type, weather, and so forth), farming type, etc., at a farm location. The spatial data services offer feature objects and associated thematic attributes to the outer wrapper which turns them into the FSSIM model inputs, which are then passed to the inner wrapper. The run-time initialization of parameters for the FSSIM model is done from the inner wrapper. Finally, to generate the WPS response to the client, the simulation results are delivered to a generator in the outer wrapper. For example, (Figure 4) shows the recommended cultivated area for the commercial sorghum at a household farm location to meet the aforementioned objective function. Following this objective function, (Table 2) describes other output decision variables that are defined in the FSSIM model.

Table 1: FSSIM model inputs offered by the agricultural services, at a household farm location in Burkina Faso

Service	FSSIM inputs	Description
Farm	*f*	household farm *f* at a location
	farm size(f)	size of the household farm *f*
	capstart(f)	starting capital or value of livestock of the household farm *f*
	famlab(f)	available family labor (hours) to the household farm *f*
Labor	*i*	crop *i*
	aez	agro ecological zone *aez* (soil type, weather, etc.,) at the household farm *f* location
	labreq(f,i,aez)	labor required (hours) per crop *i* for land quality *aez* of the household farm *f*
Price	*price(f,i)*	price of crop *i* at the household farm *f* location (CHF)
Parcel	*cropland(f,aez)*	quantity of cropland (ha) of quality *aez* available to the household farm *f*
	totland(f,aez)	total land (ha) of quality *aez* available to the household farm *f*
	initland(f,i,aez)	initial land of quality *aez* allocated (ha) per crop *i* by the household farm *f* (observed crop allocation maps)
Production	*inputs(f,i,aez)*	non-labor costs (N or K or pests) for producing crop *i*, depending on the land quality *aez* and farming type of the household farm *f*
	yield(f,i,aez)	yield (kg/ha) per crop *i* of the household farm *f* depending on land quality *aez* (yield maps)

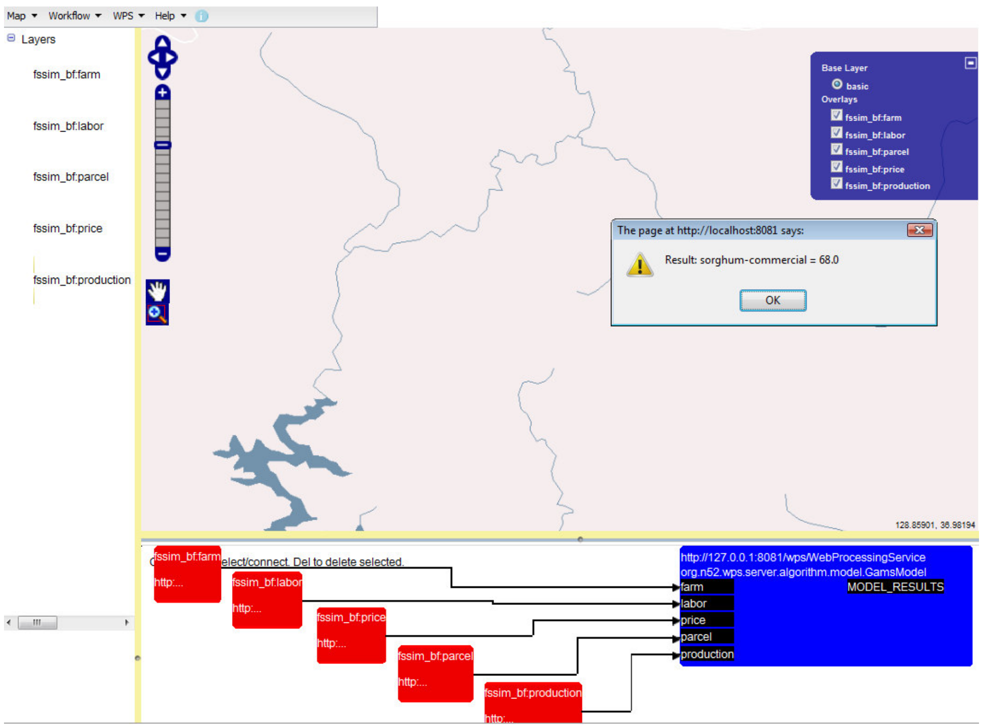

Figure 4: An orchestrated service for the GAMS-based bio-economic farm model (FSSIM) with other agricultural services in an SDI framework, for household farms in Burkina Faso

Table 2: FSSIM model output decision variables

Objective function	FSSIM decision variables	Description
To maximize the production of a houshold farm in Burkina Faso, after securing enough food for all household members	*A(f,i,aez)*	land allocated (ha) to crop *i* per land quality *aez* of the houshold farm *f*
	C(f,i,aez)	capital assigned to crop *i* per land quality *aez* of the household farm *f*
	P(f,i)	food in the form of crop *i* purchased from market by the household farm *f*
	TLABOUR(f,i,aez)	hired temporary labor (hours) for crop *i* per land quality *aez* by the household farm *f*

Various FSSIM model scenarios have been successfully tested using Workflow Modeller[7]. The workflow modeller is a client application that provides a visual environment for constructing flexible geoprocessing workflows. The WPS service for the FSSIM optimization model is chained and equipped with WFS services for model inputs for a given farm location (Figure 4). For a household farm, the optimized solutions can be delivered in a variety of formats, both generic (i.e., with GoogleEarth KML, GML) and domain-specific (i.e., with agroXML).

