Software Platform Interoperability Throughout EnviroGRIDS Portal

Gorgan, Dorian; Bacu, Victor; Mihon, Danut; Stefanut, Teodor; Rodila, Denisa; Cau, Pierluigi; Abbaspour, Karim; Giuliani, Gregory; Ray, Nicolas; Lehmann, Anthony

How to cite

GORGAN, Dorian et al. Software Platform Interoperability Throughout EnviroGRIDS Portal. In: IEEE journal of selected topics in applied earth observations and remote sensing, 2012, vol. 5, n° 6, p. 1617–1627. doi: 10.1109/JSTARS.2012.2211580

This publication URL: https://archive-ouverte.unige.ch/unige:23042

Publication DOI: 10.1109/JSTARS.2012.2211580

Software Platform Interoperability Throughout EnviroGRIDS Portal

Dorian Gorgan, *Member, IEEE*, Victor Bacu, Danut Mihon, Teodor Stefanut, Denisa Rodila, Pierluigi Cau, Karim Abbaspour, Gregory Giuliani, Nicolas Ray, and Anthony Lehmann

***Abstract*—The Black Sea Catchment area is well known for subjects such as ecologically unsustainable development or inadequate resource management. The EnviroGRIDS project addresses these issues by using emerging information technologies. The enviroGRIDS Web Portal allows the users to access the geospatial functionality given by Web infrastructure, and to high power computation resources given by Grid infrastructure. The Black Sea Catchment Observation System portal provides a single point of access to the enviroGRIDS applications and tools. Both the vertical and horizontal interoperability are available between the platforms and applications throughout the portal. The horizontal interoperability is accomplished through services, meaning the applications are working together by the exposed services. The vertical interoperability is supported by the communication between the layers of end user applications, Web infrastructure, and Grid infrastructure. The basic solution of interoperability is accomplished by services, messages, and data. The paper highlights the solutions developed by the Black Sea Catchment Observation System portal to support various types of interoperability between the modules of geospatial data management, hydrological model calibration and running, satellite image processing, spatial data visualization, and virtual training center.**

***Index Terms*—Grid computing, interoperability, OGC services, SWAT calibration.**

I. Introduction

ENVIROGRIDS (Building Capacity for a Black Sea Catchment Observation and Assessment System supporting Sustainable Development) is a 4-years project funded by the European Commission (EC) through 7th Framework Programme (FP7), aiming to address the subjects of ecologically unsustainable development and inadequate resource management in the Black Sea Catchment area [1]. The project develops a Spatial Data Infrastructure (SDI) targeting this region and using the European Web and Grid infrastructures. A large catalogue of environmental data sets (e.g. land use, hydrology, and climate) is gathered and used to perform distributed spatially-explicit simulations to build scenarios of key environmental changes.

The Black Sea Catchment Observation System (BSC-OS) portal developed by the enviroGRIDS project is the single point of access to resources such as geospatial data, hydrologic models, environmental scenarios, data processing tools and applications, visualization facilities, environmental reports, and training materials. The applications available through the portal provide functionalities such as data management, hydrological model calibration and scenario execution, satellite image processing, geospatial data visualization, environmental report generation and visualization, training material development, and lesson execution (Fig. 1). The Web portal is addressed to Earth Science specialists, decision makers and citizens and offers tools and applications specific for each category.

SWAT (Soil and Water Assessment Tool) is used as the primary hydrological model in enviroGRIDS for the Black Sea Catchment. It is a large-scale hydrological model that simulates the impact of various management practices on water, sediment, and agricultural chemical yields in ungauged watersheds [2]. It requires complex input data such as weather, hydrology, soil temperature, plant growth, nutrients, pesticides, land management, bacteria and pathogens information.

A few software platforms are primarily involved in the implementation of the portal architecture. The data management is accomplished mainly by URM (Uniform Resource Management) that provides the user with management and operations on geospatial data [3]. The user may enter data and metadata, visualize, modify, update, and remove spatial data from data repositories. The hydrological model oriented platforms and applications, gSWAT (Grid based SWAT) [4] and gSWATSim (gSWAT Services Implementation) [5], support the specialists on Grid based hydrological model configuration, scenario and model development, model calibration, and scenario execution. Using the GreenLand satellite data processing platform and application the specialists may process satellite data and images in order to search for relevant information (e.g. land cover, vegetation, water, soil composition, etc.) [6]. The geospatial data visualization and report generation is supported by the BASHYT toolset enabling the specialists to visualize various spatial data in different formats and views, and compose environmental reports for decision makers and citizens [7]. The virtual training center implemented by eGLE [8] allows the teacher to author training materials by retrieving through services the geospatial content from the URM and BASHYT repositories.

This paper mainly concerns with interoperability between different platforms implementing the tools and the applications provided by the portal as well as the interoperability between

Manuscript received October 03, 2011; revised February 11, 2012 and May 04, 2012; accepted July 13, 2012. Date of publication September 19, 2012; date of current version December 28, 2012. This work was supported by the FP7 enviroGRIDS Project funded by the European Commission, through Contract 226740.

D. Gorgan, V. Bacu, D. Mihon, T. Stefanut, and D. Rodila are with the Technical University of Cluj-Napoca, Computer Science, 400114, Cluj-Napoca, Romania.

P. Cau is with the Center for Advanced Studies, Research and Development, Computer Science, Pula, Italy.

K. Abbaspour is with EAWAG, Swiss Federal Institute for Aquatic Science and Technology, Department of System Analysis, Integrated Assessment and Modeling, Dübendorf, Switzerland.

