Abstract: The enviroGRIDS project has developed and provides through the BSC-OS portal a set of tools, applications and platforms concerning with the processing of huge spatial data for the Black Sea catchment region. The paper highlights the main issues of interoperability between Geospatial and Grid infrastructures, and between different platforms supporting the Earth Science oriented tools and applications. The BSC-OS portal provides end user applications for spatial data management, hydrological model calibration, environmental scenario development and execution, workflow based satellite image processing, data reporting and scenarios visualization, and development of Earth Science oriented training materials.

Keywords: spatial data; grid computing; hydrological model; satellite images; Earth Science.

1. INTRODUCTION

Ecologically unsustainable development and inadequate resource management, in the context of climate, land cover and population changes, in the Black Sea catchment region, are the main concern of the enviroGRIDS (Black Sea Catchment Observation and Assessment System supporting Sustainable Development) FP7 project [1]. The quantity and quality of waters are extremely important as well as understanding the evolution of the complex environmental systems over the coming decades. The enviroGRIDS project aims to develop, calibrate, and make available the hydrological model of the Black Sea catchment region by four main achievements:

- Collection of large transnational data sets;
- Adequate management and sharing processes of the environmental data by a dedicated Spatial Data Infrastructure (SDI);
- Distributed computing in order to allow running a high-resolution hydrological model;
- Providing a set of tools and applications to specialists and citizens in order to access data processing and visualization, and analyze environmental scenarios.

One main challenge of the project is to experiment and implement the interoperability between different technologies, platforms, and applications. One such a case is the interoperability between the Geospatial and Grid infrastructures, in order to extend the features provided by the both technologies. The Geospatial technologies offer very specialized functionalities for Earth Science oriented applications meanwhile the Grid technology is able to support high performance computation by scalability, and distributed and parallel processing.

The resources of the enviroGRIDS system are accessible to the large community of users through the BSC-OS (Black Sea Catchment Observation System) portal. By single sign-on authentication technique the portal provides Web applications for data management, hydrological models calibration and execution, satellite image processing, report generation and visualization, and virtual training center.

This presentation focuses on the BSC-OS portal architecture, and the main challenges and issues regarding the development of environmental tools and applications regarding the Black Sea catchment.

The paper is structured as follows. Section 2 presents the works and achievements related with the enviroGRIDS project. Section 3 sketches the portal architecture and the set of tool and application
categories. Each of the next six sections describes a tool and application category such as data management, SWAT model calibration and scenario execution, satellite image processing, spatial data visualization and reporting, two demonstrator applications oriented to citizens and decision makers, and virtual training center. The last section concludes on the portal development and future work.

2. RELATED WORKS

The enviroGRIDS project develops the SWAT model as a high-resolution (i.e., sub-catchment spatial and daily temporal resolution) water balance model to the entire Black Sea catchment region. The model is calibrated and validated by using river discharge data, river water quality data, and crop yield data as in [2]. The Black Sea watershed related hydrological model is very complex due to the highly interconnected and continuously evolving interactions at many spatial and temporal scales, and requires to gather and integrate different sets of environmental data (e.g. physical, chemical, biological) [3]. Other European projects aim environmental related subjects [4]. IS-ENES project develops the European Network for Earth System Modeling (ENES), which calls together the European climate/Earth system modeling community in order to work on understanding and prediction of future climate change. The ENSEMBLES project was a joined effort to develop an ensemble prediction system for climate change based on the principal state-of-the-art, high resolution, global and regional Earth System models developed in Europe. The METAFOR project addresses the fragmentation and gaps in availability of metadata for climate data that are currently found in existing repositories. The goal of the DRIHMS project is to systematically build a bridge between the HMR (Hydro-Meteorological Research) and ICT (Information and Computing Technology) communities, and to identify requirements of HMR users and match them to capabilities of the newly developed ICT infrastructure. The GENESI-DEC project aims to provide guaranteed, reliable, easy, effective access to a variety of data, facilities, tools and services to an ever increasing number of Digital Earth users from all disciplines.

