

Remote Sensing Services

Title	Remote Sensing Services
Creator	Filiz BEKTAŞ BALÇIK (İstanbul Technical University, Turkey)
Creation date	2012.02.10
Date of last revision	2012.03.31
Subject	ESIP Platform and Hot Spot Inventory case studies
Status	Finalized
Type	Word document
Description	Reporting Hot Spot Inventory Case studies and ESIP Platform
Contributor(s)	Filiz Bektaş Balçık, Çiğdem Göksel, Seval Sozen, Karin Allenbach, Mamuka Gvilava, Kazi Rahman, Dorian Gorgan, Vasile Mihon.
Rights	Public
Identifier	EnviroGRIDS D2.11
Language	English
Relation	enviroGRIDS D2.1 (Interoperability guideline) enviroGRIDS D2.2 (Data storage guidelines) enviroGRIDS D2.4 (Remote sensing data use and integration guidelines) enviroGRIDS D2.6 (Transition plan for grid infrastructure) enviroGRIDS D2.7 (Grid-enabled Spatial Data Infrastructure) enviroGRIDS D4.1 (Data Collection) enviroGRIDS D6.1 (BSC network of GEO partners)

Abstract:

This document aims to provide a technical description of the Environment Oriented Satellite Data Processing Platform (ESIP). Built over the gProcess platform, ESIP offers a set of Web and Grid services, and image oriented basic operators. This platform is accessible through a web-based application called GreenLand, especially developed for remote sensing applications. Specific workflows have been requested through GreenLand and are currently being tested in three detailed case studies: a change detection analysis on İstanbul (Turkey) by computing vegetation indices and land cover/use classifications on Landsat 5 TM imagery, the hydrological modelling on Rioni River basin (Georgia) and the retrieval of large remote sensing products datasets for the entire Black Sea catchment.

GreenLand Grid-based application offers scalability when dealing with a large number of users and/or a large processing data volume. Due to the computing and storage capabilities offered by the Grid infrastructure, the workflow execution times are significantly reduced in comparison with standalone/cluster processing. Therefore this powerful tool will certainly be very useful for sustainable management of the Black Sea watershed.

An update of this deliverable will be provided by the end of the project and will include a detailed description of each fully-processed case study.

This deliverable is part of work package 2 Spatial Data Infrastructure. It is covering the following points:

CONTENTS

REMOTE SENSING SERVICES	1
LIST OF FIGURES	4
LIST OF TABLES.....	5
1. INTRODUCTION.....	6
1.1 DESCRIPTION OF THE WORK	6
1.2 CONTRIBUTORS.....	6
2 ESIP PLATFORM AND GREENLAND APPLICATION	9
2.1 ESIP PLATFORM	9
2.1.1 ESIP Overview.....	10
2.1.2 ESIP Architecture.....	10
2.1.2.1 Grid infrastructure.....	11
2.1.3 Workflow Based Description	12
2.1.3.1 Process Description Graph.....	12
2.1.3.2 Instantiated Process Description Graph	13
2.1.4 Basic Operators Library	13
2.1.5 GRASS Based Processing.....	14
2.2 GREENLAND APPLICATION	18
2.2.1 GreenLand Overview	18
2.2.2 GreenLand Architecture.....	19
2.2.3 Use Case Scenarios in GreenLand Application	19
2.2.3.1 Processing Description	20
2.2.3.2 Grid Based Process Execution	26
2.2.3.3 Process Execution Monitoring.....	28
2.3 ESIP AND GREENLAND SYSTEM.....	28
2.3.1 Generic Case Studies.....	28
3 HOT SPOT INVENTORIES	31
3.1 İSTANBUL CASE STUDY	31
3.1.1 Introduction.....	31
3.1.1.1 Purpose and Scope	31
3.1.2 Study Area.....	31
3.1.3 Data and Methodology.....	32
3.1.3.1 Radiometric Correction (DN to Reflectance Conversion).....	32
3.1.3.2 Atmospheric Correction (Dark Object Subtraction).....	33
3.1.3.3 Vegetation Indices (NDVI and EVI)	34
3.1.3.4 Classification (Density Slicing).....	35
3.1.3.5 Classification Accuracy Assessment	35
3.1.4 Results (ENVI).....	37
3.1.5 Results (ESIP Platform).....	42
3.1.5.1 Grid Based Process Execution	42
3.2 RIONI RIVER BASIN CASE STUDY	43
3.2.1 Introduction	43
3.2.1.1 Rioni River Basin and Case Study Description	43
3.2.2 Application Methodology	44
3.2.2.1 Users Profile	45



3.2.2.2	Applications	45
3.2.2.3	Functionality	46
3.2.2.4	Use Scenarios	47
3.2.3	Data Resources	48
3.2.3.1	Global (Remote Sensing)	48
3.2.3.2	Local (hydrometeorology, soils)	52
3.2.4	Results (SWAT)	55
3.2.4.1	Land (hard) data	55
3.2.4.2	Climate (soft) data	56
3.2.4.3	Hydrological modelling results	57
3.2.5	Recommendations for model improvements	58
3.2.5.1	Specifications for ESIP and GreenLand Based Processing	59
3.2.6	Synergy with PEGASO (Georgia CASE Site)	59
3.3	REMOTE SENSING PRODUCTS FOR SWAT OUTPUTS VALIDATION	61
3.3.1	Introduction	61
3.3.2	Remote sensing products	61
3.3.2.1	RSP commonly used in hydrologic research	61
3.3.2.3	Quality assessment	62
3.3.3	Evaluation on Vit river basin, Bulgaria	62
3.3.3.1	Data Storage	63
3.3.4	Processes	63
3.4	GREENLAND BLACK SEA CATCHMENT APPLICATION	64
3.4.1	Process description	65
3.4.1.1	Products	65
3.4.1.2	Upload	66
3.4.1.3	Processes (band extraction, conversion, mosaic, subset)	67
3.4.1.4	Export	69
3.4.2	Iterations	69
3.4.3	File size	69
3.4.4	ESIP application at Black Sea catchment extent	70
3.4.4.1	ESIP and GreenLand Based Processing	70
3.4.5	Future improvements	71
4	CONCLUSION AND RECOMMENDATIONS	71
5	ABBREVIATIONS AND ACRONYMS	72
6	REFERENCES	74
	APPENDIX 1 RIONI HYDRO-METEOROLOGY	78
	HYDROLOGY	78
	Hydrologic Stations of Rioni River Basin	79
	Description of Hydrologic Stations along Rioni River	79
	Existing Reservoirs	81
	Planned Reservoirs	82
	Metadata	82
	METEOROLOGY	82
	Description of Meteorology Stations along Rioni River	82
	Metadata	84

List of figures

FIGURE 2.1 ESIP RELATED ARCHITECTURE	10
FIGURE 2.2 PDG INTERNAL STRUCTURE	13
FIGURE 2.3 GREENLAND SYSTEM RELATED ARCHITECTURE	19
FIGURE 2.4 USER'S PROJECT DETAILS MODULE	21
FIGURE 2.5 ADD WORKFLOWS TO PROJECT	23
FIGURE 2.6 UPLOAD LOCAL FILES PANEL	24
FIGURE 2.7 OGC SERVICE BASED DATA RETRIEVAL PANEL	25
FIGURE 2.8 FTP BASED DATA RETRIEVAL PANEL.....	25
FIGURE 2.9 SPECIFY THE WORKFLOW INPUTS	26
FIGURE 2.10 GRID BASED EXECUTION MONITORING PANEL	28
FIGURE 2.11 GRASS SCRIPT FOR WORKFLOW DEFINITION	29
FIGURE 2.12 PDG AND IPDG DEFINITIONS FOR THE NDVI WORKFLOW.....	30
FIGURE 3.1 STUDY AREA (İSTANBUL)	31
FIGURE 3.2 FLOWCHART OF İSTANBUL CASE STUDY	32
FIGURE 3.3 IMAGE CLASSIFICATION.....	35
FIGURE 3.4 THE CONFUSION MATRIX AND SOME COMMON MEASURES OF CLASSIFICATION ACCURACY THAT MAY BE DERIVED FROM IT. ADOPTED FROM FOODY, (2002).....	36
FIGURE 3.5 PRE-PROCESSED DATA (MOSAICKED AND SUBSET) A) 1987 B) 2009	37
FIGURE 3.6 RESULTS OF VEGETATION INDICES (NDVI, AND EVI) A) 1987 NDVI B) 2009 NDVI C) 1987 EVI, AND D) 2009 EVI	39
FIGURE 3.7 DENSITY SLICING CLASSIFICATION A) 1987 NDVI B) 2009 NDVI.....	40
FIGURE 3.8 DENSITY SLICING CLASSIFICATION A) 1987 EVI B) 2009 EVI.....	41
FIGURE 3.9 SELECTING THE EVI INPUT DATA SET	42
FIGURE 3.10 EVI WORKFLOW USAGE RESULT	43
FIGURE 3.11 ENVIROGRIDS FUNCTIONAL ARCHITECTURE	48
FIGURE 3.12 TRMM PRECIPITATION TIME SERIES 0.25° X 0.25° GRID POINTS, FALLING WITHIN THE RIONI BASIN.	51
FIGURE 3.13 HYDROLOGICAL AND METEOROLOGICAL STATIONS IN THE RIONI BASIN (BOX INDICATES LOCATION OF NAMAKHVANI HYDROPOWER CASCADE WITH 3 PLANNED RESERVOIRS ON R. RIONI)	53
FIGURE 3.14 SOILS OF GEORGIA AND IN THE RIONI BASIN.....	54
FIGURE 3.15 GLOBAL SOURCES USED IN HYDROLOGICAL MODEL SETUP FOR RIONI (SOILS, LAND COVER, DEM). (BACKGROUND SHOWS SUMMER CHLOROPHYLL IN THE BLACK SEA AND SAMPLE TRMM GRID ON LANDWARD SIDE.).....	56
FIGURE 3.16 SWAT SETUP FOR RIONI RIVER COLLATING LOCAL HYDROMET & GLOBAL TRMM PRECIPITATION DATA. NOTE: OBSERVED DATA SEVERAL KM UPSTREAM THE SIMULATED DATA (ALPANA VERSUS OUTLET 8).	57
FIGURE 3.17 OBSERVED VERSUS SWAT SIMULATED DAILY FLOW CALIBRATION	58
FIGURE 3.18 INPUT SOURCES, DATA FLOW AND PROCESSING SUGGESTIONS FOR ESIP BASED PLATFORM	59
FIGURE 3.19 PEGASO/EG GEORGIA CASE SITE (LEFT) & INDICATOR MAP FOR TSKALTSMINDA COASTAL UNIT (RIGHT)	60
FIGURE 3.20 DATA DOWNLOAD AND PREPROCESSING WORKFLOW, EXAMPLE FOR MCD15 PRODUCT	64
FIGURE 3.21 GENERAL PROCESS WORKFLOW.....	65
FIGURE 3.22 MODIS TILES PRODUCTS TO COVER THE EXTENDED BLACK SEA WATERSHED, HORIZONTAL (H) AND VERTICAL (V) POSITION	68
FIGURE 3.23 DROP-DOWN LIST OF ACCESSIBLE PRODUCTS FOR GREENLAND'S TOOLS	68
FIGURE 3.24 NUMBER OF EXPECTED PROCESSED PRODUCTS ACCORDING WORKFLOW STEPS.....	69
FIGURE 3.25 MOSAIC OF 12 TILES OF LAI BAND (MOD15) ON 23 APRIL 2011	70
FIGURE 3.26 MOSAIC WORKFLOW USAGE RESULT	71



List of tables

TABLE 2.1 ARITHMETIC OPERATORS.....	13
TABLE 2.2 VEGETATION INDEX OPERATORS.....	14
TABLE 2.3 CORRECTION OPERATORS.....	15
TABLE 2.4 BAND MANAGEMENT OPERATORS.....	15
TABLE 2.5 VIEW AND PSEUDO-COLOR OPERATORS.....	16
TABLE 2.6 STATISTICAL OPERATORS.....	16
TABLE 2.7 CLASSIFICATION OPERATORS.....	16
TABLE 2.8 CLASS FILE FORMAT.....	17
TABLE 2.9 SCOPE OPERATORS.....	17
TABLE 3.1 REMOTE SENSING INDICES FORMULAS WERE USED IN THE STUDY TO DEVELOP NIR, RED, BLUE DENOTES REFLECTANCE VALUES DERIVED FROM LANDSAT 5 TM BANDS OF B4, B3, AND B1, RESPECTIVELY AND C 1 =6, C 2 =7.5 AND L=1 FOR EVI INDEX.....	34
TABLE 3.2 GLOBAL RS DATA RESOURCES FOR SWAT APPLICATION.....	49
TABLE 3.3 LIST OF PROVIDED PRODUCTS AND THEIR SELECTION FOR SWAT RESULTS VALIDATION.....	62
TABLE 3.4 LIST OF PROVIDED PRODUCTS, COVERED PERIOD, NUMBER OF DIFFERENT DATES SUPPLIED, TOTAL FOLDER SIZE.....	63
TABLE 3.5 TESTED PRODUCT DESCRIPTION.....	66

1. Introduction

1.1 Description of the work

The society needs and environmental predictions encourage the applications development, oriented on supervising and analyzing different Earth Science related phenomena. This task concerns with experimenting solutions by using visual descriptions of the satellite image processing over distributed infrastructures, such as Grid infrastructure. The GreenLand application is developed over the ESIP (Environment Oriented Satellite Data Processing Platform) and gProcess platforms and it is based on simple, but powerful, notions of mathematical operators and workflows that are used in distributed and parallel executions over the Grid infrastructure.

ESIP could be defined as a set of Web and Grid services, and image oriented basic operators such as arithmetic (e.g. addition, division, exponent, complement etc), radiometric transformations (e.g. histogram equalization, mean, histogram scaling etc), spatial domain transformations (e.g. blurring, convolution etc), edge and line detection (e.g. gradient transformation), pseudo coloring, etc. The gProcess platform is used as a middleware between the client application and the Grid infrastructure.

This task is responsible for two different subjects such as building the Environment Oriented Satellite Data Processing Platform (ESIP) based on gProcess for environmental applications and determining specifications for the platform based on three different hot spot inventory studies. The GreenLand application is free Geographic Information System (GIS) software for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization.

At the first stage, İstanbul case study conducted to examine basic image processing steps of medium resolution Landsat 5 TM data to derive land use and land cover categories especially vegetated regions by using different vegetation indices and density slicing classification.

A specific workflow, using GreenLand application, will be developed to facilitate the retrieval of large dataset of existing Remote Sensing Products for the entire Black Sea watershed. First test implementation will be run on two MODIS products; Leaf Area Index and Evapotranspiration as they have revealed a great interest for validating SWAT results on a test site; Vit river basin in Bulgaria. Such an application will offer an effective solution to improve data retrieval efficiency by giving an alternative of processing very large datasets without caring about storage and calculation capacities. Moreover, by connecting the Greenland application into the enviroGRIDS URM geo portal will encourage the reuse of already processed data making those datasets easily accessible to enviroGRIDS consortium and users.

Data resources and methodologies outlined above are applied to handle hydrological modelling in the case of Rioni River Basin, aiming to demonstrate the utility and compatibility of the combination of global remote sensing and local hydrometeorological and other datasets in research and characterization of the catchment hydrology. Objective of this particular case is to apply the remote sensing and modelling tools (such as SWAT & ESIP), constituting integral components of the enviroGRIDS BSC Observation System.

1.2 Contributors

İstanbul Technical University: <http://www.itu.edu.tr/>

Filiz Bektas Balciik

Dr. Filiz BEKTAS BALCIK is an Assistant Professor at İstanbul Technical University (ITU), Department of Geomatic Engineering. She received her PhD degree in 2010 from the same faculty at ITU. Her PhD subject is related with wetland vegetation discrimination and mapping by using multispectral and hyperspectral remotely sensed data. She did a part of her PhD research at International Institute for Geo-Information Science and Earth Observation (ITC), Natural Resource Department, in the Netherlands while she was a Huygens Nuffic PhD Scholar (2 years). She was involved in several environmental projects as a

researcher in the field of remote sensing applications on land use/land cover change detection, urbanization, forestry, and biophysical and biochemical characteristics of savanna vegetation (in the South Africa). Within the enviroGRIDS project, she will be involved in WP2 (Task 2.4 and Task 2.11) and WP5 (Biodiversity, Health and Energy tasks) as Remote Sensing and GIS analyst.