G. Giuliani is with the University of Geneva, The Global Resource Information Database, Geneva, Switzerland.

N. Ray and A. Lehmann are with the University of Geneva, Environmental Sciences, Geneva, Switzerland.

Digital Object Identifier 10.1109/JSTARS.2012.2211580

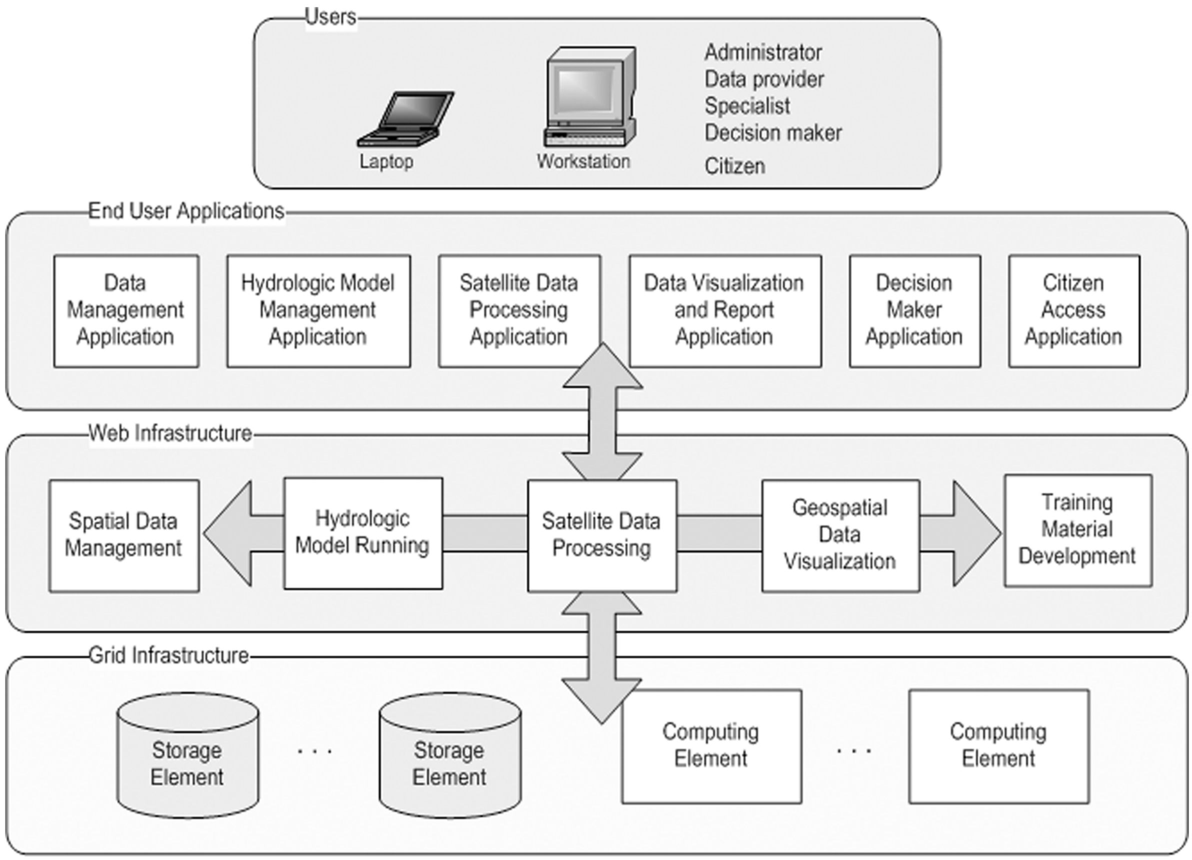

Fig. 1. Horizontal and vertical interoperability throughout the BSC-OS Portal architecture.

different architectural layers building up vertically the particular functionalities. The paper highlights the challenges and issues of the different ways of interoperability and analyses the solutions proposed and experimented in the BSC-OS portal by the enviroGRIDS project. The directions of analyzing the interoperability concern with data format and consistency, communication protocols, service based interface, compatibility on conceptual notions and methods (e.g. Web versus Grid infrastructure, Geospatial versus Grid services, security, service definition, data management, etc.), quality and performance of the interoperability, etc.

The paper is structured as follows. The next section explains the data and processing flow throughout the BSC-OS portal. Then each tool, platform and application is shortly described by highlighting the interoperability related issues and solutions in Section III. Section IV presents some use case scenarios. The next Section V analysis and details a few real use cases of interoperability between particular modules in the portal. Finally the paper sketches the conclusions and the future research work.

II. Data and Processing Flow Throughout BSC-OS Portal

The aim of the enviroGRIDS project addresses the subjects of ecologically unsustainable development and inadequate resource management in the Black Sea Catchment area. Various applications were developed in the frame of the project to accomplish these complex tasks. Starting from the individual and portal related functional and usability requirements each application has been developed independently by using specific software platforms and then integrated into the portal. Therefore the technical and technological interoperability between tools and applications is a key requirement of achieving their integration into the portal. Moreover, the quality and the performance of the interoperability are an imperative requirement and measure for the fluency of data transfer and processing control within the same working session.

Therefore the design and the implementation of the tools and the applications consider as an important objective the interoperability. The achievement of the interoperability is proved by an experimental global scenario connecting fluently the particular scenarios of each application by prototyping its functionality (i.e. data flow and processing control). This paper aims just to highlight and prove the interoperability solutions rather than to measure and evaluate its quality (i.e. usability, efficiency and performance).

There are three types of applications and tools available through the enviroGRIDS Web Portal (Fig. 2):

1) SWAT Scenarios Development Tools/Applications (gSWAT and gSWATSim, GreenLand): the user may develop various scenarios for natural phenomena (e.g. through a graph based language) and use cases, execute them over the Grid. The user may visualize the results and analyze statistical data.
2) Data Management Tools (URM): data administrators and data provider may access, upload, update, and organize spatial data.

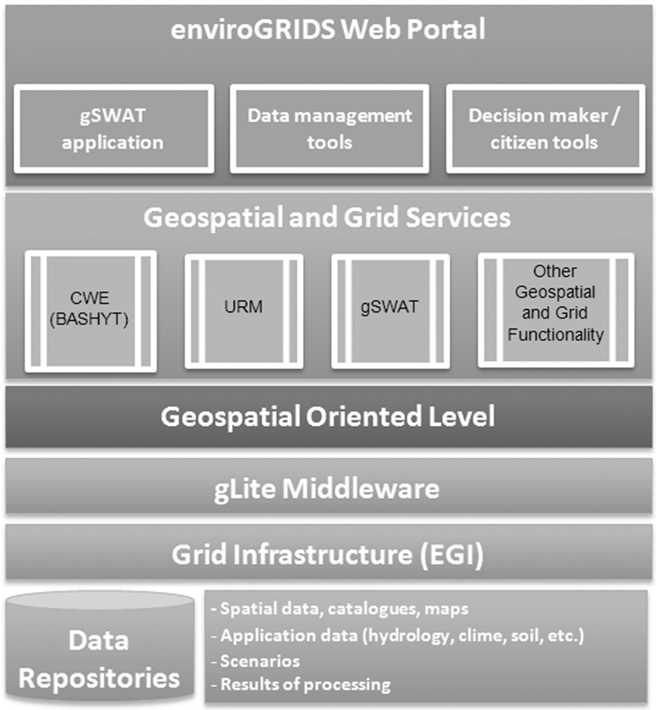

Fig. 2. enviroGRIDS portal levels.

3) Decision Maker/citizen tools:
   - Decision Maker Tools (BASHYT and gSWATSim): provide the user with the possibility to develop and execute various scenarios by different data series, and to analyze and make predictions on the phenomenon evolution. Graphical data visualization and mapping are available as well.
   - Citizen Tools (eGLE and BASHYT): provide the citizen, as an Internet visitor, the execution of a given set of scenarios by limited set of data, and graphical visualization of the results.

It is important the specialist user has an integrated control over all applications that guides him throughout the whole process. Data provider user enters, by a working session, spatial data into the URM repository. Later, another user may search, select and extract various data that can be used in the process of SWAT model development [2]. This data are accessible through standard OCG Web services. Some data retrieved in this way should be processed in order to use them. For instance by the GreenLand application the user can specify the data source (e.g. OGC Web Service provided by URM) and process those data (i.e. satellite images) by taking advantage of the computational resources provided by the Grid infrastructure.

Another particular sub-scenario could prototype, through the gSWAT application, the calibration and the running over the Grid of an already developed hydrological SWAT model. Based on a calibrated SWAT model the user may define and experiment new scenarios in order to make environmental studies and predictions. One possibility for developing scenarios is to use the tools provided by the BASHYT platform. By using the Web services provided through the gSWATSim platform the scenario oriented BASHYT application may export and execute remotely the new scenario, and finally import and visualize the results.

All the geospatial data and outputs of the previous applications may be used by the eGLE based application to build up training materials. The geospatial materials fetched by Web services from the URM and BASHYT repositories are integrated within training lessons and stored in the Virtual Training Center.