The projects EGEE, SEEGRID-SCI, and C3Grid, provide solutions for sharing complex spatial and environmental data sets, and Grid based processing tools and applications. The aim of the C3Grid project for instance, is to create a grid-based working environment for earth system research.

Manny other European projects such as SAW-GEO, CYCLOPS, GDI-Grid, GEO-Grid, DEEGREE, DORII, and GENESI-DR address the management of spatial data and environmental tools and applications.

Other EU projects such as OBSERVE, EGIDA, Balkan GEONET, enviroGRIDS, BalkanGEONet, and GEONetCab have significant contribution to the development of the environmental network and observation capacity in the South East Europe.

The enviroGRIDS project gathers solutions and experience from many of these mentioned projects in order to approach the particularity of the Black Sea catchment region in terms of SDI, platforms interoperability (i.e. Geospatial and Grid, and software platforms like URM, gSWAT, ESIP, GreenLand, gProcess, eGLE, etc), high resolution models, processing scalability, user interaction usability, and processing efficiency.

3. BSC-OS PORTAL

The BSC-OS portal consists of a set of Web applications through which the users access the system resources such as spatial data, hydrologic models, environmental scenarios, data processing tools, visualization facilities, environmental reports, and training materials (Figure 1).

There are five categories of users such as data providers, earth science specialists, decision makers, citizens, and system administrators. The user may access the features of an individual application by local authentication, or all published applications of the portal by the single sign-on authentication.

The main user tool and application categories provided by the portal are [xx]:

- **Data management** – provides the user with spatial data management and operations. The user may enter data and metadata, visualize, modify, update, and remove spatial data from repositories;
- **Hydrologic model management** – provides the Earth Science specialists with hydrologic model configuration, scenario and model development, model calibration and scenario running. One of the water quality models used in the enviroGRIDS project is SWAT (Soil Water Assessment Tool) [6]. It is a model designed to estimate impacts of land management practices on water quantity and quality in complex watersheds. The SWAT model requires specific information about weather, soil properties, topography, vegetation, and land management practices of the watershed;
- **Satellite data processing** – the specialist may process satellite data and images in order to search for relevant information (e.g. land cover, vegetation, water, land use, soil
• Data visualization and report – the specialists visualize various spatial data in different formats and views and compose environmental reports for decision makers and citizens;

• Decision maker and citizen applications – provide the decision makers with the interactive and graphical tools to access the private environmental reports. The user may visualize data that make possible statistical analysis and predictions;

• Virtual training center – supports the specialists to develop Earth Science oriented training materials and the users to execute the lessons.

The regular users visualize the reports generated by the specialists as results of executing environmental scenarios. The input data for the reports are built up by the specialists by running hydrological models of the Black Sea catchment area and by processing related satellite data. All data sets required for building up the hydrological models, environmental scenarios, and spatial models are provided and entered into the system by the data providers.

4. DATA MANAGEMENT

The URM (Uniform Resource Management System) platform [7] allows users to search and share spatial and non-spatial information, and establish a network to encourage a broader community to adopt and support the GEOSS concept of data sharing for a more sustainable environment. The URM Geoportal is not the one integrated solution, but a set of modules and services, which are able to communicate through interoperable services defined by OGC (Open Geospatial Consortium), and W3C (World Wide Web Consortium). URM Geoportal consists of four basic blocks interconnected through metadata:

1. Metadata management is supported by the MicKa toolset for editing and management of metadata for spatial information, Web services and other sources;

2. Data management by the DataMan application. It supports the import, export, and management of spatial data in files or databases, for both raster (IFF/GeoTIFF, JPEG, GIF, PNG, BMP, ECW) and vector (ESRI Shapefile, DGN, DWG, GML) data types;

3. Data visualization, provided by the MapMan software tool. It supports publication of spatial compositions from locally stored data with external WMS (Web Map Service), WFS (Web Feature Service) data services;

4. Content management for publishing in context and connections with social networks, is supported by the SimpleCMS toolset.
5. SWAT MODEL CALIBRATION AND SCENARIO EXECUTION

The SWAT model supports the specialists on making predictions on the effects of management decisions on water, sediment, nutrient and pesticide yields with reasonable accuracy on large, engaged river basins [8]. The data package of the model could be quite large (up to 20 thousands) and its running requires great storage capacity and high power computation resources.