Cigdem Goksel

Dr. Cigdem GOKSEL is an Assistant Professor at the Istanbul Technical University (ITU), Department of Geomatic Engineering. She received her B.Sc. degree from Geodesy and Photogrammetry Engineering from ITU, Faculty of Civil Engineering in 1984. She received her graduate degrees M.Sc. (1989) and Ph.D. (1996) from the same faculty in the field of Geodesy and Photogrammetry Engineering. She was visiting scholar at Murray State University's Mid-America Remote Sensing Center (Geosciences) KY – US during six months in 1999. Her research interests are in monitoring of landuse/landcover changes and in remote sensing and GIS integration for environmental studies. She supervised several research studies and projects in the field of remote sensing and integrated technologies. She has 75 scientific publications related with different remote sensing applications such as water basin monitoring, urbanization detection, coastal line and epidemiologic studies.

She will be involved in enviroGRIDS in WP2 ((Task 2.4 and Task 2.11) and WP5.

Seval Sözen

Prof. Dr. Seval SOZEN is professor at the Istanbul Technical University (ITU), Department of Environmental Engineering. She received her B.Sc. degree from Environmental Engineering from ITU, Faculty of Civil Engineering in 1985. She received her graduate degrees M.Sc. (1987) and Ph.D. (1995) from the same faculty in the field of Environmental Engineering.

Dr. Sözen has 26 years of teaching and research experience in the field of Environmental Science and Technology. Her research spans environmental biotechnology (activated sludge modeling, biological nutrient removal etc.), industrial pollution control, waste management, integrated water management and water quality modeling. She has directed and supervised numerous research studies and projects in the field of wastewater management, industrial pollution control, kinetics of biological processes and integrated water management. She holds a long list of scientific publications with over 100 papers, which received more than 500 citations.

Her key personal and academic skills include; broad knowledge of environmental and other sciences, strong skills in applied environmental sciences, strong understanding of biological processes theory gained from undergraduate and postgraduate study and her research experience, proven skill in water and wastewater treatment plant design and modeling, experience in writing research and other proposals for funding and other agencies throughout her academic career, proven organizational and academic management skills, commercial and governmental consultancy experience, extensive publishing, editing, and presentation track record.

Dr. Sözen supervised five European Commission (EC) funded projects at ITU. She is also trainer in Emergency Management certified by the Federal Emergency Management Agency, USA. She was awarded an incentive by the Scientific and Technical Research Council of Turkey in the engineering branch in 2000.

She will be involved in EnviroGRIDS as the leader of WP5 and will specifically contribute in biodiversity, health and energy tasks. She will also promote WP2 in the evaluation of remote sensing data.

UNIGE: <http://www.unige.ch/enviospace> & **UNEP/GRID-Geneva :** <http://www.grid.unep.ch>

Karin Allenbach

Karin Allenbach obtains a multidisciplinary master in Earth Sciences. She pursued a postgraduated certificate specializing in remote sensing and GIS. In 2004, she started to work at UNEP/GRID-Geneva, where she was monitoring algal bloom and identifying land-based pollution on coastal water in the eastern part of the Mediterranean using low resolution imagery (SeaWiFS). Since then, she has been involved in



numerous projects using low to very high resolution imagery to monitor and classify the environment on an object analysis basis.

- IMOS: monitoring re-flooding and vegetation growth on Iraqi marshlands using low resolution imagery (MODIS) <http://imos.grid.unep.ch/>

- SwissED: Swiss Environmental Domain classification consists on grouping together pixels of similar environment rather than geographic, creating an innovative spatial framework for environmental reporting.

- Landcover-landuse classification of Geneva's canton and surrounding based on aerial photography and medium resolution imagery (SPOT)

- Cartography of natural habitats of Geneva's canton and surrounding based on aerial photography and medium resolution imagery (SPOT)

In enviroGRIDS project, she will be involved in WP2 (Task 2.4 and Task 2.11) as a Remote Sensing Analyst.

GeoGraphic : <http://www.GeoGraphic.ge>

Mamuka Gvilava

Mamuka Gvilava is engaged by the GIS and RS Consulting Center "GeoGraphic", Georgia, as the Task Manager responsible for GeoGraphic's contribution into the enviroGRIDS Project. He is specialized in a wide range of environmental subjects with more than 10 years of experience gained with cooperative projects in Georgia and in the Black Sea region.

Physicist with education and science degree, re-expanded his field of activities to environmental and social subjects, such as integrated coastal management, environmental informatics (GIS & RS, with particular experience in setting up the coastal monitoring and information systems for Georgia), environmental and social impact assessment, as well as the application of green design and development methodologies.

He combines his practical hands-on experience in environmental management and research with wider scope of policy advice serving the Ministry of Environment as the National Focal Point on Integrated Coastal Zone Management (ICZM) and the Georgian Member of the ICZM Advisory Group to the Black Sea Commission. Mamuka Gvilava is based in Tbilisi, Georgia, and is employed by "GeoGraphic" to lead its enviroGRIDS team in implementing WP2 (Task 2.4 and Task 2.11) and WP7.

UNIGE: <http://www.unige.ch/envirospace> **ACQWA [Assessing climate change impact on quality and quantity of water. www.acqwa.ch]**

Kazi Rahman:

Expertise area: hydrology, watershed modelling, climate change impact assessment in complex environment, currently pursuing his PhD program at the University of Geneva. He obtained his M.Sc in Water Resources Engineering from Vrije university of Brussels focusing hydrodynamic model for flood frequency analysis.

Mr. Rahman has sound understanding on Soil and Water Assessment Tool (SWAT), an open source tool to analyze complex environmental issues like discharge, water quality, sediment transport and soil moisture for agricultural production.

Besides hydrological modelling he is familiar with high performance computing application in watershed modelling. He has experience working with computing language like MATLAB, FORTRAN and R. His PhD is funded by the EU large scale research project ACQWA [Assessing climate change impact on quality and quantity of water. www.acqwa.ch]. In the enviroGRIDS project he enjoys working as volunteer in data processing for developing hydrological model.

The Technical University of Cluj-Napoca (UTC) <http://www.utcluj.ro/>



Dorian Gorgan

Dorian Gorgan is professor and PhD supervisor in Computer Science at the Technical University of Cluj-Napoca. He graduated Computer Science and Automation from "Politehnica" University of Timisoara, has PhD in Graphical Modelling and Simulation, Visual Programming, and Graphical User Interfaces at the Technical University of Cluj-Napoca. For two years he gave technical and scientific consultancy in Autodesk's projects (in Milan, Italy) in the fields of location based services (LBS) and geographical information systems (GIS). He is the chair of the CGIS (Computer Graphics and Interactive Systems) Laboratory, director of the MedioGrid project, a Grid software and hardware infrastructure supporting the development of the environmental and geographical applications.

Danut Mihon

Danut Mihon is a second year PhD student at Technical University of Cluj-Napoca 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 and it is involved in several environmental and remote sensing projects funded by the European Union. His work is also confirmed by the national and international list of publications.

2 ESIP platform and GreenLand Application

2.1 ESIP platform

The Environment oriented Satellite Data Processing Platform (ESIP) (Băcu et al., 2010) is based on the gProcess (Băcu et al., 2009) platform. ESIP could be defined as a set of Web and Grid services, and image oriented basic operators such as arithmetic (e.g. addition, division, exponent, complement etc), radiometric transformations (e.g. histogram equalization, mean, histogram scaling etc), spatial domain transformations

(e.g. blurring, convolution etc), edge and line detection (e.g. gradient transformation), pseudo coloring, geometric transformations (e.g. rotation, scale etc), statistics (e.g. compute histogram, standard deviation pixel values etc), and spatial domain transforms (e.g. forward and inverse Fourier transformation etc).

2.1.1 ESIP Overview

The ESIP platform uses process description graphs to define algorithms implemented as executable workflows. A process description graph (PDG) is represented like a DAG (Direct Acyclic Graph) by composing basic operators, services and sub-graphs. The operators are always identified as atomic components and represent the smallest ESIP unit that could be executed without further decomposition. Some components can be already developed as Grid or Web services, and by combining this functions, we could benefit of code reusability. In some cases, a subset of a graph can be used in creating hyper-graphs that in the Grid execution phase are expanded into their corresponding basic operators. The PDG represents a pattern that does not specify any particular inputs for these operators. We are specifying only the input type (e.g. satellite image, integer, float, string, etc.) but not the concrete value. By attaching input resources to the PDG, we obtain an iPDG (instantiated Process Description Graph). Only the iPDGs can be executed over the Grid infrastructure.

2.1.2 ESIP Architecture

The ESIP platform is an interactive toolset supporting the flexible description, instantiation, scheduling and execution of the Grid processing. The ESIP is built over the gProcess platform that represents a collection of Grid services and tools providing the following basic functionality:

Visual manipulation based interactive description of the Grid based satellite image processing by pattern workflow like directed acyclic graph (DAG);

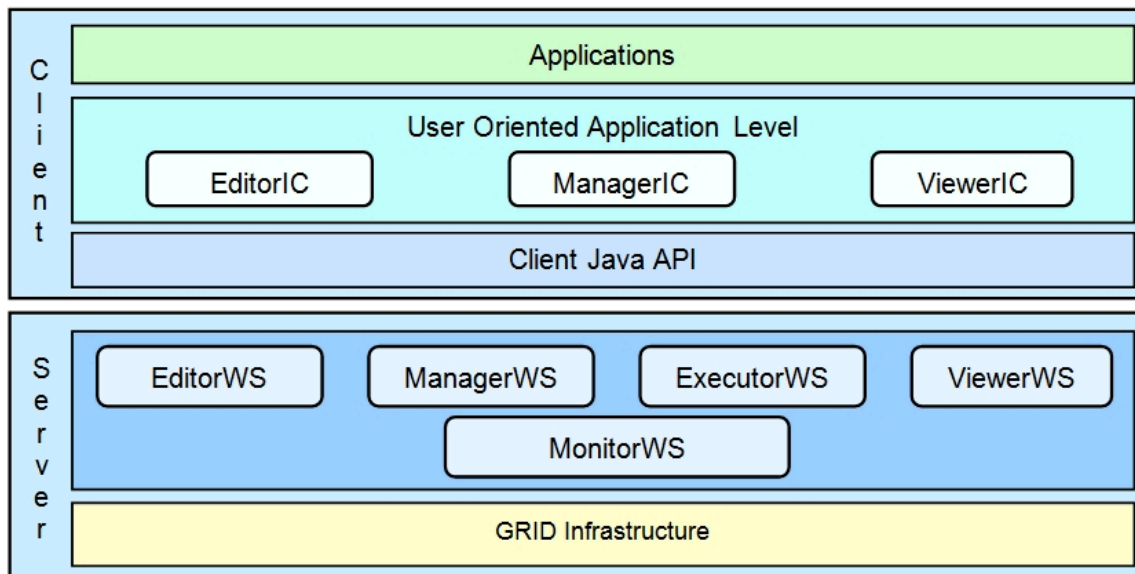
Development of hyper-graphs as combinations of basic operators, Grid and Web services, and sub-graphs;

Pattern workflow instantiation for particular satellite image;

Satellite data management, access and visualization;

Process execution control and visualization;

Optimal execution for appropriate mapping of the processing over the Grid resources. The optimal processing is achieved in terms of code optimization, total execution time, and data communication costs over the Grid.



The ESIP structure, from architectural point of view, is represented by the gProcess internal design that is based on a client-server model (Figure 2.1). Different components (Editor, Manager and Executor) that have a Web service interface and a user interaction component, make part of the ESIP platform. The Web services are used to expose different functionality and the interaction components respond to the requests initiated by users at the GreenLand graphical interface level. The EditorWS provides information about available resources (e.g. lists of operations, sub-graphs, services, satellite image types, data types etc). Information about workflows (i.e. PDG, iPDG) and fetching and uploading resources (e.g. workflow, operators, services, sub-graphs, data, etc.) are retrieved by the ManagerWS.

The ExecutorWS executes instantiated workflows (iPDG) over the Grid infrastructure. Another feature of the ExecutorWS is the monitoring of workflows execution. The EditorIC supports the user's editing operations for workflow development. The ManagerIC instantiates workflows for particular satellite data and manage the model resources (e.g. operators, services, sub-graphs, satellite data, etc.). The ViewerIC displays the input and output data (e.g. initial and processed satellite images) on the client site, and gets and displays the monitoring data.

At the server side the Grid infrastructure supports the access to computing resources and distributed databases through a set of services such as EditorWS, ManagerWS, ExecutorWS, and ViewerWS. The client side encapsulates Client Java API, User Oriented Application Level, and the software applications, such as GreenLand. The Web and desktop applications access the gProcess services by the Client Java API level. A set of user interaction supporting components, such as EditorIC, ManagerIC, and ViewerIC are exposed through the User Oriented Application Level. The last level of Applications combines the editor, manager, and viewer related functionality into complex user interaction tools.

2.1.2.1 Grid infrastructure

The Grid infrastructure (Mihon et al., 2010) is the execution environment where all the data processing takes place. Every environmental application has a characteristic set of features, corresponding to the applicability domain (e.g. meteorology, soil pollution, water pollution, etc). These applications require large input data sets, that mainly consists of satellite images, like: MODIS (Moderate Resolution Imaging Spectroradiometer), Landsat, Meris or QuickBird (some of them up to 1Gb in size).

Another important aspect regards the fact that all the applications in the Earth Science domain are using algorithms based on big sets of parameters that have to be combined in a certain way in order to obtain the correct results.

This means that on a standalone machine, obtaining reasonable execution times is almost impossible. Grid infrastructure offers the solution to this problem, by providing parallel and distributed computation methods. This means, that an application is divided into atomic components that could be executed on more than one Grid nodes. In order to run a Grid based application, all the input datasets and dependency files should be sent to a Storage Element inside the Grid infrastructure. Then the execution step follows, right after the files are copied with success.

Grid applications are, in general, very complex and time consuming. Keeping in mind the distributed nature of the Grid, the applications were divided into atomic components (jobs), which can be executed in parallel on the Grid nodes. There are many languages for the description of the environmental applications that allow the user to specify the characteristics of their applications to be run on the Grid. These languages are divided in two main categories: script-based and workflow-based. There are several languages used to describe these characteristics: JSDL (Job Submission Description Language), GXML (Guideline XML) or JDL (Job Description Language). Most of these languages assume that the applications can be represented as a set of jobs. This entire workflow is graph-structure based, where the nodes represent the jobs and the edges describe the execution path of the application.

All the GreenLand data processing is executed over the envirogrids.vo.eu-egee.org Virtual Organization (VO). The RO-09-UTCN site is part of this VO and it contains 1024 worker nodes and 13TB of storage space. Some tests were also conducted on the AM-02-SEUA site that encapsulates 48 worker nodes.

2.1.3 Workflow Based Description

2.1.3.1 Process Description Graph

The Process Description Graph developed in the scope of the gProcess platform is defined as a directed acyclic graph (DAG). The nodes describe resources, processing entities and the connections between those components. The following types can be used to define the structure of such a graph: resources, operators, services and sub-workflows (Figure 2.2).

Resource nodes: encapsulate a unique name and a general description. They may represent any of the following: images, temperature files, including primitives like strings, integers or floats;

Operator nodes: have the same basic description as above but they contain a link toward an executable file which gives the node its processing capabilities;

Service nodes: are described by service name and description as well as by method name and its corresponding description;

Sub-graph nodes: represent the workflow of an algorithm and provide a structure which is easily extendable. It facilitates the possibility to include a graph inside another graph thereby providing the possibility to combine different algorithms together and execute them as a single one.

