Earth Science Oriented Data Processing Tools and Applications for Black Sea Catchment Basin

Dorian Gorgan
Computer Science Department
Technical University of Cluj-Napoca
dorian.gorgan@cs.utcluj.ro
# Workshop subjects

<table>
<thead>
<tr>
<th>Time</th>
<th>Subject</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:45 - 14:05</td>
<td>enviroGRIDS project overview: Portal, Tools and Applications</td>
<td>Dorian Gorgan</td>
</tr>
<tr>
<td>14:05 - 14:25</td>
<td>Grid based processing and data management</td>
<td>Victor Bacu, Denisa Rodila, Dorian Gorgan</td>
</tr>
<tr>
<td>14:25 - 15:15</td>
<td>SWAT model calibration and execution by gSWAT application</td>
<td>Victor Bacu</td>
</tr>
<tr>
<td>15:15 - 15:45</td>
<td>coffee break</td>
<td></td>
</tr>
<tr>
<td>15:45 - 16:00</td>
<td>Interoperability between Geospatial and Grid infrastructures</td>
<td>Denisa Rodila</td>
</tr>
<tr>
<td>16:00 - 17:00</td>
<td>Spatial data visualization by BASHYT</td>
<td>Pierluigi Cau, Simone Manca</td>
</tr>
<tr>
<td>17:00 - 17:15</td>
<td>Remote sensing applications</td>
<td>Victor Bacu, Dorian Gorgan</td>
</tr>
<tr>
<td>17:15 - 17:30</td>
<td>Training in Earth Sciences by eGLE</td>
<td>Dorian Gorgan</td>
</tr>
</tbody>
</table>
enviroGRIDS Project

- **enviroGRIDS** - Gridifying the Black Sea catchment to support its sustainable development ([http://www.envirogrids.net](http://www.envirogrids.net))

- Founded by the European Commission FP7 framework (Theme 6: environment), April 2009 – March 2013, 27 partners, 7.9 mil EUR.

- Coordinator
  - University of Geneva, Switzerland

- Objectives:
  - Develop a SDI (Spatial Data Infrastructure) targeting the Black Sea catchment region
  - Use new international standards to store, analyze, process, and visualize important information regarding this area
  - Perform distributed spatially-explicit simulations of environmental changes
BSC-OS Portal

- Single way of the user to get into the enviroGRIDS system
- Exposes personalized tools for different category of users: data manager, earth science specialist, decision maker, citizen, and system administrator.
- Provide applications for:
  - data management
  - hydrologic models calibration and execution
  - satellite image processing
  - report generation and visualization
  - virtual training center
- Support interoperability between the Geospatial and Grid infrastructures on security, heterogeneous data access, distributed data processing
- EnviroGRIDS functionality gathers services provided by various technologies such as SWAT related modules, Collaborative Working Environment (CWE), Uniform Resource Management (URM), gProcess, ESIP, and eGLE platforms
## Portal Architecture

### BSC-OS Portal

<table>
<thead>
<tr>
<th>Data Management Tools</th>
<th>Applications/ SWAT Scenarios</th>
<th>Visualization Tools</th>
<th>Decision Maker/ Citizen Tools</th>
</tr>
</thead>
</table>

### Geospatial and Grid Services

- **URM**
- **SWAT (GANGA)**
- **ESIP, gProcess**
- **CWE (BASHYT)**
- **Other Geospatial and Grid Technologies**
- **eGLE**

### Geospatial Oriented Level

- gLite Middleware

### Grid Infrastructure (EGEE)

- Spatial data, catalogues, maps
- Application data (hydrology, climate, soil, etc.)
- Scenarios
- Results of processing

---

OpenWater symposium and workshops, UNESCO-IHE, Delft, The Netherlands, 18-19 April, 2011
Data Flow Throughout the Portal
Specific Objectives

- Link, gather, store, manage and distribute key environmental data concerning the Black Sea Catchment Basin
- Large scale and high resolution distributed hydrological models
- Gridification of tools and applications
- Model and process huge spatial data over the Grid (e.g. hydrological models, satellite images, and maps)
- Develop early warning and decision support tools at regional, national and local levels
Hydrological Models - SWAT

- SWAT (Soil Water Assessment Tool)
  - hydrological model
  - operates on a daily time step
  - used for predicting the water resources, sediment, and chemical yields in a specific watershed

- Input data: weather, soil properties, topography, vegetation, and land management practices of the watershed

- SWAT estimates the impact of land management practices on water quantity and quality in complex watersheds

- The SWAT model must pass through a careful calibration and uncertainty analysis
Black Sea Catchment Basin
Grid Based Processing

- Many input data required by the SWAT model -> store the data on Storage Elements
- The calibration process requires running a high number of iterations, each iteration consisting in a high number of simulations -> execution on different Grid Worker Nodes
- Allows the calibration of the SWAT models and the execution of different scenarios based on a calibrated SWAT model on GRID infrastructures
gSWAT Application

- Support the SWAT model development in the BSC-OS Portal
- Model calibration and execution over the Grid infrastructure
- Project management of the hydrological model
- SWAT data visualization
- Project upload and download
- Support interaction with visualization tools
- Support interaction with scenarios development tools
Processing Steps

- Pre-processing step
- Iteration n
  - Simulation 1
  - Simulation n
- Post-processing step
- Test Calibration
- Calibrated model
gSWAT Application - Architecture

Web Portal
- SWAT Calibration
- SWAT Scenarios

SWAT Services
- Management and Execution
- Data Management
- Monitoring

Grid Infrastructure (gLite Middleware)