A. gSWAT Application

The gSWAT application has been developed in enviroGRIDS project and available through the BSC-OS Portal in order to support the development, calibration and execution of the SWAT model [9]. Grid based computation infrastructure is the basic solution for parallel and distributed processing of the hydrological model in the gSWAT application.

It is developed as a Web application that hides to the user the complexity of the Grid infrastructure (Figure 2). The application provides support for scalable models in terms of geographical area, modeling resolution, and number of users. Multicore architecture and GPU cluster based solutions are explored as well in order to speed up and optimize the hydrological model in the gSWAT application.

B. SWAT Oriented Services

gSWATSim is a server side extension of the gSWAT platform that is exposed as a collection of REST Web Services supporting the user to create new projects (i.e. new scenarios), modify some information related to projects (e.g. project name, description, etc.), run environmental scenarios, upload results to visualization module (i.e. BASHYT), and visualize the execution status of scenarios.

C. SWAT Model Development and Running

The hydrological model could be developed, calibrated and run through various approaches based on the gSWAT, gSWATSim services, and BASHYT platforms. The specialist could use the following solutions:

1. gSWAT application – The specialist develops the SWAT model by using ArcSWAT and ArcView tools on his desktop. By using the gSWAT application the user uploads the model onto the gSWAT server and performs interactively the calibration of the model [9]. The user controls the convergence to the optimal calibration (i.e. parameters, simulations, and iterations) by interactive techniques provided through the Web graphical user interface (Figure 2). Finally the user may download the resulted calibrated model.

2. gSWAT and BASHYT tools – The applications “work together” by separate working sessions that are connected just at the data level. The main advantage of this solution is the independency of the tools. The user performs the following steps: develop the SWAT model just in BASHYT, and then downloads the archived SWAT files and metadata. Now, follows the calibration by gSWAT as in the first solution. Finally the user uploads the results into BASHYT and visualizes the environmental information.
3. gSWATSim services – The applications work together through a common Storage Element and dedicated Web Services. The working session is in BASHYT through which the user develops the model and defines the scenario. The user exports to gSWATSim the data onto the Storage Element, then through the dedicated Web Service the execution environment is customized, and the scenario is executed. Scenario execution progress can be monitored directly in BASHYT. Finally after the execution, the results are available automatically into BASHYT for visualization. In this solution the user does not need to switch between the applications. BASHYT accesses a new functionality available through gSWATSim services, which allows the execution and the monitoring of running scenarios.

6. SATELLITE IMAGE PROCESSING

Satellite images could reveal information on land cover, precipitations, geographic areas, pollution, and natural phenomena. Spatial and environment related data could be acquired by imagery classification that is actually a data mining throughout the multispectral bands. It is a multivariable process taking in account satellite image types (e.g. MODIS, Landsat), particular geographic area, soil composition, vegetation cover, and generally the context (e.g. clouds, snow, and season). All these specific and variable conditions require flexible tools and applications to support an optimal search for the appropriate solutions.

One of the basic platforms supporting the development of the Grid oriented applications for satellite image processing are ESIP (Environment oriented Satellite Data Processing Platform) and gProcess [11]. ESIP supports a workflow based flexible description of the satellite images complex processing over the Grid. Actually ESIP includes as well the gridified GRASS functionality [12]. The gProcess platform supports the management and execution of workflows (i.e. task distribution, management of parallel and sequential tasks) over the Grid infrastructure.