From the end-user's point of view the BSC-OS portal offers an integrated solution to the observation system of the Black Sea Catchment area. The user can run different applications to execute a specific workflow. All the applications are linked together by exchanging data or by services. In particular, the end-users perform different tasks through various tools and applications.

The main issues come from the fact that most of the applications and tools were developed individually. As most of the applications are exposing specific Web services the integration tried to linked them at this level. In other cases, the interoperability was solved by exchanging messages between them. The section related with the interoperability between platforms, tools and applications highlights the solutions.

The BSC-OS is not yet used in production but rather in the development and testing phases. Each individual tool has its own particular users. For instance, the URM is used mainly by the data providers, which feed the enviroGRIDS SDI with relevant data. On the other side, the SWAT specialists could use that data in the process of SWAT model development. The gSWAT and gSWATSim applications have not a wide target of potential users because mainly specialists in hydrology use them. The results are exposed via BASHYT and are important also for citizens not only for specialists and decision makers.

III. Platforms, Tools and Applications

This section briefly presents the applications included into the BSC-OS portal, highlighting the main functionalities and the architectural solutions. The interoperability solutions between this platforms, tools and applications are described in another section highlighting also the issues that were encountered. The Grid infrastructure that is used by some of the applications available in the portal is based on gLite [9]. Currently the enviroGRIDS VO is served by two Computing Elements (ce01.mosigrid.utcluj.ro and ce.seua-cluster.grid.am) and one Storage Element (se01.mosigrid.utcluj.ro). The computational resources provided by CE consist of 140 physical CPUs, with 1072 logical CPUs. The storage capacity is about 13 TB.

A. *gSWAT Platform and Application*

The gSWAT application [4], [5] allows the calibration of the SWAT models by a flexible and easy approach, both by the specialists and the students who are developing hydrological models. The gSWAT application uses remotely the Grid infrastructure in order to gain speed and scalability in the calibration process. The scalability concerns on increasing number of users and models running in parallel, as well as the large geographical area that could be modeled at a higher resolution.

The calibration process involves an iterative process until the user states the SWAT model is calibrated. Each iteration consists of running a set of simulations (e.g. 500 simulations), of the SWAT model but using different input parameters. Because

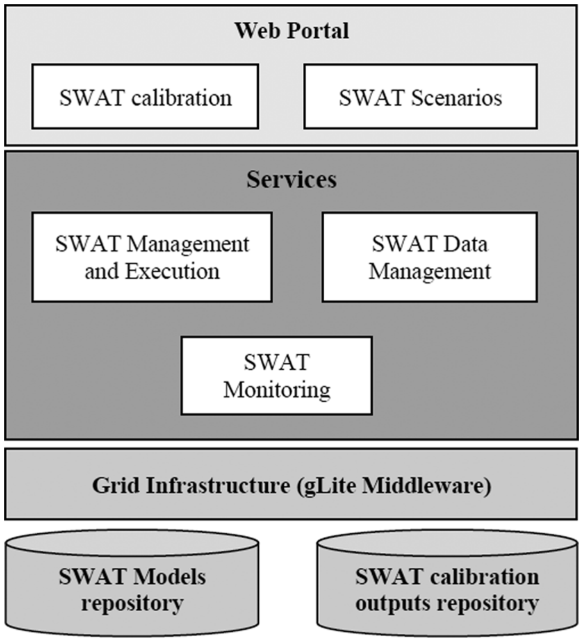

Fig. 3. gSWAT architecture.

each simulation is independent against others, all the simulations may run in parallel across the distributed Grid computing resources. Through a post processing phase the system determines the best simulation of the completed current iteration. Based on this relative optimum and an already defined objective function the specialist may search for another better calibrated status of the model by running a new iteration with other values of the input parameters [10].

The gSWAT platform consists of three tiers (Fig. 3):
1) the end-user gSWAT Web application;
2) the server side implementing the Web service related components;
3) the Grid resources.

The Web application is built in Adobe Flex 4 and hides to the user the complexity of the server and Grid based processing. It uses intuitive graphical user components that facilitate the interaction of the users with the system. The service related components expose the server functionality through Web services. There are three sets of services, such as SWAT management and execution, SWAT data management, and SWAT monitoring. The Grid infrastructure offers the computation and storing resources needed for running the parallel simulations, such as Worker Nodes (i.e. computation) and Storage Elements.

B. gSWATSim Platform

The gSWATSim platform is a collection of REST Web services supporting the remote calibration of the SWAT model and execution of the SWAT scenario over the Grid infrastructure. A SWAT scenario is defined starting from a calibrated SWAT hydrological model. The Web services expose the gSWAT functionality to other modules in the BSC-OS portal. The services support the functionality such as import new models and scenarios, execute them over the Grid, and export the results to the remote module (e.g. BASHYT based application). The execution of a scenario means the execution of a single simulation of the SWAT model. The monitoring component of the gSWATSim platform updates the status information so that the user can be informed on the progress of the execution.

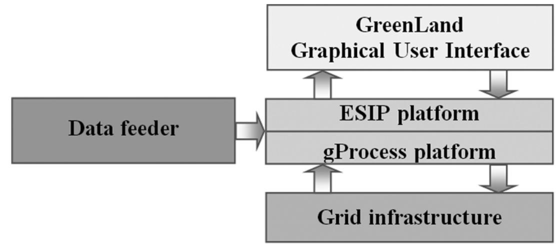

Fig. 4. GreenLand architecture.

C. BASHYT Toolset

BASHYT (The Basin Scale Hydrological Tool) offers a Web based interface for SWAT models and allows the user to produce reports in a flexible manner [7]. The SWAT models are digested by the system that generates a particular relational database. This database is then used to define reports consisting of objects such as tables, charts, forms, layers, etc. The reports concerning with environmental states can be visualized in various manners. The main core is developed in JAVA and is provided through the Tomcat servlet container. The use of Apache Velocity allows to define in real-time the look and feel of the application. The SPRITE technology supports the integration of the SWAT hydrological models into BASHYT and gSWAT. SPRITE is used to extract a minimum dataset from SWAT models, transform the model in the desired format, and upload the SWAT model onto the BASHYT and gSWAT servers.

D. GreenLand Application

The GreenLand application supports the satellite image processing, that is actually a searching of information across various combinations of the multispectral bands. The main processing is based on the Grid infrastructure, and the gProcess and ESIP (Environment Satellite Image Processing Platform) platforms [6], [11], [12]. Another important feature of the GreenLand application is the gridification of the GRASS (Geographic Resources Analysis Support System) operators, which allows the user to develop different workflows and to execute them over the Grid infrastructure.

The components of the GreenLand architecture are presented in Fig. 4:
- Data Feeder—allows the retrieval of data from different data repositories;
- Graphical User Interface—exposes the functionality to the user;
- ESIP and gProcess platforms—provide the basic operators and workflows, and allow the execution of the complex workflows over the Grid infrastructure;
- Grid infrastructure—provides the computational and storing resources.