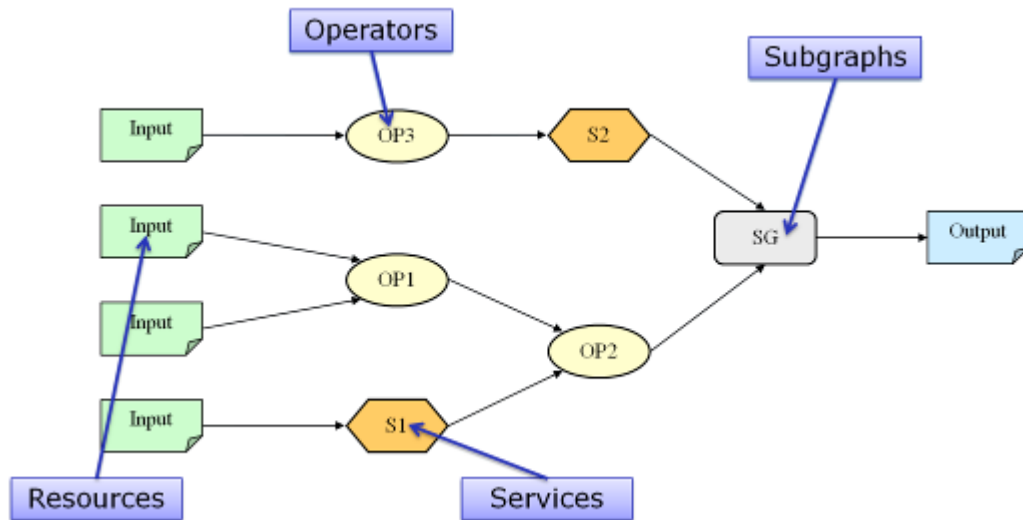


Figure 2.2 PDG internal structure

2.1.3.2 Instantiated Process Description Graph

The instantiated Process Description Graph is a variant of the PDG; it provides additional information to the later, instantiating the resource nodes by specifying the values for the inputs. There are two types of resources which can be used: files or primitive types. The previous are instantiated by a path towards the input file while the second represents a string which can be parsed to the type mentioned within the PDG description file.

2.1.4 Basic Operators Library

The GreenLand application is based both on simple and complex operators, which gives the user the possibility to choose the granularity that he desires. Complex operators can be created from simpler ones by virtue of the Process Description Graphs, which may be used to compose the fine grained operations based on the simple or basic operators.

To be noted is the fact that all operations performed in GreenLand are done through Grass operators (Neteler and Mitasova, 2004), be they simple/basic operators or alternatively complex/GRASS based processing. Currently there exist a series of basic operators included within the GreenLand application. They represent mainly arithmetical operators as listed in the following tables:

Table 2.1 Arithmetic operators

Arithmetic Operators		
1	Add- Adds the value of 2 single band images pixel by pixel.	Inputs: 1.Single band GeoTiff. 2.Single band GeoTiff. Outputs: 1. Single band GeoTiff.
2	Subtract-Subtracts the value of the second single band image from the first single band image pixel by pixel.	

3	Multiply- Multiplies the value of 2 single band images pixel by pixel.	Notes: Operations are done at pixel level.
4	Divide-Divides the first single band image by the first one pixel by pixel.	
5	AddConstant-Adds a constant value to the single band image.	Inputs: 1. Single band GeoTiff. 2. Constant value represented by a floating point number. Outputs: 1. Single band GeoTiff.
6	SubtractConstant- Subtracts a constant value form the single band image.	
7	MultiplyConstant- Multiplies the single band image by a constant value.	
8	DivideConstant- Divides a single band image by a constant value.	

Future operators may include: Convolution Operators, Forward Fourier Transform, Inverse Fourier Transform and further Mathematical Operators (OR, XOR, LOG, EXP, etc). The list of operators is not a complete one, since no specific request for implementation has been issued so far.

2.1.5 GRASS Based Processing

Grass based processing also includes the basic operators set. What detaches the operators described in this section from the rest is the fact that they operate exclusively on satellite images. The ones described above can operate on any given type of image.

The project so far has been interested in developing operators for raster images but future iterations and developments may also include vector images and 3D raster images. The raster Grass Operators implemented so far could be classified as: Vegetation Index, Correction, Band Manipulation, View and Pseudo-Color, Statistics, Classification, Scope and Miscellaneous Operators.

Table 2.2 Vegetation index operators

Vegetation Index		
1.	NDVI-Normalized Difference Vegetation Index Takes two singleband images and creates a Normalized Difference Vegetation Index image	Inputs: 1.Image representing the Red band (Geotiff) 2.Image representing the NIR band (Geotiff) Outputs: 1. NDVI image band (Geotiff). Each pixel is in the [0,1] range
2.	EVI-Enhanced Vegetation Index Takes three singleband images and creates an Enhanced Vegetation Index image, highlighting areas of increased vegetation	Inputs: 1.Image representing the Blue band (Geotiff) 2.Image representing the Red band (Geotiff) 3.Image representing the NIR band (Geotiff)

		Outputs: 1. EVI image band (Geotiff). Each pixel is in the [0, 1] range
3.	GEMI-Global Environmental Monitoring Index Generates a Global Environmental Monitoring Index image, based on two singleband images	Inputs: 1. Image representing the Red band (Geotiff) 2. Image representing the NIR band (Geotiff) Outputs: 1. GEMI image band (Geotiff). Each pixel is in the [0, 1] range

Table 2.3 Correction operators

Correction		
1.	DOS-Dark Object Subtraction Removes haze caused by atmospheric phenomena. It operates on a singleband image	r.mapcalc – Removes haze from image using a dark object. One may specify the size of the object (pixels).
2.	Landsat Atmospheric Correction- operator is used in removing any kind of atmospheric interference	Inputs: 1. Landsat singleband image (Geotiff) 2. MTL or MET file representing atmospheric conditions as detected by the satellite or added at a later stage Outputs: 1. Geotiff image with atmospheric interference removed

Table 2.4 Band management operators

Band Manipulation		
1.	Extract Band- Extracts bands from Geotiff multiband images	Inputs: 1. Multiband image from which a certain band will be extracted (Geotiff) 2. Band to be extracted (Integer) Outputs: 1. Extracted singleband image (Geotiff)
2.	Combine Bands- Generates a multiband image, based on two single banded (or multibanded) input images	Inputs: 1. Multiband or singleband image (Geotiff) 2. Multiband or singleband image (Geotiff) Outputs: 1. Multiband Geotiff image, which represents a combination of the previous in the order that they were given

Table 2.5 View and pseudo-color operators

View and Pseudo-Color		
1.	View- Normalizes the image in the [0,255] pixel range	Inputs: 1. Image to be made viewable (Geotiff) Outputs: 1. Viewable image created by automatic transformation (Geotiff)
2.	ViewInterval- Normalizes the image in the [0,255] pixel range	Inputs: 1. Image to be made viewable (Geotiff) 2. Minimum value of pixel band (Float) 3. Maximum value of pixel band (Float) Outputs: 1. Viewable image created by manual transformation (Geotiff)
3.	Pseudo-Color- Combines three single banded images into a 3 layer colored one	Inputs: 1. Image representing the Blue band (Geotiff) 2. Image representing the Green band (Geotiff) 3. Image representing the Red band (Geotiff) Outputs: 1. Resulting pseudo-Color multiband image (Geotiff)

Table 2.6 Statistical operators

Statistics		
1	Statistics- Returns the statistic information about an image.(histogram)	Inputs: 1. Single band image (Geotiff) Outputs: 1. Statistics file containing a histogram given in pixels and the covered area
2	Information- Returns the general information about an image.	Inputs: 1. Single band image (Geotiff) Outputs: 1. File containing general information about the image

Table 2.7 Classification operators

Classification		
1	Supervised Classification- Generates a classification map based on a multibanded image(representing the feature	Inputs: 1. Multibanded image (Geotiff) 2. Vector image, providing supervised

	vector) and spectral signatures provided through a vector map.	classification control points. (ascii) Outputs: 1.single band image(Geotiff)
2	Unsupervised Classification- Generates a classification map based on a multibanded image (representing the feature vector) and a settings file (representing the parameters of the classification ex: nr. of classes, iterations, convergence, etc.)	Inputs: 1. Multibanded image (Geotiff) 2. Classification Arguments file (Class). Outputs: 1.single band image(Geotiff)

Table 2.8 Class file format

Format of Class File	
[classes={integer} NEWLINE]	DEFAULT VALUE=10
[min_size={integer} NEWLINE]	DEFAULT VALUE=17
[iterations={integer} NEWLINE]	DEFAULT VALUE=30
[convergence={floating point (50.0-100.0)} NEWLINE]	DEFAULT VALUE=98.0
[separation={floating point} NEWLINE]	DEFAULT VALUE=0.0

The operators described above use the maximum likelihood algorithm in order to classify each unit of information (pixel). Further operations to be implemented within the scope of the current group are: Recoding, Density Slicing and Accuracy Assessment.

Table 2.9 Scope operators

Scope		
1	SetProjection- Reprojects and changes the geographic reference of images.	Inputs: 1. Geographically referenced Image (Geotiff) 2. Projection File (PROJ4) 3. Region File
2	GetProjection-	Inputs:

	Extracts the projection of a given image	1.Singleband image (Geotiff) Outputs: 1.Projection file (PROJ4)
3	SetRegion- Clips images to a given region.	Inputs: 1.Image to be either clipped or reprojected (Geotiff) 2.Region to which the image shall be clipped (internal grass region file) Outputs: 1.Clipped image (Geotiff)
4	GetRegion- Extracts the region of a given image	Inputs: 1.Singleband image (Geotiff) Outputs: 1.Region file
5	Mosaic- Mosaics two adjacent singleband images	Inputs: 1.First adjacent image (Geotiff) 2.Second adjacent image (Geotiff) Outputs: 1.Mosaicked singleband image (Geotiff)

All images within scope operations must be geographically referenced. Images not containing any kind of geographic reference cannot be used as input for the scope operators. No operators belonging to the miscellaneous group have been implemented yet. Future operations include coordinate sampling.

2.2 GreenLand Application

The GreenLand application is free Geographic Information System (GIS) software for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization. In particular this application represents a case study for the Istanbul geographical area and for the Rioni river, situated in the western part of Georgia. Because GreenLand operates mainly on satellite images – containing in some cases high confidential data – the Grid infrastructure adds an extra security layer (based on Grid certificates), besides the Web user authentication mechanism.

2.2.1 GreenLand Overview

The majority of the operators are based on mathematical formula. The GreenLand application tries to hide their complex definitions by describing them as nodes within a graph, called workflow. The main objective of the GreenLand application is the generation and execution of workflows, based on satellite images. In this case each workflow could be 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 GreenLand is a Grid based application that offers scalability when dealing with large number of users and large processing data volume. Due to the computing and storage capabilities offered by the Grid infrastructure, the workflows execution times are significantly improved in comparison with standalone/cluster processing. Various inputs are supported in the framework of this application, such as: satellite images (Landsat, MODIS, Aster, etc.), executable files, primitive inputs (numbers, strings, boolean values).

2.2.2 GreenLand Architecture

The GreenLand is a web based application, implementing the client-server standard (Figure 2.3). The client side is developed using the Adobe Flex 4 technology. The server side is Java based, and uses MySQL database for storing useful information. The connection between the client side and the Grid infrastructure is based on rest services, resident on the GreenLand application server.

The only way for a user to access the backend functionality of the GreenLand application is through its graphical user interface. In order to hide the complexity of the Grid mechanisms, the GreenLand application could be easily accessed as a Web application.

The rest services exposed by GreenLand are resident on the application server. When needed, they could be easily transferred to other locations without affecting the application functionality. These services manage the execution and monitoring of the workflows over the Grid infrastructure, but also the analysis and interpretation of results. The development of customized workflows is another issue related to this architectural level.

The gProcess platform interprets user requests in terms of Grid infrastructure commands. The ESIP platform is represented in this architecture as a repository that collects the general information about each workflow. A detailed description of these workflows is stored in the GreenLand repository.

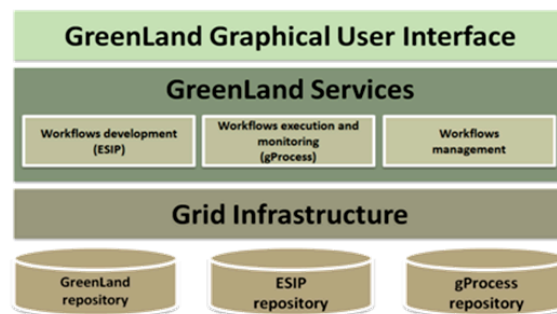


Figure 2.3 GreenLand system related architecture

2.2.3 Use Case Scenarios in GreenLand Application

This section describes the actions flow that a user needs to follow in order to execute different kinds of processes over the Grid infrastructure (Gorgan et al., 2010). Because this application runs over the Internet layer, it allows the users the possibility to access it all over the world, from different physical machines, even from mobile devices. There are three major functionalities offered throughout the GreenLand application:

- Preprocessing phase: allows the user to prepare all the input data that is going to be used in further Grid execution processes. At this stage the user selects all the workflows included in these processes, specifies different attributes, groups workflows into projects, etc.;
- Execution phase: closely related with the Grid based processing. This step is performed in the background and tries to hide the complex functions used to launch such processes;
- Monitoring phase: displays the Grid execution status in a friendly manner by using different user interaction techniques at the application graphical interface level. All the time the user is notified about the execution progress and it has the possibility to interfere in the overall execution process.

This phase also includes the results management, where all the outputs generated by Grid executions are at the user discretion.

The following sections detail all the phases of this use case scenario, by highlighting theoretical and practical solution implemented within the GreenLand application. All the required steps that allow the user a correct usage of the platform are also described in the following paragraphs.

2.2.3.1 Processing Description

Before starting any kind of processing over the Grid infrastructure the user should verify if all the input data are properly specified for all the workflows attached the selected project. By default the GreenLand application set standard values for each type of input:

- The first image, from the available user images list, in case the input type is identified as GeoTIFF;
- For integer or float types the default value is set to 0;
- For text based inputs, the empty string is used.

This processing description phase allows the user to prepare all the required data that is going to be used in Grid processing. The steps that are described in the following paragraphs are organized as phases in the GreenLand usage scenario:

A. User authentication

In order to access the GreenLand functionalities the user should first sign in the application. The username and password are mandatory in this type of process. It is worth to mention that when a new user account is created, also a new project is automatically generated by the application (called Default project) that could be further used in the Grid based process executions.

B. Graphical user interface – general overview

The user interface consists of six main modules that are organized in such way in order to increase the degree of usability for the entire application. A short description of each module is presented below:

- User information: display the basic information of the logged in user;
- Application logos: visual representation of the GreenLand application and the enviroGRID project;
- Menu options: contain the most used set of operations. Other actions will be provided as sub-option of this menu set;
- Projects list: one of the concepts that are regularly used within the GreenLand framework is the project one. A project represents the basic structure of the application that could be processed over the Grid infrastructure. It has a unique name, provides a suggestive description and allows multiple workflows attachment. It has a status that offers feedback about its behavior (e.g. “Executing project” means that the project is in the Grid execution phase and when this phase ends it will automatically set the “Execution completed” status).

Each user has a list of active projects that could be updated with new data or could be deleted together with the content and their output results. Creating a new project is another available option for the logged in user;

- Workflows list: the workflow is another key-concept introduced by the GreenLand application. A workflow is a graphical representation of a mathematical formula encapsulated within operators. In its simplest form, a workflow identifies with a basic operator. Complex workflows contain a group of operators, or sub-workflows linked together by uni-directional communication channels. Each channel describes the order used for executing the workflow and connects the output of an operator/workflow with the input of another operator/workflow.

Cyclic graphs are not allowed within the workflow's internal structure. Multiple nested levels are supported by each workflow (e.g. a workflow could contain another workflow that in turn it could have other inside workflows, and so on).