The ESIP based applications have been developed according with the methodology reported in [13]. The BSC-OS portal publishes the GreenLand end user application that is accessible by Web browsers (Figure 3). The GreenLand application layers on the gProcess and ESIP platforms and extends the satellite image processing related functionality:

- Supporting the scalability, in terms of number of users, number of projects, number of workflows;
- By using the OGC Web services in order to search, visualize, fetch, and store the satellite images;
- Interoperability between GreenLand and URM is supported by standard OGC services (e.g. WMS, WCS, and WFS);
- GreenLand publishes satellite data by OGC services provided by GeoServer, and registered on the URM server;
- The GreenLand functionality and operators are published as WPS services (e.g. NDVI, EVI, and Accuracy Assessment);
- Two editors support the development of Basic Operators and Workflows. The first editor includes into the GreenLand platform the

![GreenLand platform and application for Grid based satellite image processing](image_url)
basic operators, which are used later to develop complex functionalities as workflows. The Workflow editor supports the diagrammatic description of complex processing to be executed over the Grid.

The main features of the GreenLand cover the requirements defined three main use cases in enviroGRIDS project: land cover monitoring for the Istanbul area in Turkey, Rioni River in Georgia, and Mosaic scenario related with the Black Sea catchment region.

7. DATA VISUALIZATION AND REPORTING

BASHYT (The Basin Scale Hydrological Tool) [14] is a Web based interface to SWAT that works together with ArcSWAT and AvSWAT [15]. It can be used to manage many watersheds/scenarios at once and exposes on the Web a template to produce environment oriented applications. The applications can be edited directly through the browser. BASHYT implements the Driving forces-Pressures-States-Impacts-Responses paradigm and is able to produce reports on environmental states that can be visualized in different ways.

In BASHYT the SWAT models are stored into a relational database. A preprocessing step is required to import raw data (vector, raster and tabular data) into the system. After importing SWAT models BASHYT could offer tables, charts, and maps in a transparent way to the end users.

8. CITIZENS ORIENTED APPLICATIONS

Two demonstrator Web applications for citizens have been developed within the enviroGRIDS project and available through the BSC-OS Portal. The first application, which is related to near real time dissemination of environmental data to citizens, a flood forecasting demonstrator is applied on the Somes Mare catchment in northern Romania. For the second application, related to long-term planning in river basins a demonstrator for long-term planning of remediation strategies regarding flooding, sediment and ecosystem problems along the Danube River section between the towns of Braila and Tulcea has been selected.

The first application is supported by the HEC-HMS [16] hydrological model executed over the Grid infrastructure, and the second one is supported by the SOBEK 1D/2D [17] hydrodynamic model of flow and sediment transport. Geospatial data is available through the enviroGRIDS URM Portal by standard OGC services, while for water-related time series data the emerging WaterML standard is used.

On the client side, for both applications the main interfaces are map-based (e.g. OpenLayers, Google maps, and Google Earth platforms), over which the additional data are overlaid as spatially distributed data, or point data containing time series of modeled results.

9. TRAINING MATERIALS

The BSC-OS Portal provides the virtual training center based on eGLE (GiSHEO eLearning Environment), developed initially through the GiSHEO project [18]. The training system has as generic users the teacher and the student. The teacher is the Earth Science specialist who authors teaching materials and coordinates the training sessions. The student is the trainee who accesses the teaching objects organized by lessons in order to get presentations, experiment algorithms on spatial data, process satellite images, execute environmental scenarios, and visualize reports already prepared by the specialists.

The teaching material is built as lessons in terms of templates, patterns, and tools. The Earth Science related content of the lessons may be fix or dynamically fetched from data repositories by standard OGC services such as WMS and WCS, (Figure 4).

The teacher may use the Grid based execution to process satellite images, to execute specific algorithms through workflow descriptions or to visualize previously created teaching resources (i.e. already processed satellite images, geographical maps, diagrams, algorithm workflow descriptions, etc.). The students have only the ability to execute the lessons according to the constraints established by the teacher. Depending on the interaction specified level, they could as well be allowed to describe and experiment new workflows (i.e. algorithms, scenarios) or choose different input data (e.g. satellite images, parameters) for existing ones.