E. eGLE and Virtual Training Center

eGLE (eLearning Environment) is built on a distributed architecture that supports the management and processing of large amounts of data (see Fig. 5) [8]. Most of the complex operations needed to interact with Grid infrastructure, distributed

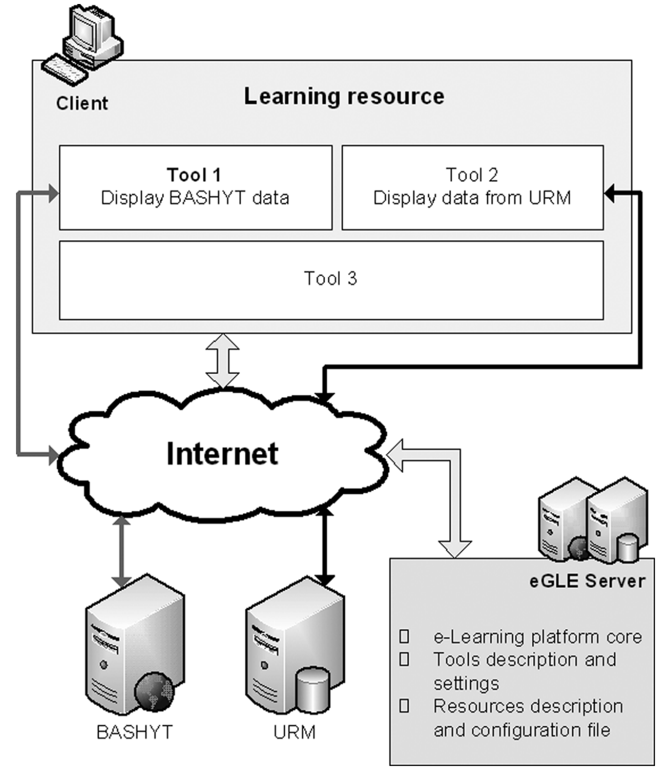

Fig. 5. eGLE eLearning environment.

databases, Web and multimedia services, etc. are presented to the teachers through dedicated and very simple user interfaces. These allow them to access all the functionalities without requesting advanced technical knowledge. eGLE provides mechanisms for knowledge presentation and assessment based on Grid processing capabilities, both for teachers and students. Also, the lesson creation process has been designed to maximize the efficiency of the teacher by allowing him to reuse, as much as possible, previously developed learning patterns and resources, created either by him or by other teachers.

F. URM Platform

The implementation in URM of enviroGRIDS SDI is based on analysis of INSPIRE and GEOSS principles [3], [13]. The system design is based on the principles of the Service Oriented Architecture (SOA) and is INSPIRE compliant. The INSPIRE requirements give to the overall system architecture a loosely coupled integration based on OGC standard usage, which allows to use any OGC compliant software component and easy replace of it by another if it is required. In order to achieve interoperability, the main software interface among each particular component has to be based on ISO standards and OGC specifications, following the INSPIRE directive and GEOSS architecture. The main architectural components of the enviroGRIDS Geoportal are (Fig. 6):

- Discovery services using Micka technology [14]—It implements search and discovery services by exposing as well the catalogue services. MicKa is a complex system for metadata management used for building Spatial Data Infrastructure (SDI) and geoportal solutions. It allows the editing and management of metadata for spatial information, Web services and other sources (documents, Web sites, etc.). MicKA is compatible with obligatory standards for European SDI building (INSPIRE). HSLayers is a JavaScript Mapping framework used for building rich mapping portals.
- Metadata management using Micka technology—Provides to users and other system components the access to enviroGRIDS metadata. It supports internal Metadata Management.
- View services based on HSlayers [15]—Visualise the results supplied by the application service components. The portrayal services perform the rendering of generic data (e.g. catalogue entry, map image) into an output format that will be delivered to the user through the horizontal service and then through the application services.
- Uploading and downloading services based on Geohosting [16]—This component implements view and download services by exposing the map and feature services. View services allow display, navigate, zoom in and out, pan or overlay viewable spatial data sets and display legend information and any relevant content of metadata. Download services allow extracting copies of spatial data sets, or parts of such sets, to be downloaded and, where practicable, accessed directly.
- Content management services using Simple CMS—The Content Management System (CMS) is a collection of procedures used to manage the workflow in a collaborative environment.

IV. Examples of Scenarios

This section shortly presents some use-cases or scenarios that are available and which are involving different applications and tools from the enviroGRIDS Web Portal.

A. Integrated Scenario Regarding SWAT Models

This section presents an example of integrated scenario (Fig. 7), involving different applications available through the Web portal, related with SWAT hydrological models (from the creation of the un-calibrated SWAT models to reports and lessons). Hydrological specialists develop SWAT models in specific applications such as ArcSWAT or AvSWAT. The ArcSWAT application is an ArcGIS-ArcView extension providing a graphical user input interface for the SWAT. The next step is to calibrate the SWAT models before defining various scenarios. The calibration process requires huge computation capabilities (we need to run hundreds and even thousands of simulations) that can be offered by the Grid infrastructures. For the calibration process the user is using the gSWAT application. After this process is completed and the model is calibrated we need to upload the calibrated model to BASHYT. This new calibrated model is identified in BASHYT based on the fingerprint of the watershed. BASHYT offers tools to define new scenarios based on calibrated SWAT models. New scenarios, which represent new SWAT models, are fetched to the gSWATSim component that can execute scenarios. gSWATSim application runs the scenario and then uploads the results to BASHYT. The user can define different reports in BASHYT, a report aggregating different information presented as maps, charts, tables, images, text, etc. eGLE is able to include different information exposed

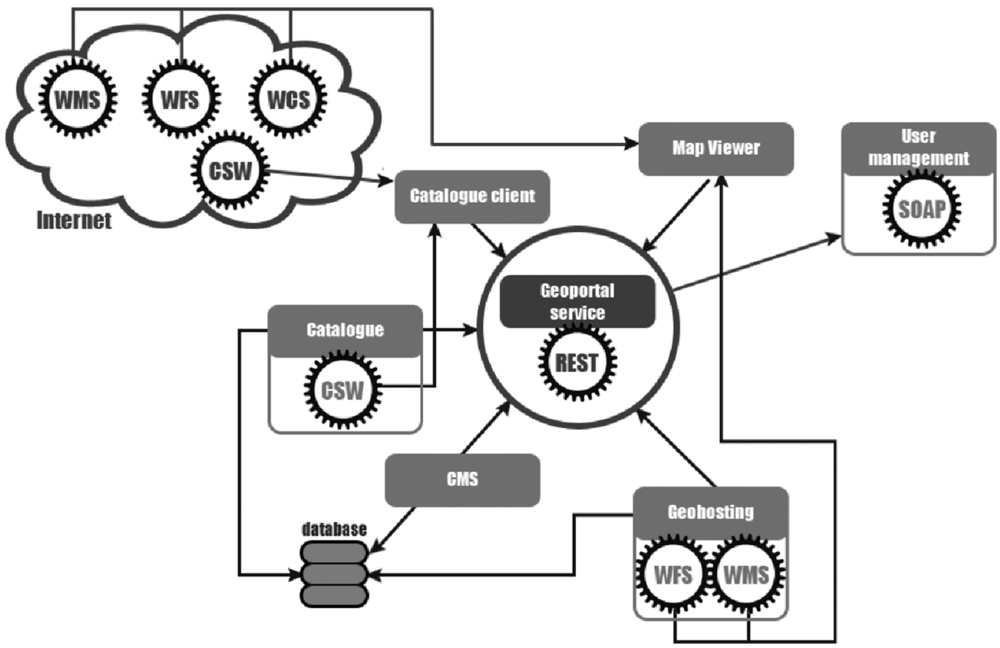

Fig. 6. enviroGRIDS Geoportal architectural components (figure taken from [1] and [13]).