Each project contains a list of workflows. When a new project is created this list is empty. That is why the default status for the project is set to "Empty project". It is up to the user to specify what kind of workflows it would like to add to the current selected project;

- General information window: details all the actions performed by the user at the graphical interface level as highlighted in Figure 2.4 (e.g. when a project is selected it offers information about it. Or when a workflow is selected this module allows the user to view all the inputs of the workflow and to add them the required values).

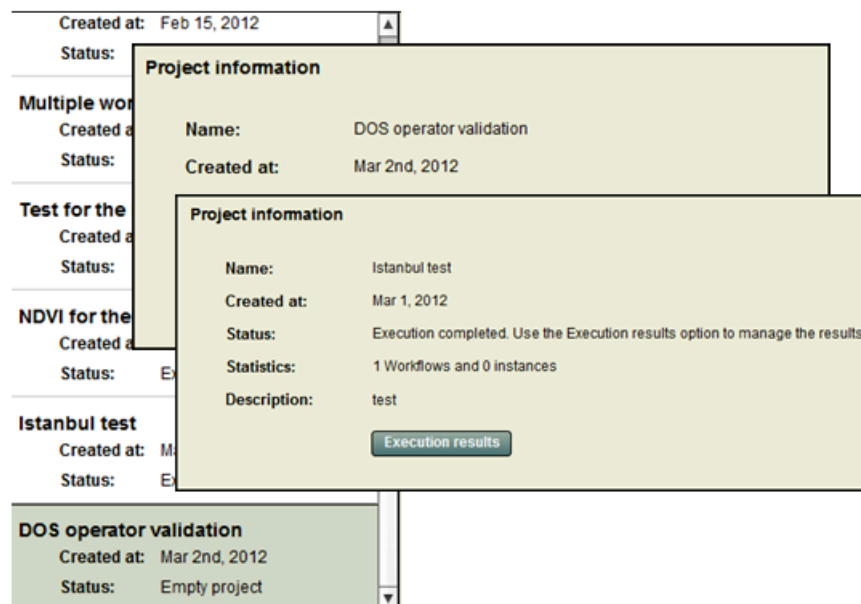


Figure 2.4 User's project details module

C. Project management

The next step of this use case scenario consists in creating a new project. This is done by using the **Create project** menu option. All the user actions performed at the graphical interface level correspond with a specific server side service. In this case the **createProject** service will be

invoked. Such a communication is always initiated by the GreenLand interface that sends a list of parameters that describe the action behavior. The server side script is executed in the background and after it finishes, it generates a response that is send back to the application graphical interface. The response is processed and displayed to the user in different formats: message, execution progress, list of operators, list of available input data, etc.

Each project requires a unique name and a short description in order to be easily identified when the projects list grows in size. After that it is automatically added to the user's projects list where it displays the date when it was created and the "Empty project" status (Figure 2.4). Other possible statuses for a project are listed below:

- Content provided: the project contains a non-empty workflows list;
- Pending execution: the user request for a Grid processing, but the data is not completely sent to the processing nodes;
- Executing project: all the input data and the external dependencies were successfully transferred to the Grid worker nodes and the execution is in progress;
- Execution completed: this status is generated after the project execution finishes and the results are available for analysis and interpretation;
- Incomplete execution: generated by network failure. Once the situation is back to normal the data processing is automatically resumed from the last point of execution;
- Error: generated by sever technical issues regarding the Grid infrastructure. In this case the execution should be start all over again.

Updating a project with a new name or description is another option provided by the GreenLand application. Also for an existing project new workflows could be added or the user is allowed to remove some of the old ones.

All projects could be deleted by the user that created them. The "Pending execution" and "Executing project" projects are exceptions for the deletion rule.

D. Workflow management

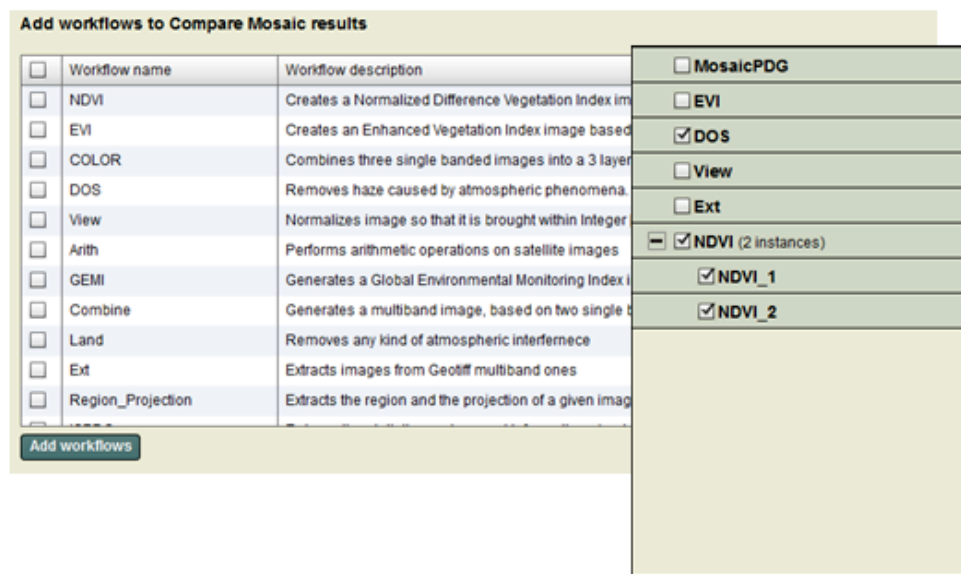
The "Empty project" status is displayed each time the user creates a new project. All the items listed in the **Workflows list** could be used as content for the new generated project. In order to access this list the user could use the **Add workflows** menu bar option. After that it may perform multiple workflows selection from the right hand side list. The list item click action takes the user to a help page that describes with technical details the inputs, outputs and the behavior of this workflow (Figure 2.5).

It is worth to mention that multiple instances of the same workflow could be added to the same project. Because each item added to the project should have a unique name (Figure 2.5), all these instances will be automatically renamed by adding an underline character (" _ ") followed by the index of that instance (e.g. NDVI_1, NDVI_2, NDVI_3, etc.).

Removing workflows from a project is another feature implemented within the GreenLand application. This operation is performed through the **Remove workflow** menu option and it affects all the selected workflows from the list attached to the current project.

Adding and removing workflows to a specific project is not a persistent operation. This means that on the next sign in action the user will find this project to be empty. In order to achieve this purpose, the **Start project** button should be used. By using this option all the relations established at the graphical interface level are also defined in the GreenLand database.

For each workflow that is added to the current selected project the user is able to specify a list of inputs. Update the workflow name and description is also supported by the GreenLand application. More information about adding these inputs are described in the chapter entitled *Grid based process execution*.



<input type="checkbox"/>	Workflow name	Workflow description
<input type="checkbox"/>	NDVI	Creates a Normalized Difference Vegetation Index image
<input type="checkbox"/>	EVI	Creates an Enhanced Vegetation Index image based
<input type="checkbox"/>	COLOR	Combines three single banded images into a 3 layer
<input type="checkbox"/>	DOS	Removes haze caused by atmospheric phenomena.
<input type="checkbox"/>	View	Normalizes image so that it is brought within integer
<input type="checkbox"/>	Arith	Performs arithmetic operations on satellite images
<input type="checkbox"/>	GEMI	Generates a Global Environmental Monitoring Index i
<input type="checkbox"/>	Combine	Generates a multiband image, based on two single b
<input type="checkbox"/>	Land	Removes any kind of atmospheric interference
<input type="checkbox"/>	Ext	Extracts images from Geotiff multiband ones
<input type="checkbox"/>	Region_Projection	Extracts the region and the projection of a given ima

- ☐ MosaicPDG
- ☐ EVI
- ☒ DOS
- ☐ View
- ☐ Ext
- ☒ NDVI (2 instances)
- ☒ NDVI_1
- ☒ NDVI_2

Figure 2.5 Add workflows to project

E. Upload external data

After the project contains all the needed workflows, the next step consists of specifying the inputs for all these workflows. In case an operator requires a satellite image as input the user should select it from its own data, or from public data uploaded by other users. There are three options for uploading data into the GreenLand application, based on OGC services, FTP or http data transfer. It is worth to mention that after all these satellite images are completely transferred a conversion mechanism will try to convert them into multiple single-banded images. If the conversion succeeds then these bands of image could be used as inputs for the Grid executions, otherwise an error message will be displayed.

In order to upload external data the user should access the **DataMng.** menu option. The following features are supported by this option:

- Upload local files**

Allows the user the possibility to upload into the GreenLand data repository the files stored onto its own physical machine. This action is performed based on the http communication protocol. Multiple files upload is also a supported feature in the framework of this application.

In order to attach these files the user should be able to browse the folders and files structure. This operation is done through the **Attach files** button. At any time the user is

allowed to cancel the upload process, by using the **Cancel upload** button. The data transfer progress is represented through a progress bar (Figure 2.6).

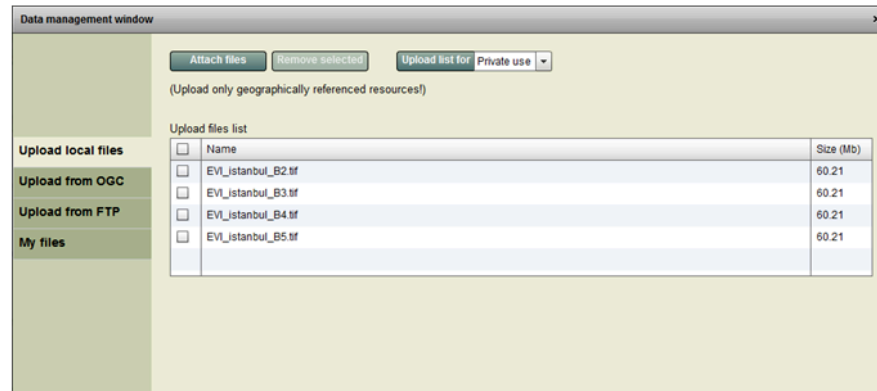


Figure 2.6 Upload local files panel

- **Use OGC services for data retrieval**

Data transfer based on OGC services is another feature implemented within the GreenLand application. This operation requires from the user only the http address of the server that hosts these services. By requesting the GetCapabilities xml file, the application will display a list of layers provided by this service and allows them to be overlapped over an interactive map. This operation is identified as WMS (Web Map Service) and its one of the three most important features provided by the OGC services. The WMS implementation has only an interactive data visualization effect. The actual data transfer is based on the WCS (Web Coverage Service) that is also supported by the GreenLand application. When selecting an item from the layers list, some details about this layer are automatically extracted from the GetCapabilities file: width and height of the geographical area, type of projection, OGC version, etc. Each of this information is editable by the user. Generally speaking, each layer covers a large geographical area. The user, in most cases, is interested in collecting data for a smaller region. That is why the GreenLand application allows the map based geographical area selection that will be used in the data retrieval process (Figure 2.7).

The data that will be generated by the WCS service will have the GeoTIFF format and it will be available as a static image. All the previous mentioned parameters (image width and height, geographical area boundaries, etc.) are used in the request call of retrieving the corresponding data from the WCS service.

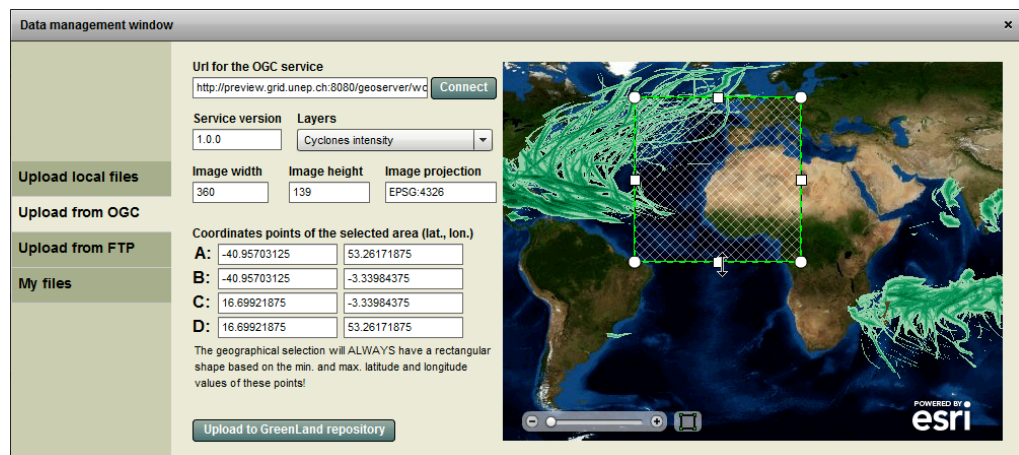


Figure 2.7 OGC service based data retrieval panel

- **Use FTP protocol for data retrieval**

By using the FTP protocol the user is able to specify remote files in the Grid based data processing. This requires the input of a FTP server address, followed by a username and password authentication or by an anonymous authentication (Figure 2.8). At this moment only a single file could be transferred at a time, but the improvement of this feature is already one concern for the next GreenLand version. At any moment the user is able to cancel the data transfer.

Specifying the file access type (private or public) is also supported. If the user has another image with the same name an error message will be displayed and the FTP data transfer will not be performed.



Figure 2.8 FTP based data retrieval panel

No matter what option for data uploading will be performed by the user, the conversion mechanism will be applied on each new image transferred to the GreenLand data repository. Based on the conversion result one of the following statuses will be set up for each uploaded image: *Checking* (in progress operation), *Valid* (operation succeeds), *Error* (the image is not compatible with the GreenLand internal format). Only images that display the *Valid* status could be used as inputs for the workflows execution. The user has the possibility to perform data management operations on all of its uploaded images, by using the **My files** option. This feature allows him to search files by keywords, to change the image privacy or to download a group of images as an archive.

This section (*Processing description*) highlights the main steps that should be followed by a user in order to set up valid data that is going to be used in the Grid based execution processes. Also some technical descriptions were detailed in order to provide a better understanding of the internal mechanisms of the GreenLand application.

2.2.3.2 Grid Based Process Execution

After all the required workflows were properly added to the new created project, the next step consists in specifying valid inputs for all these workflows. As mentioned in the previous sections the user has the possibility to upload its own data. Each such data should be mapped over an input that requires a satellite image. This is done manually, by filtering all the images own by the current logged in user and the public images uploaded by other users. The action of specifying the workflows' inputs should be preceded by the workflow selection (Figure 2.9).

In some cases a workflow requires a text based input (number, string, boolean value) instead of satellite image. In this particular case the data uploaded by the user is left aside, and it should specify a value that is compatible with the type requested by this input (e.g. the EXT workflow needs to know the band index that is going to be extracted from a multiband image). By using the **Save** button placed at the bottom of the inputs list, the workflow will be updated with the user's data. Renaming the workflow before the execution is another feature provided by the GreenLand application.