10. CONCLUSIONS

The development of the BSC-OS portal and generally the research through the enviroGRIDS project have revealed a lot of challenges regarding the gathering data into a dedicated SDI, interoperability between Geospatial and Grid infrastructures, connections through standard OGC services, and interoperability between platforms developed by different partners (e.g. URM, gSWAT, ESIP, gProcess, GreenLand, gLite, BASHYT, and eGLE), huge spatial data sets involved in the development of hydrological models and environmental scenarios (e.g. Danube, Mosaic, Black Sea Catchment, Istanbul, and Rioni River in Georgia), security and access management in different platforms, application development in distributed and heterogeneous systems, etc.
Another issue the portal development has to face is the compatibility with new technologies and functional requirements. One main concern is the compatibility with the new European Middleware Initiative (EMI) that aims to improve and standardize the dominant existing middlewares in order to produce one simplified and interoperable middleware [19]. EMI attempts to unify a few Grid platforms such as ARC, gLite, Unicore and dCache. The EMI and Globus platforms will empower the EGI (European Grid Infrastructure) with more stable, usable and manageable software.

The main work aims to develop extended and high resolution models and scenarios, to improve the tool and application functionality, and to improve the user interaction techniques with spatial data models.

The service oriented architecture, multicore, GPGPU based systems, Cloud processing are other technologies that are explored in order to extend the scalability, interoperability, standard connectivity, functionality, usability of end user applications, system efficiency, and to improve the performance of data processing.

11. ACKNOWLEDGMENT

This research is supported by the FP7 enviroGRIDS Project funded by the European Commission, through the Contract 226740.

12. REFERENCES


334


---

**Prof. Dorian Gorgan** is PhD supervisor in Computer Science and Information Technologies at the Technical University of Cluj-Napoca. He is the chair of the Computer Graphics and Interactive Systems Laboratory. His fields of interest concern with graphical modeling and processing based on graphics clusters, Grid and Cloud 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, developing the first national Grid infrastructure for research and academic community. He is involved as well in other Grid related international research projects such as EnviroGRIDS, GiSHEO, and SEE-GRID-SCI and has developed satellite image processing tools, applications and platforms such as: gProcess, eGLE, ESIP, GreenView, and Greenland.

**Victor Bacu** is junior lecturer in Computer Science at the Technical University of Cluj-Napoca. He is the member of the Computer Graphics and Interactive Systems Lab. His research contributes to national and international projects including MedioGrid, SEE-GRID-SCI, GiSHEO and enviroGRIDS. His research interest focus on Grid and Cloud application development, cluster based visualization, computer graphics, graphical model and simulation, tool packages for flexible classification in the satellite images.

**Danut Mihon** is Researcher in Computer Science and Information Technologies at the Technical University of Cluj-Napoca. His fields of interest concern with satellite image processing over distributed architectures, spatial data processing and visualization, Earth Science oriented application development, Grid and Cloud computing. He has been involved in the SEE-GRID-SCI, GiSHEO and
enviroGRIDS research projects. He is the main developer of the GreenLand platform and Web application.

Denisa Rodila is Researcher in Computer Science and Information Technologies at the Technical University of Cluj-Napoca. She is a member of the 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, Cloud, cluster and service oriented technologies.

Teodor Ţtefanuţ is junior lecturer in Computer Science at the Technical University of Cluj-Napoca. He is the member of the CGIS Laboratory. His research contributes to national and international projects including mEducator, I-Trace, SEE-GRID-SCI,GISHEO, and enviroGRIDS. 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.

Dr. 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.

Dr. Gregory Giuliani, after obtaining a degree in Earth Sciences, 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 Developer in a local Swiss GIS company. He works at UNEP/GRID-Europe since 2001 and is the focal point for Spatial Data Infrastructure (SDI).

Dr. Nicolas Ray is the enviroGRIDS project manager. He holds a Master in Environmental Sciences, a PhD in Biology from University of Geneva in the field of Human Population Genetics, and a postgraduate certificate in Computer Sciences. Nicolas’ 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.

Dr. Anthony Lehmann is the EnviroGRIDS project initiator and coordinator. He holds a Masters Degree and a PhD 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.

Pierluigi Cau is an hydraulic engineer graduated at the University of Cagliari. Researcher for the Water Resources Management group, Energy and Environment Program of CRS4, since 2000.