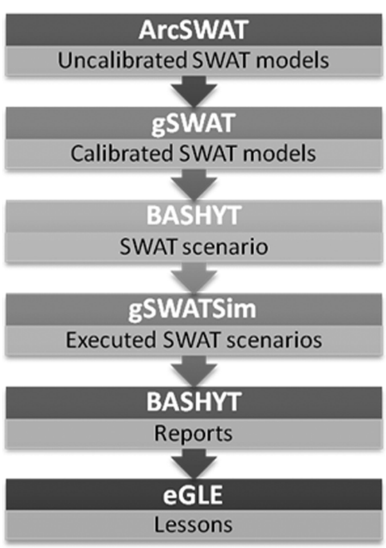

Fig. 7. SWAT related integrated scenario.

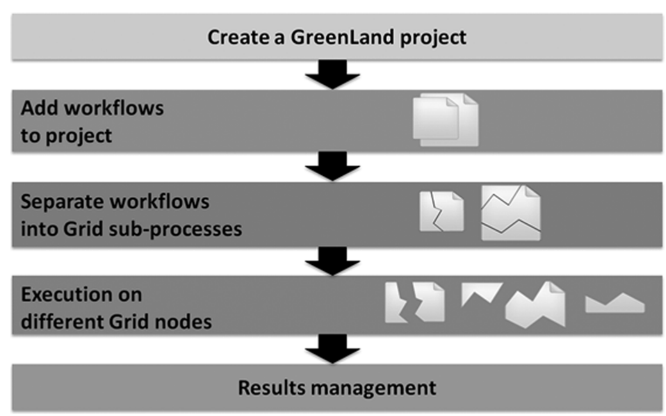

Fig. 8. Project life cycle in GreenLand application.

through BASHYT in lessons that could be used in the process of teaching for instance the SWAT calibration procedure.

The main outcome of this integrated scenario is that the users are guided to all these steps within the same main umbrella (being the BSC-OS Web Portal) in an integrated manner. All the communication and interaction between all the applications is done transparently from the user's point of view.

B. Spatial Data Management Scenario

The data providers may enter, visualize and update spatial data according to the INSPIRE and GEOSS specifications by using the URM platform. Spatial data consists of geospatial data, satellite data and images, and hydrological data files. Data is available to any other module of the BSC-OS portal by Web geospatial services according to the OGC WS specifications.

C. SWAT Model Development

SWAT is a large-scale hydrologic model designed to simulate the impact of various management practices on water, sediment, and agricultural chemical yields in ungauged watersheds [2]. Main model components consist of weather, hydrology, soil temperature, plant growth, nutrients, pesticides, land management, bacteria and pathogens. The program ArcSWAT is used to develop a SWAT model. The ArcSWAT project is developed within the ArcGIS 9.3 framework. The basic SWAT data requirements include elevation, soil, landuse, and climate data. River discharge data and water quality information such as sediment concentration, and levels of nitrate, and phosphate, etc. are used to calibrate the model. Depending on the objective of the model information such as dams and reservoirs, agricultural management practices (fertilizer application, planting, harvesting, irrigation, etc.) may also be required.

The SWAT model is developed based on spatial data supplied by the enviroGRIDS data repositories through OGC standard services. The calibration and the execution of the SWAT model are supported by the gSWAT application as it is described in the

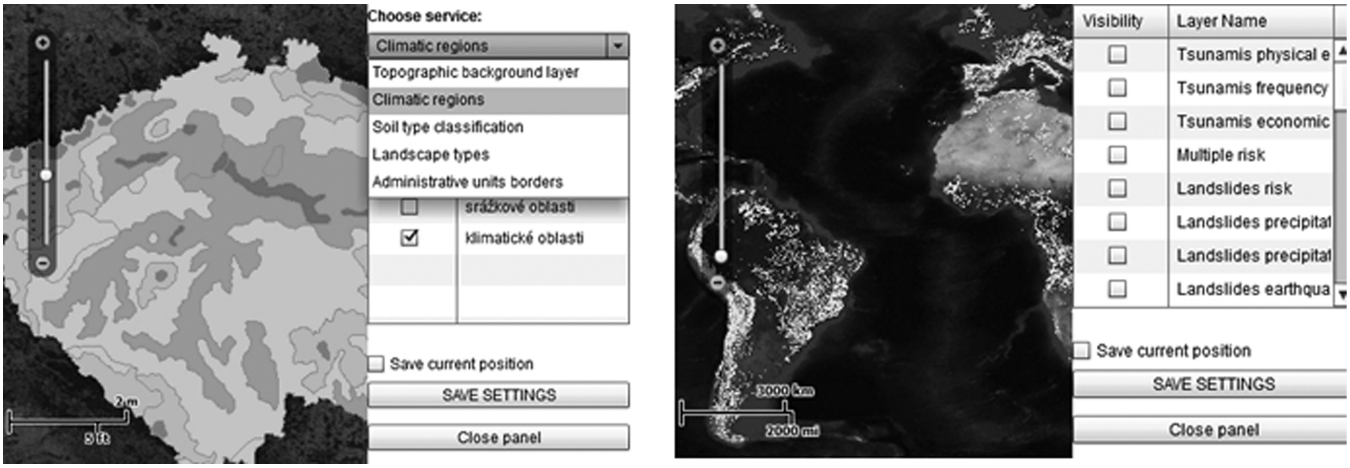

Fig. 9. Tools for OGC Web services integration in teaching materials.

previous sections. The results are visualized by BASHYT and included by eGLE into the training materials.

D. Satellite Data Processing Scenario

The satellite data is processed over the Grid. The users generate and execute workflows, based on satellite images. In this case each workflow is represented as a DAG (Direct Acyclic Graph). The nodes of the graph could take the form of a simple operator or a sub-workflow. Multiple nested levels are allowed within the GreenLand application. The processed data can be visualized or integrated in different reports. The project lifecycle is presented in Fig. 8. Both the initial and processed satellite data are stored within the Grid in Storage Elements, and are accessed through OGC WS, as any data from Web infrastructure. The input data could also be specified as an URM resource.

E. Report Generation and Visualization

The reports can be composed of data provided by data suppliers, hydrological model execution results, processed satellite data. The data is stored in Web repositories and in the Grid's Storage Element. The reports are stored outside the Grid in data repositories in the Web infrastructure. The reports, from the Web's data repository, are visualized by the decision makers and citizens through applications (BASHYT or eGLE).

F. Integration of OGC Web Services and BASHYT API in EnviroGRIDS Training Center

The BSC-OS portal provides virtual training capabilities with support for earth science specialists that want to create or to consult learning materials. The users will be recognized by the system either as teachers—trainers that will create training materials, or students—users that will have only the ability to consult materials available into the Training Center.