The Grid execution of the selected project starts when using the **Start project** menu option. Before proceeding to this step the user should double check the correctness of data. The Grid data processing consists of the following actions:

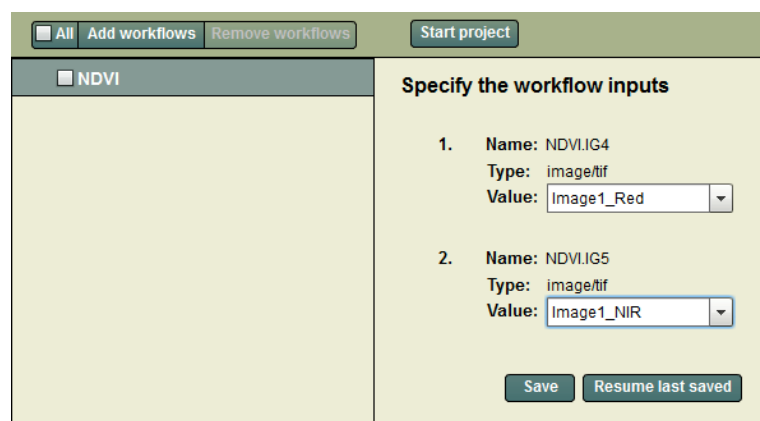


Figure 2.9 Specify the workflow inputs

- Project validation: projects with no attached workflows could not be processed. The GreenLand application offers the possibility to update an existing project with new workflows and to execute



- it with this new information. In such cases the previous results will be overridden by the new ones, generated by this second Grid execution. Before proceeding to this step a warning message will be displayed, asking the user if it agrees with the terms of this processing;
- Verify the Grid infrastructure availability: before sending data to be processed over the Grid infrastructure, an internal GreenLand service verifies if the Grid is accessible and if it could perform the actions required by the user;
 - Update the GreenLand database: only now the actions described by the user at the graphical interface level are defined in the GreenLand database. The content of the project is trans-coded into an internal ESIP format, called iPDG (instance Processing Description Graph) that stores all the relations defined between the workflows. In this step each workflow is divided into its' atomic components, in order to accomplish a better execution management. An atomic component is the smallest executable unit of a workflow. More information about this mechanism is described in the section entitled *Process description graphs*;
 - Set up the Grid resources: depending on the project complexity a specific number of worker nodes are selected to process the user's request;
 - Data transfer to the Grid nodes: data required by each atomic component of the workflows will be transferred to a specific Grid worker node;
 - Grid based process execution: this step performs the data processing. It is important for the ESIP platform to be aware of the relations established between the atomic components of the workflows. An atomic component will not start the execution if it requires as input a result that is generated from another job that is in the execution stage. The graph based representation of the GreenLand project (stored in the iPDG file) allows the ESIP platform to know the status of the execution and the atomic components that should start in the next processing phase;
 - Generate the final output result: a workflow could have multiple outputs. These final outputs are generated based on all the intermediate results of the atomic components of the workflows. When such a final result is available it is displayed in the graphical user interface of the GreenLand application. The project execution completes when all the inner processes are completed.

There are five statuses involved when executing data over the Grid infrastructure. Each status is suggestive for the progress of execution:

- Submitted: displayed when allocating Grid resources and sending the project inputs and the external dependencies (additional libraries in particular) to the Grid worker nodes;
- Running: stated throughout the entire execution of the GreenLand project;
- Completed: this means that the Grid based process execution finished and the results are available for further analysis;
- Error: occurs when technical errors interfere with the project execution. These errors consist in: network failure, worker nodes inaccessibility, sending wrong data to the Grid nodes, data that do not have a physical correspondence (e.g. files that do not exist), incomplete workflows, etc. In such cases the process should be started all over again;
- Canceled: caused by the user request in cancelling the Grid based data processing.

2.2.3.3 Process Execution Monitoring

The Grid data execution monitoring phase offers the user the feedback about the status of the GreenLand project execution. This is based on a general timer, defined at the application level that periodically updates this status. All this information is displayed to the user in a table similar to the one presented in Figure 2.10. The first column displays the workflows name attached to the project, while the status column represents the state of the execution. When a workflow finishes its execution and the output results became available, in the **Result** column a list box and a download button allows the user to select the needed result that is going to be downloaded to its local machine. At this stage of development the user is able only to download the results and to analyze them offline. The GreenLand future work also includes online visualization and interpretation of results.

Project execution details

Workflow name	Start time	Status	Result
NDVI	2011-10-13 10:43:50.0	Running	
EVI	2011-10-13 10:43:50.0	Completed	5a6ac4f8-ba84-4a43-4... Download

Figure 2.10 Grid based execution monitoring panel

2.3 ESIP and GreenLand System

ESIP and GreenLand systems represent two sides of the same application working in synergy to offer satellite image processing, visualization and management functionalities. Services provided by ESIP have been created purposely for the GreenLand application and all the backend functionality of GreenLand is implemented on the ESIP platform. These two platforms may be understood within the framework of the classic client-server paradigm. ESIP represents the server side, while GreenLand represents the client side of the pair.

GreenLand gives the user access to the following functionality: project management and execution and workflow and data management. It also requires authentication in order to give any user access to its interface. All of these functionalities are implemented by the ESIP server side application, which communicates with the GreenLand interface by means of HTTP messages, forming a kind of restful services.

2.3.1 Generic Case Studies

The GreenLand application supports the definition of workflows of predefined operators. Future developments of the application will allow the user to describe its own workflows and operators, based on a workflow editor tools. The operators will be described by Grass scripts.

To be mentioned is the fact that currently only the entity developing these operators can add new operations to the already existing set. Starting from a given algorithm such as NDVI which has the following formula, one can start to implement the desired functionality.

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

To solve this equation one needs to create the basic operators for the addition, subtraction and division (+, -, /) mathematical operations and only after that it can bring them together in complex workflows to provide the needed functionality. The procedure employed to add new functionalities to the GreenLand database is the following:

A. Create the Grass Script

This describes the general behavior of the workflow that is going to be implemented. As a further note to this step is that the user should define the most basic operations, which may be later composed to form a complex workflow (Figure 2.11);

```
1 #!/bin/bash
2 #Grass script file ADDITION
3
4 #Grass Import Section
5 r.in.gdal input=input1.tif location=grass_location output=input1
6 g.gisenv set=LOCATION=grass_location
7 r.in.gdal input=input2.tif output=input2
8
9 #Grass Functionality Section
10 r.mapcalc output=input1+input2
11
12 #Grass Export Section
13 r.out.gdal input=output output=output1.tif
```

Figure 2.11 Grass script for workflow definition

- B. Integrate the script into the GrassOperators application, which will be used by the gProcess platform;
- C. Add the arguments which will be passed to the GrassOperators application within the gProcess database;
- D. Add the operator mapping from gProcess in the ESIP database;

Create a workflow within the ESIP database to highlight the functionality added to the application. After creating the workflow at the GreenLand graphical interface level the ESIP backend services convert the visual representation to a textual one which the workflow executor (represented by the ESIP and gProcess platforms) can understand. The textual representation is under the form of a PDG, described in XML format (Figure 2.12).



Figure 2.12 PDG and iPDG definitions for the NDVI workflow

3 Hot Spot Inventories

3.1 İstanbul Case Study

3.1.1 Introduction

3.1.1.1 Purpose and Scope

This chapter reports on application of remote sensing image processing steps to derive land cover/land use categories changes especially vegetated areas in İstanbul, Turkey by using commercial remote sensing software and enviroGRIDS based developed ESIP tools (Gorgan et al., 2009). Remote sensing technologies have been utilized to multidisciplinary applications by many scientists. Remotely sensed data is the major source of spatial information of the earth's surface cover and constitution (Curran, 1994; Schmidt and Skidmore, 2003). The technology accommodates accurate and reliable information as base input of many research subjects. Different vegetation indices have been developed and used to determine land cover categories and occurred changes by using Landsat 5 TM data. In this chapter after short description of İstanbul, case study processing steps and requirements are defined. Results of the application are then reported, with recommendations for subsequent developments.

3.1.2 Study Area

The population of Turkey was 13 648 270 in 1927 and it increases approximately five fold in seventy three years and reached to 67 803 927 in 2000 and 74 724 269 in 2011. According to Turkish Statistical Institute, the population of Turkey will reach to 85 407 000 by 2025. İstanbul, with the population of over 12 million, is one of the most important cities in the world due to its historical, cultural, industrial and natural characteristics. İstanbul has been affected because of huge immigration, the city population was 3 million in 1970s, it became 7.4 million in 1990s and current population is around 13 million (Turkish Statistical Institute, 2012). After 1980, İstanbul has experienced rapid land use/cover changes connected to population growth, industrialization and urbanization.

İstanbul is located in north-western Turkey within the Marmara Region on a total area of 5,343 square kilometres (Figure 3.1). There are several reasons why İstanbul is considered as the test case study for deriving land cover categories phenomena. Humans are increasingly disturbing natural resources, ecosystems and the environments. As a result, the city is facing serious water quality problems, deforestation, desertification, soil erosion, degradation of land productivity, and the disappearance of biodiversity and sensitive regions. There is an emergent need to assess the environmental impact of the rapid urban expansion on natural resources such as forest and green area.

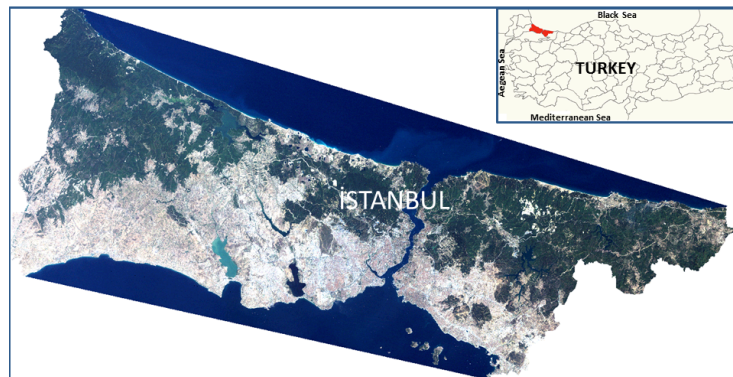


Figure 3.1 Study Area (İstanbul)

3.1.3 Data and Methodology

Landsat Program has been extensively used for ecosystem monitoring (Cohen and Goward, 2004; Pizarro et al., 2012) and the almost 45 year long record of Landsat imagery provides a rich data set to map vegetation composition. Landsat images are available freely and that makes the data feasible to analyse both larger areas and denser time series. Landsat TM and ETM images with their 30 m spatial resolution do not have potential to map vegetation in species level. However, Landsat images have been successfully used for vegetation mapping. To map coverage of vegetation and to determine the changes in vegetated areas four LANDSAT 5 TM images (Row/Path: 180/31; 180/32) dated at 25th of September 1987 and 5th September of 2009 were used. Remotely sensed data were obtained from NASA The Warehouse Inventory Search Tool (WIST) (<https://wist-ops.echo.nasa.gov/api/>). These images have 30 m spatial and 8 bit radiometric resolution with 6 spectral bands. In addition, 1:25,000 scale topographic maps used for image processing steps such as geometric correction and accuracy assessment. Field survey was conducted to gather ground truth data.

The image processing techniques conducted in this research to obtain past and present accurate vegetated area of the study site were explained in the following sections. The procedure applied in this study includes application of image pre-processing, calculation of remote sensing indices, density slicing classification and assessment of accuracy by using commercial image processing software (ENVI) and ESIP Platform (Figure 3.2).

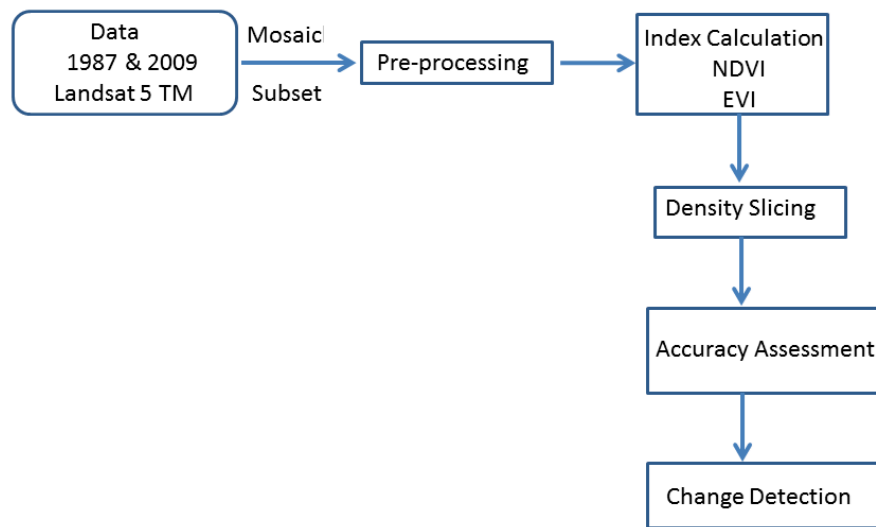


Figure 3.2 Flowchart of İstanbul Case Study

3.1.3.1 Radiometric Correction (DN to Reflectance Conversion)

The electromagnetic radiation signals travelling through the atmosphere are altered because of the process of absorption and scattering caused by gases, and aerosols within atmosphere. The radiation reaching the top of the atmosphere (TOA) from the Sun (Extraterrestrial irradiance), is altered as it travels through the atmosphere to the surface of the Earth; and the radiation reflected or emitted from the surface of the Earth is not equal to the radiation measured by sensors in orbit around Earth. As a result, remotely sensed data are not only dependent on the spectral properties of targets on the surface of the Earth, but also on the content of the atmosphere. Changes in scene illumination, atmospheric conditions, viewing geometry and instrument response cause radiometric distortions over satellite image. Therefore, satellite sensor images are radiometrically and atmospherically corrected to eliminate system errors and to minimize

contamination effects of atmospheric particles through absorption and scattering of the radiation from the earth surface (Song et al, 2001; Liang, 2004). Atmospheric contamination is a major source of error in several vegetation monitoring applications (Liang, 2001; Townshend et al., 1992) hence, it is common practice to remove it in cases when quantitative measurements are needed, or data collected over different dates and/or areas need to be used together. The objective of radiometric correction is to recover the “true” radiance and/or reflectance of the target of interest (Lathrop, 1988).

Conversion from Digital Number (DN) to radiance (analogue signal) was conducted by using calibration parameters such as gain and offset. These are available in published sources and image header files.

The following equation is used for the calculation of radiance values from DN values:

$$L_{\lambda} = C_0 + C_1 * DN$$

L ($W\ m^{-2}\ sr^{-1}\ \mu m^{-1}$) = Top of atmosphere (TOA) upwelling radiance

C_0 and C_1 ($mW\ cm^{-2}\ sr^{-1}\ \mu m^{-1}$) = Offset and Gain values

DN = Digital numbers

L was converted to TOA reflectance, R^* (unitless) using the following equation

$$R = (\pi * L_{\lambda} * d^2) / (ESUN_{\lambda} * Z)$$

Where

Z = Solar zenith angle in degrees

$ESUN_{\lambda}$ ($W\ m^{-2}\ \mu m$) = Mean solar Exoatmospheric irradiances

R = Unitless planetary reflectance

d = Earth Sun distance

L_{λ} = Spectral radiance at the sensor apertures

The Earth-Sun distance has been calculated from the given formula.

$$d = 1 - 0.01674 * \cos(0.9856 * (Jday - 4))$$

Jday = Julian day for day of satellite pass

3.1.3.2 Atmospheric Correction (Dark Object Subtraction)

Atmospheric distortion affects remotely-sensed imagery, contributing (erroneously) to pixel values. Atmospheric correction is necessary where reflectance values are desired (as opposed to simple DNs) and where images are being compared over time. There are many methods for atmospheric correction, none of which is perfect and some of which are very complex. Relatively simple and common methods include; dark object subtraction and histogram matching.

The most common atmospheric effect on remotely-sensed imagery is an increase in DN values due to haze, etc. This increase represents error and should be removed. Dark object subtraction simply involves subtracting the minimum DN value in the image from all pixel values. This approach assumes that the

minimum value (i.e. the darkest object in the image) should be zero. The darkest object is typically water or shadow.

A “dark object subtraction” method was used to correct for atmospheric scattering in the path. Dark object subtraction is an image-based approach that assumes dark objects exist within an image and these objects should have values very close to zero (such as water bodies), and that radiance values greater than zero over these areas can be attributed to atmospheric scattering and thereby subtracted from all pixel values in an image (Moran et al, 1992). In fact, very few targets on the Earth’s surface are absolute black, so an assumed one-percent minimum reflectance better than zero percent (Chavez, 1996). The pixel values are selected for each individual band with the histogram method and subtracted from all pixel values for the corresponding band across an image to remove haze from the image. To employ this method, brightness values were examined in an area of shadow or for a very dark object (such as a large clear lake) and the minimum value was determined. The minimum DN value was selected as the darkest DN with at least a thousand pixels for the entire image. The correction is applied by subtracting the minimum observed value, determined for each specific band, from all pixel values in each respective band. Since scattering is wavelength dependent the minimum values will vary from band to band, therefore each band must be evaluated independently.