4 CONCLUSION AND FUTURE WORK

In this paper, we explored how an optimization modeling system offered as a service can take benefit of various spatial data services in an SDI framework. To this end, we initially set out to provide a standard wrapper over GAMS (General Algebraic Modeling System) to allow its optimization models to be exposed as OGC web processing services. Model inputs (i.e., parameters/variables) do not need to be known a priorily, i.e., the data inputs are not hard-coded in the model. For a location, the model inputs/parameters can be initialized on-the-fly with the orchestration of alternative spatial data services offered by national SDI. Since SDI has emerged as a technology to overcome technical and conceptual barriers to support sharing of existing spatial data sets, this prototype framework has potential for parametrization, calibration, sensitivity analysis, and so forth, of the optimization models.

In the current implementation, only the GAMS-based bio-economic optimization model is offered as a service in the agricultural domain, which is then orchestrated with other OGC spatial data services. Future work should tackle chaining various predictive and optimization models as services, to provide an open service platforms for integrated environmental modeling. Furthermore, for the model parameterizations in SDI framework, scale and uncertainty associated with (geospatial) data sets and models will be incorporated. For instance, spatial data services should be replaced automatically or transformed otherwise, if certain parameters/variables do not meet the model input quality requirements. For this, a WPS interface should make explicit the semantics of model input data and parameters. Moreover, this interface should explicitly describe model functionality based on model assumptions, model spatial and temporal extent and support units, and so forth. This is quite difficult to provide with present WPS, since it is rather generic and does not provide options to describe complex processes in detail. Presently, we deployed an integrated schema to describe FSSIM model inputs and parameters pertain to the model assumptions. Using this integrated schema as the target data model, the required transformations are embedded inside the WPS wrapper. Ideally, the transformation services should not be provided inside the wrapper. Therefore, to make explicit the particular descriptions of a complex process, the development of WPS application profiles is encouraged by OGC (2007). Following this, the metadata for the integrated schema can be described in a target WPS application profile, and a transformation service could be in place if the source data models do not meet the quality of the target profile. Following this direction, future work should focus on examining solutions that will (1) close the conceptual gap in data type definition of model inputs/outputs and spatial data set, and (2) describe model semantics (for example, model assumptions) to enable automatic composition of OGC data and model services in an SDI framework.

[7] http://52north.org/

In this context, this prototype framework will be extended to provide generic web services for various spatio-agricultural models that can be easily scaled/parameterized at a location.

ACKNOWLEDGEMENTS

The authors are thankful to Sander Janssen (SEAMLESS research group, Wageningen university) for providing the FSSIM model. We are also thankful to Lutz Westermann (GAMS Software GmbH) for providing the GAMS evaluation license for this research and the further support provided. Finally, we are thankful to Bastian Schäffer (52°North) for the WPS platform-specific support.

BIBLIOGRAPHY

[1] Bian, L., Object-Oriented Representation of Environmental Phenomena: Is Everything Best Represented As An Object? *Annals of the Association of American Geographers*, 97(2): 267–281, 2007.

[2] Brimicombe, A., *GIS, Environmental Modelling and Engineering*. Taylor & Francis Group, second edition, 2009.

[3] Ferencika, I., Niemia, T., Jolmaa, A., On Site Environmental Modeling and Monitoring: the Nordic Scenario in Hydrosys Project. In D.A. Swayne, W. Yang, A. Voinov, A. Rizzoli, T. Filatova (eds.). *2010 International Congress on Environmental Modelling and Software, International Environmental Modelling and Software Society (iEMSs)*, 2010.

[4] Fitch, P., Bai, Q., A Standards Based Web Service Interface for Hydrological Models. In *18th World IMACS / MODSIM Congress*, pages 873–879, 2009.

[5] Fourer, R., Ma, J., Martin, K., OSiL: An Instance Language for Optimization. *Computational Optimization and Applications*, 45: 181–203, 2010.

[6] Geller, Gary N., Melton, F., Looking forward: Applying an Ecological Model Web to Assess Impacts of Climate Change. *Biodiversity*, 9(3&4): 79–83, 2008.

[7] GeoServer., http://geoserver.org/display/GEOS/Welcome, Accessed on December 30, 2010, 2010.

[8] Jakeman, Anthony J., Voinov, Alexey A., Rizzoli, Andrea E., Chen, Serena H., Developments in Integrated Environmental Assessment. Elsevier, 2008.

[9] Janssen, S., Andersen, E., Athanasiadis, Ioannis N., van Ittersum, Martin K., A Database for Integrated Assessment of European Agricultural Systems. *Environmental Science & Policy*, 12(5): 573–587, 2009.

[10] Knapen, Rob M. J., Athanasiadis, Ioannis N., Jonsson, B., Huber, D., Wien, J., Rizzoli, Andrea E., Janssen, S., Use of OpenMI in SEAMLESS. In *AgSAP Conference 2009, Egmond aan Zee, The Netherlands*, 2009.

[11] Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Lansink, A., van Ittersum, M., FSSIM, A Bio-economic Farm Model for Simulating the Response of EU Farming Systems to Agricultural and Environmental Policies. *Agricultural Systems*, 103(8): 585–597, 2010.

[12] OGC., Web Processing Service Specification. http://www.opengeospatial.org/standards/wps, Accessed on December 30, 2010, 2007.

[13] Parker, P., et al., Progress in Integrated Assessment and Modelling. *Environmental Modelling & Software*, 3(17): 209–217, 2002.

[14] Rothman, Dale S., Robinson, John B., Growing Pains: A Conceptual Framework for Considering Integrated Assessments. *Environmental Monitoring and Assessment*, 46(1): 23–43, 1997.

[15] van Ittersum, M., et al., Integrated Assessment of Agricultural Systems — A Component-based Framework for the European Union (SEAMLESS). *Agricultural Systems*, pages 150–165, 2008.

[16] 52°North., 52 north. http://www.52north.org/wps, Accessed on December 30, 2010, 2007.