The information available for inclusion into training materials (e.g. data in text/table formats, images, reports, maps, etc.) is stored in different databases throughout the entire BSC-OS portal: BASHYT, URM, gSWAT, GreenView, etc. At the same time, the methods required for connecting to these data sources and for retrieving information are quite different. For example, the information from BASHYT, gSWAT and GreenView applications is exposed through specialized APIs, while URM or UNEP data repositories can be accessed only through OGC Web services.

For example, the integration of OGC Web services is implemented through two specialized tools (Fig. 9). According to the different scenarios that could involve this type of data types, one of the tools allows multiple services management and connection while the other one has been tailored to one specific service that cannot be changed either by trainer or trainee.

V. Interoperability Throughout Portal

Interoperability means the ability of different systems to work together in order to achieve a specific goal. In general, the interoperability is accomplished by exchanging information between two or more systems. The syntactic interoperability is related with the ability of the systems to communicate and exchange data. For the semantic interoperability the systems should refer to a common information exchange reference model.

A. Vertical and Horizontal Interoperability

The interoperability between different platforms, tools, and applications composing the enviroGRIDS portal concerns with functionality provided by different modules to each other, data format and consistency transmitted between modules, communication channels and protocols, compatibility on conceptual notions and methods, and quality and performance of the interoperability. Therefore there are various directions of studying the interoperability, as well as different solutions and evaluation criteria of the implementation.

One classification of the interoperability concerns on direction toward which the functional modules and layers collaborate: horizontally and vertically (see Fig. 1). The horizontal interoperability deals with collaboration between different platforms and modules within the same functional and technological layer. The vertical interoperability concerns with collaboration between different architectural layers building up vertically the particular functionalities. Therefore, the horizontal interoperability is accomplished between platforms and modules within the same architectural layer, unlike the vertical interoperability that deals with components or modules of different layers within the same platform or application. The basic solution of interoperability is accomplished by services, messages, and data.

All the applications that compose the enviroGRIDS portal were developed independently, making in this way the process of linking them together not so straightforward. The next sections describe how we accomplished the interoperability between these applications and tools.

B. Service Based Interoperability

The URM platform exposes data and functionality through standard OGC Web services. The GreenLand application and the eGLE toolset may use spatial data provided by URM through the geospatial services. The interoperability between gSWAT and BASHY platforms is supported mainly both for functionality and data, by the Web services provided by the gSWATSim platform.

The OGC Web services use standards to support sharing, interaction, and visualization of geospatial data. The OGC services, available in the enviroGRIDS portal, such as WMS (Web Map Service) retrieves maps in a standardized manner, WFS (Web Feature Service) offers access to raw geospatial data, and WCS (Web Coverage Service) offers access to raster datasets.

In other cases, the interoperability between applications concerning huge data sets is accomplished at the data repository level. That means that the applications exchange data to perform the desired task by storing and fetching them through the repository, such as Storage Element available in the gLite based Grid infrastructure.

C. Geospatial and Grid Infrastructures Interoperability

The interoperability between the Geospatial and Grid infrastructure is a main challenge inside the enviroGRIDS projects as it promises to solve the integration of Geospatial services and data with the Grid platform, obtaining better performances and solving a lot of issues such as the management of large volumes of geospatial data in distributed environments. The interoperability is supported by the gridification of the OGC services such as WMS (Web Map Service), WFS (Web Feature Service), WCS (Web Coverage Service), WPS (Web Processing Service), and CSW (Catalog Service for Web).

The gridification of the OGC Web services is a complex task concerning mainly at least with the service related conceptual differences between the Geospatial and Grid platforms such as invocation interface (HTTP vs. SOAP messages), description (capability document vs. WSDL), discovery (CWS service vs. UDDI), communication (messages, HTTP, XML vs. SOAP, GridFTP), security (Authentication/Authorization not included in OGC service vs. Grid Certificate), and state information (stateless vs. stateful).

The approach chosen to solve the interoperability between Geospatial and Grid infrastructures is a mediation approach and it is achieved through the introduction of a new component called Mediator, able to mediate the major differences between the two infrastructure components and to create an interoperable communication while targeting the data parallelism. Using the Mediator component, the Geospatial services can be adapted to follow different execution flows, either over the Web or over the Grid, considering both the input parameters and the execution platform characteristics.

Considering the complexity of the problem, the interoperability between Geospatial and Grid infrastructure has been analyzed and discussed in conference proceedings [17], [18] as well as a journal paper [19], where the authors present in details the interoperability issues, the approaches chosen to solve some of them, the architecture and functionalities of the Mediator component as well as the obtained results.

D. Interoperability Between gSWAT and BASHYT

The gSWAT application supports the calibration of the SWAT hydrological models, meanwhile BASHYT is able to support the development of scenarios and the visualization of the outputs resulted from the SWAT scenario execution. The interoperability between gSWAT and BASHYT is supported differently for each of two use cases. The first use case uses just the interoperability at the data level unlike the second case that uses the service based interoperability as well as the data based one.

In the first case the gSWAT and BASHYT applications are used individually. The two applications collaborate just by data repository, instead of exposing functionality by Web services. The user calibrates the hydrological SWAT model by gSWAT and stores the results of processing on the Storage Element within the Grid infrastructure. Then, anytime later the user may visualize the results just by fetching in BASHYT the spatial data from the Storage Element.

The second case deals with both types of interoperability—service and data based. The BASHYT application invokes functionality provided by gSWAT through the gSWATSim services. The BASHYT application may export the SWAT model, run and read the processing status information, and finally fetch the results. The huge geospatial data are transferred between the applications through the data repository (i.e. Storage Element within the Grid).

For instance, the interoperability workflow between BASHYT and gSWAT consists of the following steps:

1) Develop an initial SWAT hydrological model by ArcSWAT (i.e. M1 model);
2) Import the SWAT model into BASHYT and gSWAT;
3) Calibrate the model remotely in gSWAT by using the gSWATSim services and the Grid infrastructure (i.e. CM1 calibrated model);
4) Import the calibrated model to BASHYT;
5) Develop and edit the new SM1 scenario based on CM1;
6) Transfer scenario SM1 to gSWAT by gSWATSim services;
7) Execute SM1 scenario using gSWATSim monitoring the processing, and store the results on SE;
8) Download and visualize in BASHYT the results of scenario execution.

We need a way to link the SWAT models from gSWAT with the ones from BASHYT. This is solved at step 2 from the interoperability workflow where the SWAT model is uploaded to both applications. SPRITE creates a unique fingerprint of the SWAT model that will be stored together with the model. Based on this unique fingerprint we can identify the same model on both applications.