3.1.3.3 Vegetation Indices (NDVI and EVI)

A vegetation index is a number that is generated by some combination of remote sensing image bands and may have some relationship to the amount of vegetation in a given image pixel. Description of vegetation indices tested in this study such as Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) is shown in table 3.1.

Table 3.1 Remote sensing indices formulas were used in the study to develop NIR, RED, BLUE denotes reflectance values derived from Landsat 5 TM bands of B4, B3, and B1, respectively and C 1 =6, C 2 =7.5 and L=1 for EVI index.

Index	Abbv.	Reference	Formula
Normalized Difference Vegetation Index	NDVI	Rouse et al. 1973	$(\text{NIR}-\text{RED})/(\text{NIR}+\text{RED})$
Enhanced Vegetation Index	EVI	Huete et al. 1999	$2.5 * (\text{NIR}-\text{RED}) / (\text{NIR} + \text{C1} * \text{RED} - \text{C2} * \text{BLUE} + \text{L})$

The Normalized Difference Vegetation Index (NDVI) is one of the oldest, most well-known, and most frequently used Vis (Vegetation Indices). The combination of its normalized difference formulation and use of the highest absorption and reflectance regions of chlorophyll make it robust over a wide range of conditions. It can, however, saturate in dense vegetation conditions when leaf area index (LAI) becomes high. The value of this index ranges from -1 to 1. The common range for green vegetation is 0.2 to 0.8 (Rouse et al., 1973)

The enhanced vegetation index (EVI) was developed as an alternative vegetation index to address some of the limitations of the NDVI. The EVI was specifically developed to be more sensitive to changes in areas having high biomass (a serious shortcoming of NDVI), reduce the influence of atmospheric conditions on vegetation index values, and correct for canopy background signals. EVI tends to be more sensitive to plant canopy differences like leaf area index (LAI), canopy structure, and plant phenology and stress than does

NDVI which generally responds just to the amount of chlorophyll present. The value of this index ranges from -1 to 1. The common range EVI value for green vegetation is 0.2 to 0.8.

3.1.3.4 Classification (Density Slicing)

Digital image classification uses the spectral information represented by the digital numbers in one or more spectral bands, and attempts to classify each individual pixel based on this spectral information. This type of classification is termed spectral pattern recognition. In either case, the objective is to assign all pixels in the image to particular classes or themes (e.g. water, coniferous forest, deciduous forest, corn, wheat, etc.) (Figure 3.3). The resulting classified image is comprised of a mosaic of pixels, each of which belong to a particular theme, and is essentially a thematic "map" of the original image (Roy and Ravan, 1994).

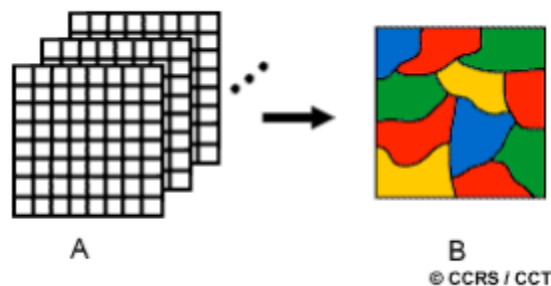


Figure 3.3 Image classification

A method of digital data interpretation in which information from an individual band is enhanced is known as density slicing. Density slicing (also known as double threshold) is a classification technique using computer processing of digital data. The range of brightness of an image is divided into sub-intervals. The enhancement procedure may be applied to one band or a multi-band image. The method works best if the range of brightness values covers a single band of frequencies. Each interval is then assigned a color. The intervals may be defined based on the application.

The original imagery is typically black and white, whereas the imagery after the density slicing is in color. Density slicing allows the user to define sub-intervals for characterizing the data. The advantage of density slicing is that it allows one to gain a greater degree of variability of brightness within the remotely sensed image compared to the original image (e.g. black and white imagery).

The range of input pixel values is assigned a single output pixel value in a density sliced image. The range of pixel values may be defined by the user. If each interval range corresponds to a different color, the result is called a pseudo-color image. Often, density slicing is most effective when the value of particular pixels have significance to a physical variable. Sometimes it is most effective for the user if a particular range of values of intensity may be isolated within the image. In some applications this technique is quite effective. In satellite remote sensing, density slicing is an important technique for classification of remotely sensed images. Density slicing has been used very often to classify the vegetation index.

3.1.3.5 Classification Accuracy Assessment

Classification accuracy has been a focus of attention for a considerable period of time and is a topic that has developed considerably in recent years (Foody, 2002). Classification accuracy is the main measure of the quality of thematic maps produced and required by users, typically to help evaluate the fitness of a map for a particular purpose (Foody, 2008). Although seemingly a simple concept, classification accuracy is a very difficult variable to assess and is associated with many problems.

Ground truth and classified classes are compared to assess classification accuracy. Confusion matrix is constructed for this comparison. This matrix is a row by column table, with as many rows as columns. Each

row of the table is reserved for one of the information classes used by the classification algorithm. Each column displays the corresponding ground truth classes in an identical order. The diagonal elements of the confusion matrix show the number of pixels classified correctly in each class (Foody, 2002).

In this research, to assess the accuracy of classification, the confusion matrix and some common measures derived from this matrix namely, overall accuracy, user's accuracy, producer's accuracy and kappa coefficient are used. The confusion matrix is a simple cross-tabulation of the mapped class label against that observed in the ground or reference data for a sample of cases at specified locations.

The design of confusion matrix is presented in Figure 3.4. The bold elements represent the main diagonal of the matrix that contains the cases where the class labels depicted in the image classification and ground data set agree, whereas the off diagonal elements contain those cases where there is a disagreement in the labels (Figure 3.4). In the following figure 1, 2, and 3 represent the classes of water, forest, and urban, respectively. The number of classes, q , is 3 in this example. Below equations used for overall accuracy assessment and Kappa Statistic calculations.

		Actual Class			
		1	2	3	Σ
Predicted Class	1	n_{11}	n_{12}	n_{13}	n_{1+}
	2	n_{21}	n_{22}	n_{23}	n_{2+}
	3	n_{31}	n_{32}	n_{33}	n_{3+}
	Σ	n_{+1}	n_{+2}	n_{+3}	n

Figure 3.4 The confusion matrix and some common measures of classification accuracy that may be derived from it. Adopted from Foody, (2002).

$$\text{Percentage correct} = \frac{\sum_{k=1}^q n_{kk}}{n} * 100$$

$$\text{User's accuracy} = \frac{n_{ii}}{n_{i+}}$$

$$\text{Producer's accuracy} = \frac{n_{ii}}{n_{+i}}$$

$$K = \frac{n \sum_{i=1}^q n_{kk} - \sum_{i=1}^q n_{k+} n_{k+i}}{n^2 - \sum_{i=1}^q n_{k+} n_{+k}}$$

Where

q = number of classes in error matrix

n = total number of observations in error matrix

n_{ii} = major diagonal element for class i

n_{i+} = total number of observations in row for class i (right margin)

n_{+i} = total number of observations in column for class i (bottom margin)

3.1.4 Results (ENVI)

Remotely sensed data calibrated, mosaicked and subset by using image pre-processing (figure 3.5 a and 3.5 b).

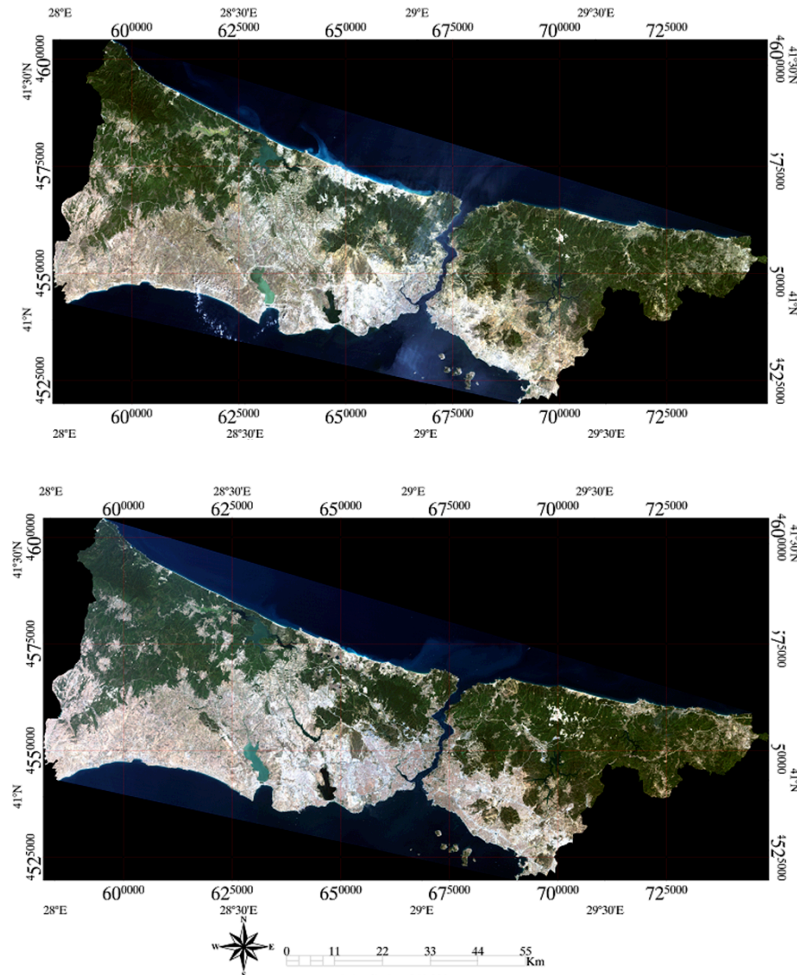


Figure 3.5 Pre-processed data (mosaicked and subset) a) 1987 b) 2009

Each vegetation indices (NDVI, and EVI) were applied on both 1987 and 2009 images. The results are shown in figure 3.6 a, b, c, d, respectively.

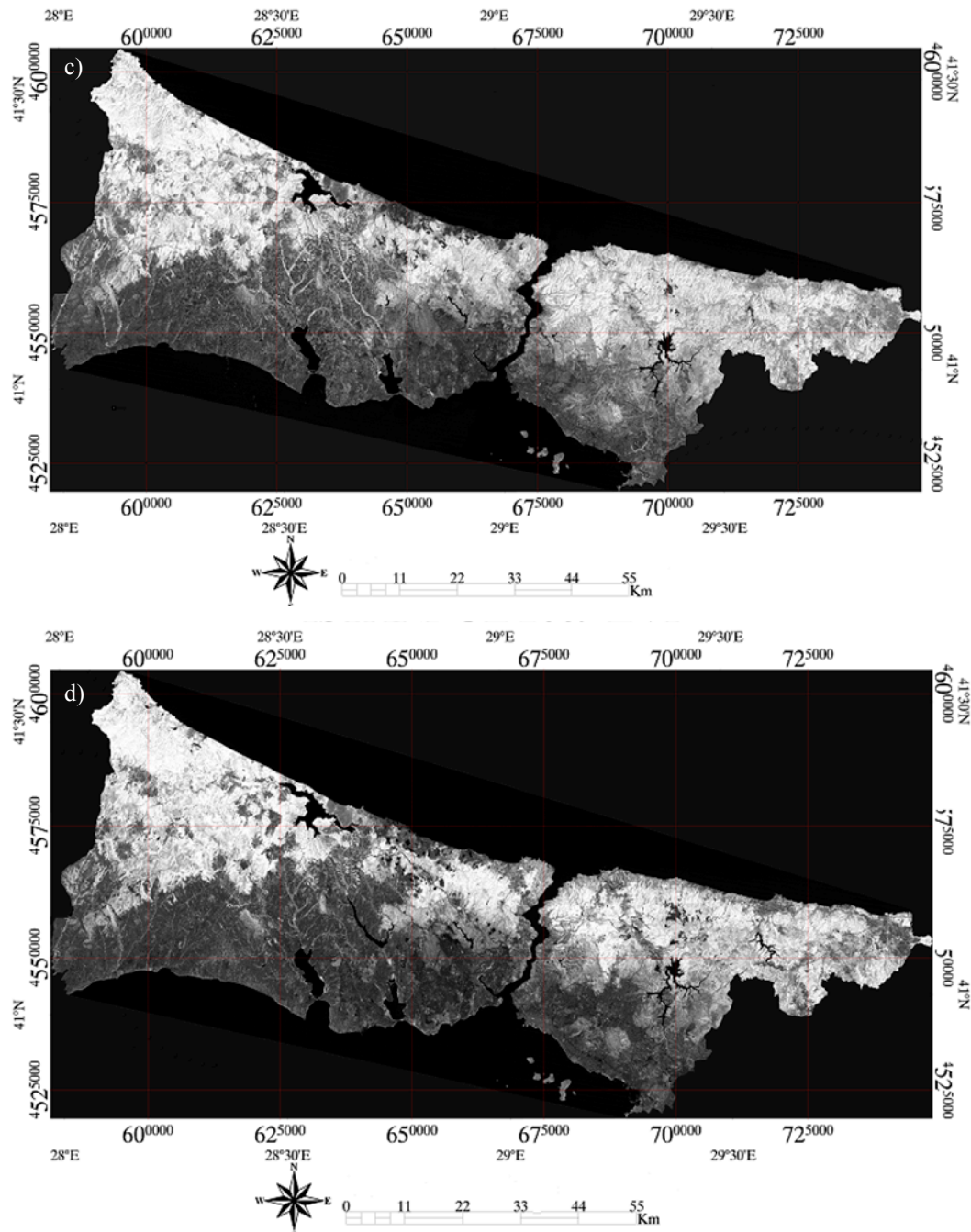


Figure 3.6 Results of vegetation indices (NDVI, and EVI) a) 1987 NDVI b) 2009 NDVI c) 1987 EVI, and d) 2009 EVI

After vegetation indices were applied these images were classified by density slicing method (figure 3.7).