Model repository
Calibration outputs repository
gSWAT Control Flow

User

- gSWAT GUI

- gSWAT Services

- gSWAT DB

Ganga

DIANE master

Monitoring component

Worker

Worker

Storage Element

GRID oriented Tools

GRID

gSWAT system
gSWAT User Interface

- Project list and detailed information on the selected project
Modify SUFI2 Calibration Parameters

OpenWater symposium and workshops, UNESCO-IHE, Delft, The Netherlands, 18-19 April, 2011
Calibration Output Visualization
Spatial Data Visualization by BASHYT

Grid

User flow

Internal flow

User and internal flows are not simultaneous.
Each color is a different set of data

BSC-OS Portal

Preprocessing

Web Portal

Local Storage System (Cache)

End-Users
The output is shown on the portal using widgets like maps, charts or tables. They can be organized using HTML and Javascript.
Spatial Data Visualization

OpenWater symposium and workshops, UNESCO-IHE, Delft, The Netherlands, 18-19 April, 2011
Satellite Image Processing

- Explore images to identify objects and give them a significance
- Analysis: detect, identify, classify, measure and evaluate the significance of physical and cultural objects, their patterns and spatial relationships
- Applicability in different Earth Science domains
  - Land cover
  - Air pollution
  - Hydrology
  - Ecology
  - etc
Satellite Image Processing

- Satellite image types used for different purposes
  - Landsat
  - MODIS
  - Aster
  - Quickbird, etc

- Size: depending on the geographical region size and on the containing information, satellite images could reach a couple of Gb in size
GreenLand Vegetation Indices

- Vegetation index based image classification
  - NDVI (Normalized Vegetation Index)
  - EVI (Enhanced Vegetation Index)
  - IPVI (Infrared Percentage Vegetation Index)
  - GEMI (Global Environment Monitoring Index)
  - SAVI (Soil-Adjusted Vegetation Index)

- Classification process - combines different bands in the Landsat satellite images in order to correctly make the classification
GreenLand Application

- Classification output (right image) based on input satellite image (left image)
Virtual Training Center - Lessons

Lesson Visual Structure

Pattern 1: Lesson Title

Pattern 2: Static image Static text

Pattern 3: Dynamic text Video Slideshow

Pattern 4: Dynamic image Processing graph editor

Pattern 5: Animations - step by step results of complex processing algorithms

Images

Videos

Documents

3D Objects

Sounds

Other multimedia content
Teaching Material Description

Teaching materials of various topics and subjects.

Lesson description, execution, and management.

Workflow based process description.

Service description, search, discovery and composition. Services, operations provided through Web.

Teaching materials of various topics and subjects.

Lesson description, execution, and management.

Workflow based process description.

Service description, search, discovery and composition. Services, operations provided through Web.

Teaching materials of various topics and subjects.

Lesson description, execution, and management.

Workflow based process description.

Service description, search, discovery and composition. Services, operations provided through Web.

Teaching materials of various topics and subjects.

Lesson description, execution, and management.

Workflow based process description.

Service description, search, discovery and composition. Services, operations provided through Web.
This lesson exemplifies interactive scenario for teacher and students.
eGLE Tools

Conceptually, the water detection algorithm is quite simple and was published by Gond. The algorithm uses SWIR (Short-Wave Infrared), Red, and NIR (Near Infrared) spectral bands for detecting water areas. The implemented algorithm uses Landsat images as input and returns a GeoTIFF image that highlights the water boundaries. The spectral contrast of the band 3 red, band 4 near-infrared (NIR), and the band 5 mid-infrared (SWIR). The SWIR frequency is absorbed by water and is very sensitive to moisture. It is used to detect vegetation and soil moisture. In the near infrared band (NIR) the water absorbs nearly all the light and that makes the water becomes visible very dark in this band. The NIR band may be used to detect the land-water boundary. The NDWI index (Normalized Difference Water Index) is used to assess water content. The NDWI index increases with vegetation water content or from dry soil to free water. The difference NDWI-HDVI reinforces the perception of free water areas.
Lesson Samples – GRID Oriented Tools

The Normalized Difference Vegetation Index (NDVI) is a simple numerical indicator that can be used to analyze remote-sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not.

Live green plants absorb solar radiation in the photosynthetically active radiation (PAR) spectral region, which they use as a source of energy in the process of photosynthesis. Leaf cells have also evolved to scatter it, reflecting it in the blue and red wavelengths. The latter is why green appears so prevalent in the visible region. The blue wavelengths are absorbed approximately half of the total incoming solar energy, because the energy level per photon in that region (wavelength longer than about 700 nanometers) is not sufficient to be used by photosynthetic molecules. A strong absorption at these wavelengths would result in overheating the plant and possibly damaging the tissues. Hence, live green plants appear relatively dark in the PAR and relatively bright in the near-infrared. By contrast, cloud and snow tend to be rather bright in the red (as well as other visible wavelengths) and quite dark in the near-infrared.

Since early instruments of Earth Observation, such as NASA’s ERTS and NOAA’s AVHRR, acquired data in the red and near-infrared, it was natural to exploit the strong differences in plant reflectance to determine their spatial distribution in these satellite images. The NDVI is calculated from these individual measurements as follows:

\[
NDVI = \frac{NIR - RED}{NIR + RED}
\]
Lesson Execution Monitoring
Future Works

- Improve the functionality and the performance of the applications
- Tools and applications for the interactive development of hydrological scenarios
- Interoperability between different technologies involved in the enviroGRIDS project
- Include and develop new tools and applications required by the partners (e.g. RIONI River hydrological model, Vegetation development in Istanbul, etc)
- Develop training materials in the domain of Earth Sciences
Thank you for your attention!

Questions?

Dorian Gorgan
Computer Science Department
Technical University of Cluj-Napoca
dorian.gorgan@cs.utcluj.ro