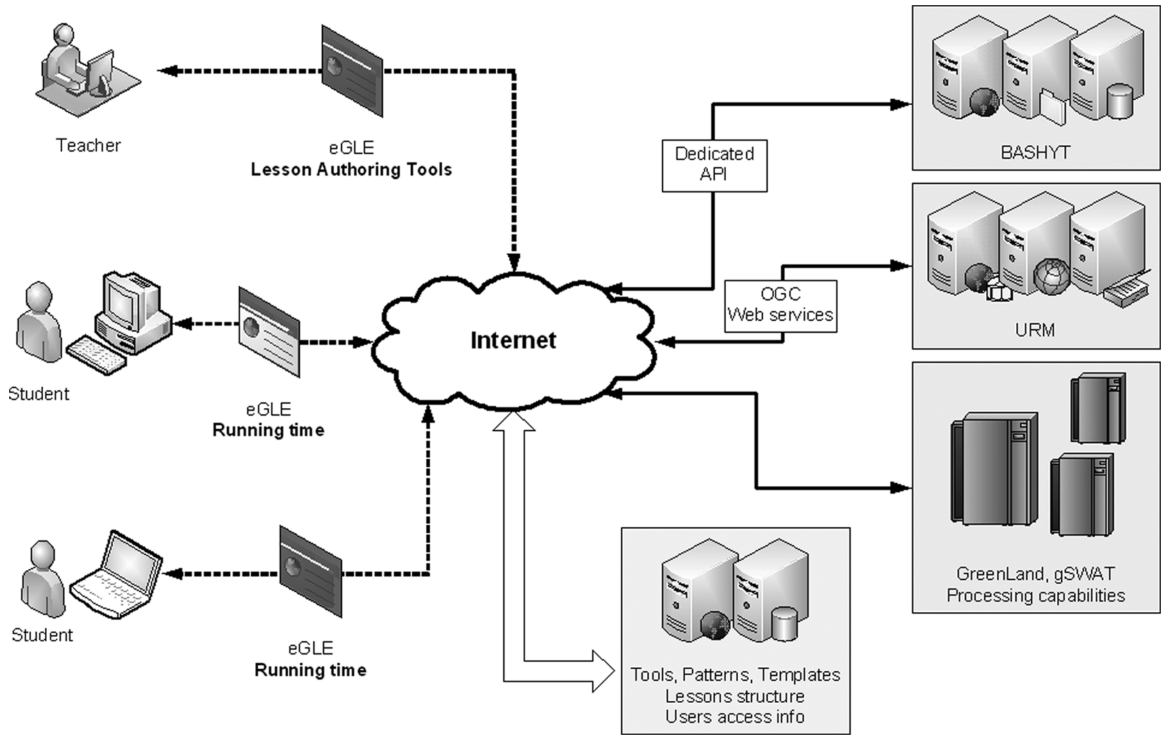

Fig. 10. Service based interoperability between eGLE and other platforms within the portal.

E. Interoperability With eGLE

Information that will be included into the training materials (e.g. data in text format, images, etc.) through the eGLE platform is stored in different repositories of BASHYT, URM, gSWAT, and GreenLand. At the same time, the methods of connecting to these data sources, in order to retrieve this information are different for each case: dedicated API—for BASHYT, gSWAT and GreenLand—and OGC Web services—for URM (Fig. 10).

As the data acquiring process can thus become very complex, it is necessary the interface provides functionalities that will reduce the complexity of operations for the user. The teacher should be able to browse the available resources without any knowledge about the location, storage or retrieval mechanisms.

The aim of the interoperability between eGLE and BASHYT is the ability to retrieve data from BASHYT (e.g. watershed representations, reports, graphics, etc.) through the dedicated API that allows any third party with proper credentials to extract information about specific scenarios that have been executed through Bashyt. Using these services one can visualize reports, information in tabular form, charts, maps and other available information. Another objective is to use eGLE capabilities to allow the teacher to include as well information retrieved from the URM through OGC Web services, in the same training material, through specific tools approach [8].

It is not a trivial action to query this specialized API and retrieve the data from it. Technical knowledge is required for a proper information exchange, and code programming is also necessary for information retrieval and later management. As a result, it gets difficult for specialists in hydrology or Earth Observation to include information from BASHYT in their own training materials.

The dedicated tool developed and included intro eGLE application allows both trainer and trainee (according with the interaction restrictions established by the teacher) to extract a certain type of information from Bashyt, through a simple to use interface.

F. Interoperability With GreenLand

The GreenLand application offers a Grid based solution to process satellite images of different types and formats. The data can be fetched from different sources (e.g. URM), and one solution is to gather data through the OGC Web services. The workflow steps are the following:

1) Introduce the URL of the service in the text field for data source;
2) Parse the XML document returned by the GetCapabilities request, and select the layer to be processed;
3) Retrieve the data with a suitable request and then process them through the GreenLand application.

In this case, the interoperability between GreenLand and URM could be bidirectional. In the first step the URM is the provider of data for GreenLand, and finally the processed data may be stored back into URM. The OGC Web services allow retrieving data in different data formats. In this case, the GreenLand feeder component is responsible for choosing the most suitable data format. Actually data are transferred directly from URM to GreenLand, and back, through HTTP requests.

VI. Conclusion

The paper analyses and exemplifies the interoperability available between different platforms that implement the tools and the applications in the BSC-OS portal, as well as the interoperability between different architectural layers building up vertically the particular functionality of the applications. The paper highlights the challenges and issues of the different ways of interoperability and analyses the solutions proposed and experimented in the BSC-OS portal by the enviroGRIDS project. The directions of analyzing the interoperability concern with data format and consistency, communication protocols, service based interface, and compatibility on conceptual notions and methods.

From the enviroGRIDS end-users point of view the portal offers an integrated solution of the various tools and applications developed in the frame of the enviroGRIDS project. By integrating all the applications in the Web portal the user can access them in a flexible manner and without the need of installing anything on their own computers. In this way, it offers a platform independent solution. For SWAT specialists it offers a flexible and interactive solution for the calibration process, minimizing the execution time by taking advantage of the Grid infrastructure. The Earth Science specialists can define the processing of satellite images as workflows and execute them on the Grid infrastructure.

The future works will concern with definition, experimentation, and metrics based evaluation of the quality and performance of the interoperability. The high fluency of data transfers, functionality exposing and availability, and processing control throughout the BSC-OS portal are among the main objectives of the enviroGRIDS project.

References

[1] enviroGRIDS Project. [Online]. Available: http://www.envirogrids.net/, 2012

[2] SWAT (Soil and Water Assessment Tool) Model. [Online]. Available: http://swatmodel.tamu.edu/ 2012

[3] K. Charvat, S. Kafka, M. Splichal, M. Alberts, A. Martinsone, P. Horak, M. Vlk, and P. Bruins, "URM concept for sharing information inside of communities," in *WCCA 2008*, Tokyo, Japan, 2008.

[4] V. Bacu, D. Mihon, D. Rodila, T. Stefanut, and D. Gorgan, "Grid based architectural components for SWAT model calibration," in *Proc. 2011 Int. Conf. High Performance Computing and Simulation (HPCS)*, Jul. 2011, pp. 193–199.

[5] V. Bacu, D. Mihon, T. Stefanut, D. Rodila, D. Gorgan, P. Cau, and S. Manca, "Grid based services and tools for hydrological model processing and visualization," in *Proc. 13th Int. Symp. Symbolic and Numeric Algorithms for Scientific Computing*, 2011, pp. 291–298.

[6] D. Gorgan, V. Bacu, D. Rodila, F. Pop, and D. Petcu, "Experiments on ESIP—Environment oriented satellite data processing platform," *Earth Science Informatics*, vol. 3, no. 4, Dec. 2010.