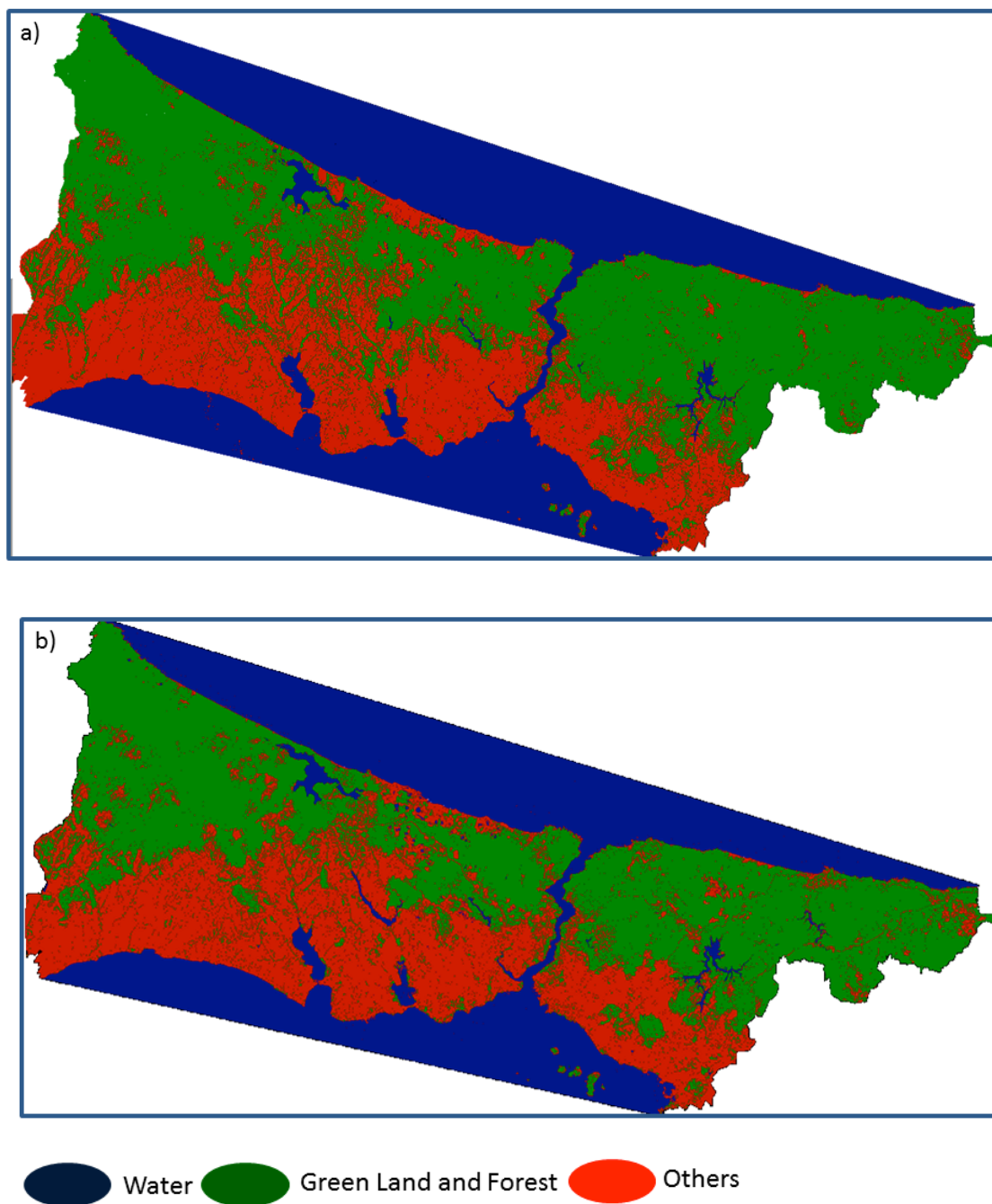


Figure 3.7 Density slicing classification a) 1987 NDVI b) 2009 NDVI

Density slicing results of the study indicate that the category of others that includes urban, artificial, agricultural and barren lands have been increased 77476 ha and green land and forest areas have been

decreased 79562 ha. Environmental changes have been occurred such as land degradation and degeneration of ground water quality.

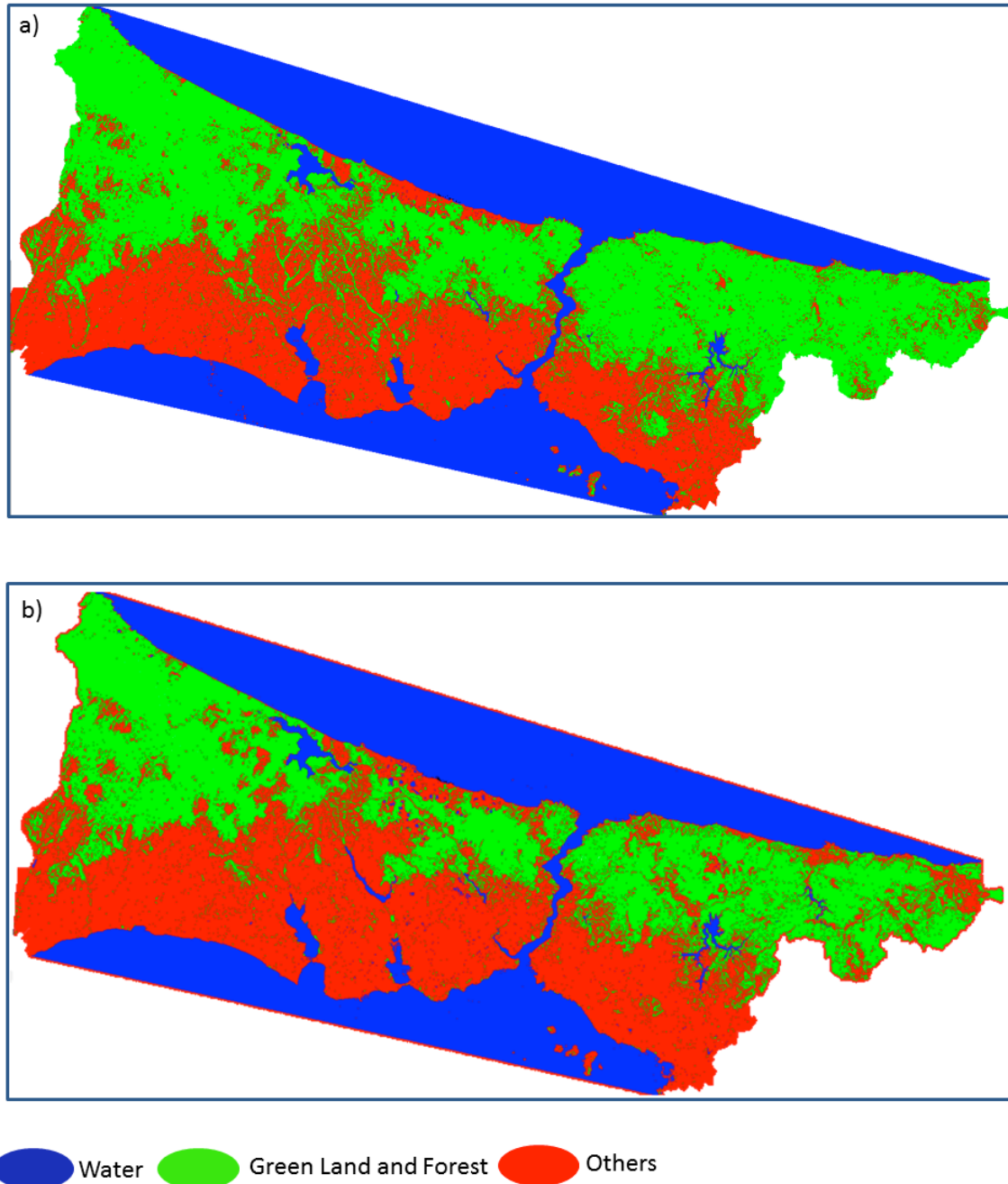


Figure 3.8 Density slicing classification a) 1987 EVI b) 2009 EVI

Overall accuracy and Kappa statistics calculated more higher 0.75 using error matrix. Two vegetation indices have different capabilities on monitoring the vegetation areas according to the situation of the areas and vegetation. Rise in urban population and rapid industrialization is one of the major causes of decreasing

vegetation areas. It is important to ascertain the alteration of vegetation area because it is directly related to human health. In this study it is understood that in heterogeneous urban area like Istanbul Enhanced Vegetation Index (EVI) performs better than others on monitoring the vegetation areas.

3.1.5 Results (ESIP Platform)

3.1.5.1 Grid Based Process Execution

A detailed description of the ESIP platform and the GreenLand application was already highlighted in the previous paragraphs. This section describes the general steps that a user should follow in order to define, execute, monitor and visualize the results for the EVI workflow. We choose this vegetation index because it is intensively used for the Istanbul case study. This operator works with Blue, Red and NIR bands of the same satellite image and it is represented by the following mathematical formula:

$$EVI = \frac{2.5 \cdot (NIR - RED)}{NIR + 6 \cdot RED - 7.5 \cdot BLUE + 1}$$

The methodology of creating such operators was already described. So, we take into consideration only the actions that are going to be performed at the GreenLand graphical interface level. The first step consists in creating a new GreenLand project. The EVI workflow could be selected from the available workflows list. Multiple instances of the same operator are allowed within the same project.

The next step requires the user to input the needed data. In this case all the inputs of the EVI vegetation index are identified as satellite images, in GeoTIFF format. The action of specifying these inputs is done manually, by browsing all the data own by the current logged in user and the public data uploaded by other users (Figure 3.9 Selecting the EVI input data set

After all the inputs are properly specified the next step is the project execution over the Grid infrastructure. This may take a while, depending on the number of EVI instances contained in the selected project and on the size of the input satellite images.

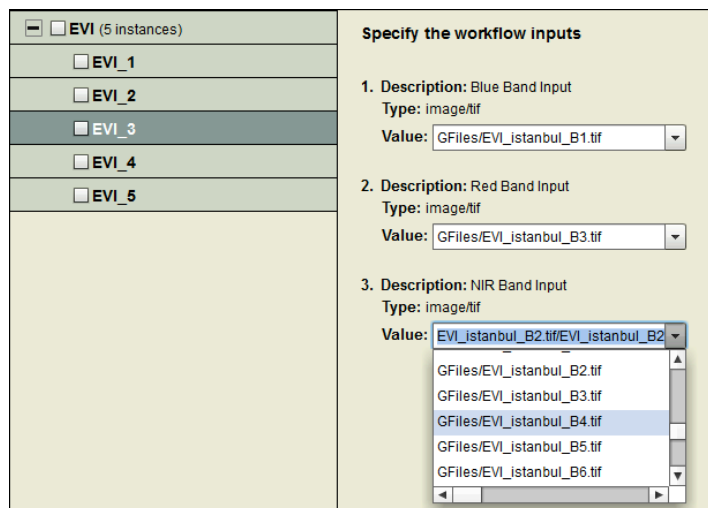


Figure 3.9 Selecting the EVI input data set

The monitoring phase of the Grid execution is also supported by the GreenLand application, and it offers the user the feedback about the execution status. A general timer periodically updates this status, starting from “Running” all the way to “Completed”. When the final output of this workflow is available, the GreenLand project will display the “Completed project” status. At this point, the user is able to download

the result and to further analyze it offline. One of the features that are going to be implemented within the GreenLand application is to give the user the possibility to interpret these results online, through specialized Web based applications.

A practical example of the EVI workflow is described in Figure 3.10 EVI workflow usage result. We use as inputs the Blue, Red and NIR bands of the same satellite image that represents the Istanbul geographical area. The result is highlighted in black and white, but a pseudo-colored representation is also provided.

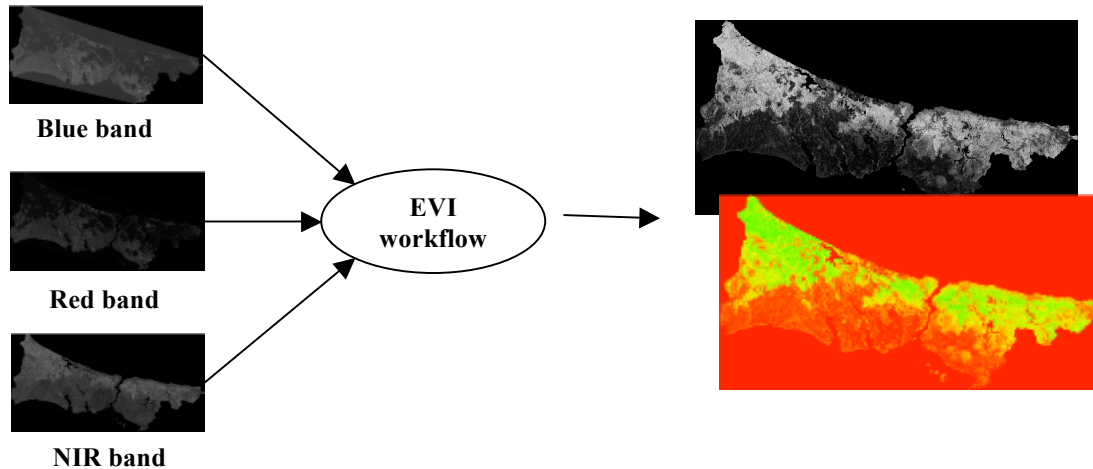


Figure 3.10 EVI workflow usage result

3.2 Rioni River Basin Case Study

3.2.1 Introduction

This section reports on application of enviroGRIDS tools, such as SWAT (Arnold et al., 1998) and ESIP (Gorgan et al., 2009), to Rioni case study in Georgia, contributing into enviroGRIDS deliverable D2.11 (Remote Sensing Data Services). After short description of Rioni River Basin (further details moved into Appendix 1), case study requirements are defined, and potential specifications for ESIP outlined, organized around the themes such as potential user profiles, data sources, functionalities, use scenarios, applications (see Allenbach and Gvilava, 2010). Results of the application are then reported, with recommendations for subsequent developments.

3.2.1.1 Rioni River Basin and Case Study Description

Brief description of the case study site selected for application of enviroGRIDS tools such as SWAT and ESIP follows enviroGRIDS D2.4, subsection 6.3. Further descriptions are provided in Appendix 1.

The Rioni River (see maps on Figures 3.12 and 3.13) is the largest one in Western Georgia with the length of approximately 330 km and catchment area of some 13,400 Km². Tributaries of Rioni mainly flow from mountainous areas: the right tributaries Lukhumi, Tskhenistskali and Tekhuri drain from the southern slopes of the Great Caucasus Range, while left tributaries Kvirila, Dzirula and Khanistskali have their sources at Lesser Caucasus Range. Glaciers contribute as well, with over 2 million cub. m of ice estimated to be accumulated in Rioni watershed.

Downstream of Kutaisi, the largest town in the river basin with 250 thousand inhabitants (located some 170 km from the source), the river flows on Kolkheti lowland and enters the Black Sea near port city of Poti with a population of 50 thousand. The population of the entire basin is estimated to exceed 500 thousand.

River flow regimes are controlled by series of reservoirs located both upstream and downstream of Kutaisi (see Appendix 1 for existing and planned hydropower reservoirs). In the lower reaches, along the Kolkheti lowland, the river is mostly diked with 3-5 m height soil berms. Major flooding events were recorded in every 40-50 years, while relatively small floods are usual phenomena and happen every 4-5 years. Deforestation and inappropriate agricultural practices within the catchment are also the source of concern, leading to slope instabilities, soil erosion and excessive discharges of sediments and pollutants into the Black Sea.

There are several reasons why catchment of the Rioni River is considered as the test case study for modelling the specific hydrological phenomena taking place in this watershed, but also for resembling essential issues of relevance for the Black Sea Catchment (BSC):

- Despite of relatively small size, the Rioni Basin is spreading from highest mountain areas mostly fed with glaciers, continues its precipitation fed flows on the lowland and joins the Sea, essentially mimicking the hydrological processes of much larger catchments and rivers elsewhere in the Black Sea Catchment. The river and catchment processes strongly change character from mountainous areas into the coastal lowland. It is worth mentioning that the edge of the Rioni catchment touches the eastern boundary of the Black Sea Catchment.
- The rivers in the Western Caucasus are characterized by high sediment loads which merit specific investigation in terms of the impact on coastal zones and the Sea, by means of coupling remote sensing data on coastal changes with catchment data on land use, soil, vegetation, hydrology (including glacier mass balance and sediment transport).
- Series of hydropower projects are planned both on the Rioni and in its basin which would drastically alter hydrology of the entire catchment, making the need in the development of river basin management policies and tools ever more required.
- Due to relatively small area compared to the Black Sea Catchment the application of remote sensing instruments and testing the sensitivity of SWAT modelling with regard to data from various sensors could be performed before extending the coverage.
- Deforestation and other types of land cover/use changes are apparent in the watershed and these processes could potentially be detected with remote sensing and coupled with sediment and nutrient loads through SWAT modelling.
- Historic pollution, hydrological and climate data of variable quality is mostly available, as well as the results of hydrological modelling with different tools (such as Sobek), while precipitation data collection improved recently.

Objective for case studies is to use and test enviroGRIDS BSC-OS Portal (see enviroGRIDS D6.1), by using in particular the ESI platform for remote sensing analysis, jointly with other tools. Similarly tools developed could be applied to neighbouring Guria region, which will be particularly important to make synergies between enviroGRIDS and FP7 PEGASO projects (<http://pegasoproject.eu>). This coastal area has been chosen as an example to apply integrated coastal zone management in PEGASO. Collaboration will be particularly useful to link the watershed with the coastal area (see Lehmann et al., 2009).

3.2.2 Application Methodology

This section below specifies for the case study application the potential users (their profile and needs), describes applications and their functionality, and potential use scenarios, while potential data sources (input data, types of available satellites images, etc.) are identified in the next section.