[7] S. Manca, C. Soru, P. Cau, G. Meloni, and M. Fiori, "A multi model and multiscale, GIS oriented Web framework based on the SWAT model to face issues of water and soil resource vulnerability," presented at the Fifth Int. SWAT Conf., Aug. 3–7, 2009 [Online]. Available: http://www.brc.tamus.edu/swat/4thswatconf/docs/rooma/session5/Cau-Bashyt.pdf

[8] T. Stefanut, V. Bacu, and D. Gorgan, "eLearning lesson development and execution based on gProcess workflow description platform and eGLE E-learning platform," in *3rd Int. Workshop on High Performance Grid Middleware, HiPerGRID—Proc. CSCS-17 Conf.*, Bucharest, Romania, May 28, 2009, vol. 2, pp. 431–436.

[9] gLite Middleware, 2012 [Online]. Available: http://glite.cern.ch/

[10] K. C. Abbaspour, J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, and R. Srinivasan, "Modelling hydrology and water quality in the pre alpine/alpine Thur watershed using SWAT," *J. Hydrol.*, vol. 333, pp. 413–430, 2007.

[11] V. Bacu, D. Rodila, D. Mihon, T. Stefanut, and D. Gorgan, "Error prevention and recovery mechanisms in the ESIP platform," in *Proc. IEEE 6th Int. Conf. Intelligent Computer Communication and Processing*, 2010, pp. 411–417.

[12] D. Rodila, V. Bacu, and D. Gorgan, "Integration of satellite image operators as workflows in the gProcess application," in *Proc. 5th IEEE Int. Conf. Intelligent Computer Communication and Processing, ICCP2009*, Cluj-Napoca, Romania, Aug. 27–29, 2009, pp. 355–358, 2009.

[13] K. Charvat, S. Kafka, M. Splichal, O. Cerba, and J. Hielkema, "Uniform resource management," in *IST Africa 2008*, Winshoc, Namibia, 2008.

[14] MicKa Metadata Editor, 2012 [Online]. Available: http://www.envirogrids.cz/

[15] HSLayers, 2012 [Online]. Available: http://bnhelp.cz/hslayers/

[16] Geohosting. [Online]. Available: http://www.geohosting.cz/eng/index, 2012

[17] D. Rodila and D. Gorgan, "Integration of spatial data infrastructures with grid environment," in *Proc. 12th Int. Symp. Symbolic and Numeric Algorithms for Scientific Computing*, 2010, pp. 269–277.

[18] D. Rodila and D. Gorgan, "A mediation approach in geospatial Web services gridification," in *Proc. IEEE Int. Conf. Intelligent Computer Communication and Processing, ICCP2011*, 2011, pp. 541–548.

[19] D. Rodila and D. Gorgan, "Geospatial and grid interoperability through OGC services gridification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS), Special Issue on Interoperability Architectures and Arrangements for Multi-Disciplinary Earth Observation Systems and Applications*, 2012, this issue.

Dorian Gorgan is Ph.D. supervisor in Computer Science and Information Technologies at the Technical University of Cluj-Napoca, Romania. He is the chair of the CGIS (Computer Graphics and Interactive Systems) Laboratory. His fields of interest concern with graphical modeling and processing based on graphics clusters, grid computing, interactive and distributed applications, Earth Science applications and tools development, geospatial and Grid infrastructures interoperability. He was the initiator and the director of the MedioGrid project and national research and academic Grid infrastructure. He is involved as well in other Grid related international research projects such as EnviroGRIDS, GiSHEO, and SEE-GRID-SCI.

Victor Bacu is Ph.D. student in computer science and a Junior Lecturer at the Technical University of Cluj-Napoca, Romania. He is the member of the CGIS (Computer Graphics and Interactive Systems) Laboratory. His research contributes to national and international projects including MedioGrid, SEE-GRID-SCI, GiSHEO and enviroGRIDS. His research interests focus on Grid application development, cluster based visualization, tool packages for flexible classification in the satellite images.

Danut Mihon is a Ph.D. student at the Technical University of Cluj-Napoca, Romania, with extensive experience in developing interactive applications for processing and visualization of spatial data over the Grid infrastructure. He is part of the CGIS (Computer Graphics and Interactive Systems) Laboratory, involved in three projects funded by the European Union: SEE-GRID-SCI, GiSHEO and enviroGRIDS. His work has resulted in more than 10 scientific publications, including one journal publication.

Teodor Stefanut received the M.S. degree in computer science in 2009 and continues his Ph.D. studies in the same domain at the Technical University of Cluj-Napoca, Romania. He is currently a member of the Computer Graphics and Interactive Systems Laboratory (CGIS) and a Junior Lecturer at the same institution, in the Computer Science department. His research interests include new user interactions techniques, development of Grid applications for non-technical specialists, eLearning platforms based on distributed architectures, graphical annotation as user interaction technique in the context of a lesson.

Denisa Rodila is a Ph.D. student in computer science and information technologies at the Technical University of Cluj-Napoca, Romania, and she received her Diploma and Master degree at the same university. She is a member of the CGIS (Computer Graphics and Interactive Systems) Laboratory and her research contribution is included in national and international projects such as GiSHEO and enviroGRIDS. She has edited several papers in the areas of distributed computing and her research interests include distributed computing and intelligent systems, with focus on Grid, cluster and service oriented technologies.

Pierluigi Cau works at CRS4 as a researcher in the Environmental Sciences program of the Energy and Environment sector since 2000. His research topics are focused on computational hydrology, GIS and WEB ICT technologies for the management of water and soil resources. He organized international workshops/conferences and taught in various University advanced courses on hydrological (both superficial and groundwater) modeling. He has tutors several stagers and early stageresearchers.

Karim Abbaspour is a soil scientist and hydrologist. He has studied at UBC in Vancouver, Canada, and ETHZ in Switzerland. He has been working at the Eawag, Swiss Federal Institute of Aquatic Science and Technology since 2000. His main specialization is in application of flow and transport models at small and large scales. He is also working on methods of model calibration, validation, and uncertainty analysis.

Gregory Giuliani After obtaining a degree in Earth Sciences, he went on to complete a master in Environmental Sciences, specializing in remote sensing and GIS. He previously worked as a GIS Consultant for the World Health Organization, as a University tutor in remote sensing and GIS and as a GIS Developper in a local Swiss GIS company. He works at UNEP/GRID-Europe since 2001 and is the focal point for Spatial Data Infrastructure (SDI).

Nicolas Ray is the enviroGRIDS project manager. He holds a Master degree in Environmental Sciences, a Ph.D. in biology from University of Geneva in the field of Human Population Genetics, and a postgraduate certificate in Computer Sciences. His research activities focused on the modeling of animal movement and habitat, with the development of several spatial and statistical analysis tools to integrate various data types (genetic, environmental, demographic). Through his research and various consultancy works, Nicolas acquired solid competencies in GIS, spatial analysis, software development, and grid computing integration. After six years of postdoctoral research in Australia, UK, and University of Bern, he recently joined the Institute of Environmental Science at University of Geneva, and UNEP/GRID-Europe to dedicate his time to the enviroGRIDS project.

Anthony Lehmann is the EnviroGRIDS project initiator and coordinator. He holds a Master degree and a Ph.D. in aquatic biology from the University of Geneva, and a Postgraduate Master in Statistics from the University of Neuchâtel. He specialized during his career in combining GIS analyses with statistical models. At the University of Geneva he is in charge of the enviroSPACE laboratory exploring Spatial Predictions and Analyses in Complex Environments.