3.2.2.1 Users Profile

The following groups could be defined as the potential developers as well as the users of the system as its functionality proceeds (taking into account the stage of development, as well as grid security concerns):

- EnviroGRIDS Partners involved in remote sensing tasks, case studies and in the development of the ESIP functionality and BSC-OS portal architecture.
- All other enviroGRIDS Partners can become potential users once certain basic functionalities and system prototype is in place.
- Provided the system is substantially developed with certain level of quality assurance, professionals from external institutions could also be invited to test the applications (such as technical staff of the respective agencies – for instance people from relevant departments of the National Environmental Agency in case of Georgia), particularly institutions with river basin management related functions.
- Wider professional communities can be invited through ICPDR and BSC PS as found appropriate.
- Potential users could include specialists involved in environmental impact assessment studies and in preparations of integrated spatial plans, as well as respective agencies reviewing and approving these documents and plans (these users could utilize remote sensing resources to establish baseline environmental conditions and to extract historical data in harmonized and easily accessible manner).
- Funding institutions involved in watershed management projects and respective project teams could utilize results as well (see UNDP, 2011 and USAID, 2011).

Above specified profiles is compatible with the generic user architecture defined for the enviroGRIDS Portal (taken from enviroGRIDS D2.2, subsection 6.5):

- Administrator – manages user accounts, creates new users, updates information on the Web Portal, manages data resources, sets up virtual organization, administers the virtual organization membership, etc.
- Data provider – enters datasets in repositories, fills in metadata, visualizes selected data, manages stored data, etc.
- Earth Science specialist – describes and executes SWAT scenarios. Develops use cases (datasets, scenarios, data visualization, results visualization) for particular studies.
- Decision maker – visualizes and analyzes the model execution results for various scenarios. Visualizes statistics on results provided by different use cases.
- Citizen – visualizes data and reports, etc.

3.2.2.2 Applications

As described in the next section 3.2.3, data from multiple global resources are available to complement nationally available datasets (which sometimes are of questionable quality and detail), but unfortunately tedious manual processes are required to acquire raw satellite data resources, process and utilise them. In addition, downloaded and reprocessed data from multiple sources needs to be converted into format useful for modelling applications, such as SWAT.

A range of programming tools, procedures and user functions would be required to handle pre- and post-processing, such as downloading and storage, input/output, georeferencing, file management before and after running a particular user function. Another challenge is the large size of the global datasets, which makes it inconvenient for each user to download and store entire sets at multiple locations, instead of collecting them in one place and providing users with routines for data extraction and use, thus relieving each project partner and user from dealing on their own with large sizes (in the range of terabytes) and highly variable formats of data sources.

GRID-enabled platform could be very useful to download and store (or interface with) the required global datasets and to provide them together with tools to pre- and post-process input data resources, allowing users to extract them within required geographic extents and temporal ranges and into the formats directly feedable into SWAT model.

Global datasets with large sizes are involved. Entire set of TRMM data is in the order of 200 Gb, SSM/I – 1Tb, AMSR-E – 150 Gb. It is therefore necessary to download and store global datasets within one platform (such as ESIP data storage elements) and provide users with sub-setting functions both spatially and temporally, generating data files with smaller sizes. Besides large size, for these data sources it is mostly required to extract entire global dataset first and then to subset to the required area of interest (space/time).

For MODIS, AVHRR and G-DEM (to extract land cover and vegetation indices, cloud coverage, elevation data, etc.) it is not necessary to extract entire global coverage, but for regional application like BSC, it would be highly convenient to regularly extract and pre-process these datasets for convenient availability and use for the entire region or to subset for particular applications. Again, file sizes involved would be in the orders of terabytes, if mosaiced for entire BSC and stacked for entire operational period of data source.

Below briefly is described a range of issues to handle to successfully extract datasets in a form useful for hydrological investigations (based on experience presented in Milewski et al., 2009a):

- Large storage requirements (terabytes of data needed for each required global data set).
- Automation of processing (to avoid manual handling of data processing steps).
- Wide variety of raw data formats (each sensor stores data in specific format and needs specific codes)
- Need to convert digital numbers into physical units (e.g. recalculation into reflectance, radiance, temperature, albedo, etc.).
- Georeferencing (to reproject various data formats into common system/projection).
- Quality control (examining false rainfall events by checking cloud cover, etc.) and quality assurance (correct data extraction and processing steps involved and thus minimizing handling errors).
- Spatial sub-setting of global datasets into specific area of interest (e.g. BSC, or study site like Rioni).
- Temporal sub-setting (to match period under investigation) or resampling (e.g. SWAT requires daily data, while global datasets provide even 1-3-6-9-hourly sampling intervals) of global datasets into specific time period of interest / sampling frequency of interest.

3.2.2.3 Functionality

Following applications/functions could be proposed for implementation to address above issues (see again experience and methodologies described in Milewski et al., 2009a):

- Read raw binary data and export data into unified image and/or database formats (to be readable by SWAT, i.e. either image format such as GeoTIFF: LULC, DEM, Soils etc., or in tab delimited ASCII or CSV: precipitation, wind, temperature, etc.).
- Pre-process data/images into common projection.
- Stack data/images to produce data in the format required as input for hydrological model (e.g. image grid data picture elements converted into tabular data treated as time series at gauge stations).
- Eliminate spectral variations within and between scenes (sun angle, image brightness equalizing).
- Subset global data sets into user-defined spatial and temporal ranges.
- Apply quality control functions to data (e.g., establishing rainfall events through cloud detection and/or soil moisture changes before and after rainfall event).

- Routines could also be available to upload in-situ data (such as precipitation time series) to cross-correlate and to cross-calibrate them with respect to remotely sensed data sources (comparing RS data time series with in-situ data time series).

Data extraction and preparation procedures outlined above are already quite complicated operations to implement due to multiple sources/formats and multiple processing steps involved, but next logical step would be to provide entire hydrological modelling capability within the ESIP GRID-enabled environment. Substantial advantages would be gained by users in terms of both data storage and processing capabilities. Final step would include the implementation of visualization and web processing tools such as BASHYT and OGC services.

Again, specific functionalities suggested above could be built keeping into consideration generic portal functionalities specified in enviroGRIDS D2.2 (see subsection 6.5):

- Gather data resources – The Web Portal should allow integration of data resources and creation of metadata which is associated with data
- Visualize data resources – Depending on the access rights a user can search, visualize, download desired data resources
- Administrate data resources – Copy data from one location to another, rename files, create replicas, manage replicas, etc.
- Define scenarios – define scenarios for the SWAT model by defining parameters, choose input resources, etc.
- Execute Earth Science oriented applications
- Run scenarios – execute scenarios over Grid
- Administration tools – management of users, management of portal, etc.

3.2.2.4 Use Scenarios

There could be two complementary ('manual' and 'automated') scenarios proposed for testing the feasibility of approaches proposed in this section:

1. Testing the usability and sufficiency of above described data resources and their utility through the application of 'conventional' hydrological modelling tools to a particular case site of Rioni Basin. 'Manual' data extraction and model building could help better define specific knowledge on data formats as well their relevance for the required hydrological modelling application.
2. Utilise enviroGRIDS tools (or some of their elements as found most appropriate) described in the enviroGRIDS D2.7 (section 6) to build and implement data use applications within GRID-enabled environment such as ESIP.

The following enviroGRIDS tools and data use applications seems particularly appropriate for the case studies (adapted from enviroGRIDS D2.7, section 6):

Data provider user

- Data Insertion (accesses and uploads required global and national data to enviroGRIDS portal)

Specialist user

- Data Extraction (user selects desired datasets falling within required geographical and temporal extents and using Grid data management downloads in desired format)

- SWAT Calibration and Execution (building data and modelling applications either in non-gridified or in gridified environment, as appropriate)

Decision maker / stakeholder user

- Data Visualization (retrieve applications with built dataset and executed model, process results with BASHYT (see enviroGRIDS 2.2, subsection 8.4) and visualize through various OGC compatible web services and kml specifications (see enviroGRIDS D2.1))

Compatibility and integration should be insured with the overall BSC-OS portal architecture, its various components and tools, as presented in enviroGRIDS D6.1 (see Figure 3.11).

The next two sections illustrate the applicability of global data resources implementing SWAT case study for Rioni basin based on manual extraction of requisite datasets.

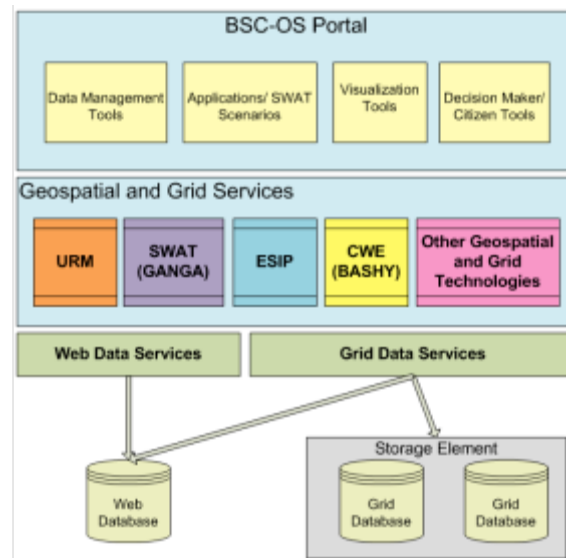


Figure 3.11 EnviroGRIDS Functional Architecture

3.2.3 Data Resources

Keynote theme for the enviroGRIDS project is to apply hydrological modelling to Black Sea Catchment and its main sub-catchments by building GRID-enabled spatial data infrastructure. SWAT hydrological model data requirements were kept in mind when developing data specifications for case applications. For detailed presentation of SWAT data input/output needs in the context of enviroGRIDS reference is made to enviroGRIDS D2.2, section 7.0 and to enviroGRIDS D4.1, section 4. Required datasets should be identified at the global level (remote sensing datasets) as well from national sources (in-situ data such as hydrology, climate, pollution, etc.). Both types of resources are specified below keeping into account applications at the BSC level, as well as for Rioni and other potential sub-catchment case study sites.

3.2.3.1 Global (Remote Sensing)

The various types of remote sensing data available from global sources can be extracted to run hydrological simulations at case study sites or even for the entire Black Sea Catchment. SWAT data needs can be satisfied with the following datasets compiled first in the table form with data uses highlighted in **bold** (reproduced and adapted from Milewski et al., 2009a and then briefly characterized further below, with reference information taken mostly from enviroGRIDS 2.4 and enviroGRIDS 2.6) (Table 3.2):

Table 3.2 Global RS data resources for SWAT application

Sensor	Operation Time(Space)	Data sets	Temporal Resolution	Spatial Resolution	Spectral Resolution	Products
TRMM	1998 - Present (50°N-50°S)	3B42.v6	3-hourly	0.25 x 0.25 degree grid (<25km)	VIR, Microwave (Passive, Active)	Instantaneous Rainfall
GSMaP	2003 - Present Near Real Time (60°N-60°S)	GSMaP_MVK GSMaP_NRT	1-hourly, daily, monthly	0.1 x 0.1 degree grid (<10km)	TRMM microwave, processed with GEO IR	Instantaneous Rainfall interpolated to 1-hourly resolution with meteo IR
AVHRR	1981 - Present	Level 1B	Sub-daily	1km, 4km, 8km	5 Bands: 1-2 = visible; 3-5 = infrared	Cloud detection, NDVI
AMSR-E	2002 - Present	AE_Land3	Daily	25km x 25km	Microwave: 6.9 - 89.0 GHz	Vegetative Water Content & Soil Moisture
SSM/I	1987 - Present	GPROF	4-6 hourly	0.5 x 0.5 deg.	Microwave: 19.4 - 85.5 GHz	Rainfall
SMOS	2009 (3 +2 years)	Soil moisture Ocean salinity	3 days 30 days	35–50 km 200 km	4% 1 pcu	Soil moisture Ocean salinity
MODIS	2002 - Present	Various MOD##	Sub-daily	250m, 500m, 1km	36 Bands: VNIR, SWIR, TIR	Land & Ocean Products
MERIS	2002 - Present	GlobCorine, GlobCover	Sub-daily	300m, 1km	15 Bands	Land & Ocean Products
QuikSCAT	1999 - Present	Wind Product	Twice daily	12.5km, 25km	Microwave: 13.4 GHz	Wind Direction & Speed
Aster G-DEM (release: 2009)	2000 ~ ongoing (83°N-83°S)	G-DEM	N/A	30m mesh in tiles of 1°-1°	N/A	Global 30m digital elevation model data



TRMM (NASA and JAXA) provides global data on **rainfall** for tropical regions (50°N-50°S) using radar, microwave and visible-infrared sensors. Spatial and temporal resolution of rainfall estimates 3 hours and 0.25° x 0.25° degrees. Launched in 1997, continuous coverage from 1998. TRMM 3B42.v6 products with 1998-Present and global spatial coverage can be downloaded from <http://daac.gsfc.nasa.gov>. TRMM can be accessed and downloaded through TOVAS (TRMM Online and Visualization and Analysis System) through <http://disc2.nascom.nasa.gov/Giovanni/tovas>. FTP sites are available as well to download TRMM products. (See TRMM data held with Remote Sensing Systems at <http://www.ssmi.com>, <http://www.ssmi.com/tmi>, <ftp://ftp.ssmi.com> and <ftp://ftp.ssmi.com/tmi>).

GSMaP (Global Satellite Mapping of Precipitation by JAXA's EORC) offers hourly global **rainfall** maps in near real time (about four hours after observation) using the combined MW-IR algorithm with TRMM TMI, Aqua AMSR-E, DMSP SSM/I and SSMIS, and GEO IR data. Near Real Time (GSMaP_NRT dataset) data from FTP is available freely through username / password access obtainable via email request (provided at <http://sharaku.eorc.jaxa.jp/GSMaP/index.htm>). Archived data (GSMaP_MVK dataset) is available since 2003 (see <http://sharaku.eorc.jaxa.jp/GSMaP/crest>, including links to FTP sources). Coverage (60°N-60°S range) is wider than TRMM and includes entire (!) Black Sea Catchment. Spatial and temporal resolution of rainfall estimates are higher as well: 1 hour and 0.1° x 0.1° degrees. Datasets are provided in lat/lon binary grids as well as in CSV ASCII format (latter is useful for hydrological model inputs). (See enviroGRIDS D2.6. Please also explore TRMM and GSMaP against BSC backdrop through Google-Earth layer, click [TRMM GSMaP.kmz](#) to launch, or download from this [link](#).)

AVHRR (Advanced Very High Resolution Radiometer sensor of NOAA's satellite) is a 5-channel scanning radiometer with 1.1km resolution in the visible and near-infrared region. AVHRR is providing since 1981 global coverage and multiple coverages on daily basis for the majority of the Earth's surface, including day and night cloud top and sea surface temperatures, ice and snow conditions, and the capability of producing **cloud coverage** and **vegetation indices**. AVHRR products can be obtained from the NOAA's CLASS website at <http://www.class.noaa.gov>.

AMSR-E (Advanced Microwave Scanning Radiometer), is the microwave radiometer mounted on the EOS' sun-synchronous Aqua satellite launched in 2002, measuring daily microwave radiation (brightness temperatures) across 12 channels. 56 km mean spatial resolution brightness temperature data is resampled into a 25 km grid. The AMSR-E data can be downloaded from NASA's DAAC at <http://daac.gsfc.nasa.gov>. Various products can be used to observe atmospheric, land, oceanic, and cryospheric parameters, including precipitation, sea surface temperatures, ice concentrations, snow water equivalent, surface wetness, wind speed, atmospheric cloud water, water vapour, **soil moisture** (<http://weather.msfc.nasa.gov/AMSR>). (See AMSR-E data held with Remote Sensing Systems at <http://www.ssmi.com>, <http://www.ssmi.com/amsre>, <ftp://ftp.ssmi.com> and <ftp://ftp.ssmi.com/amsre>). (Unfortunately, as of October 2011 AMSR-E instrument stopped producing data due to a problem with its antenna.)

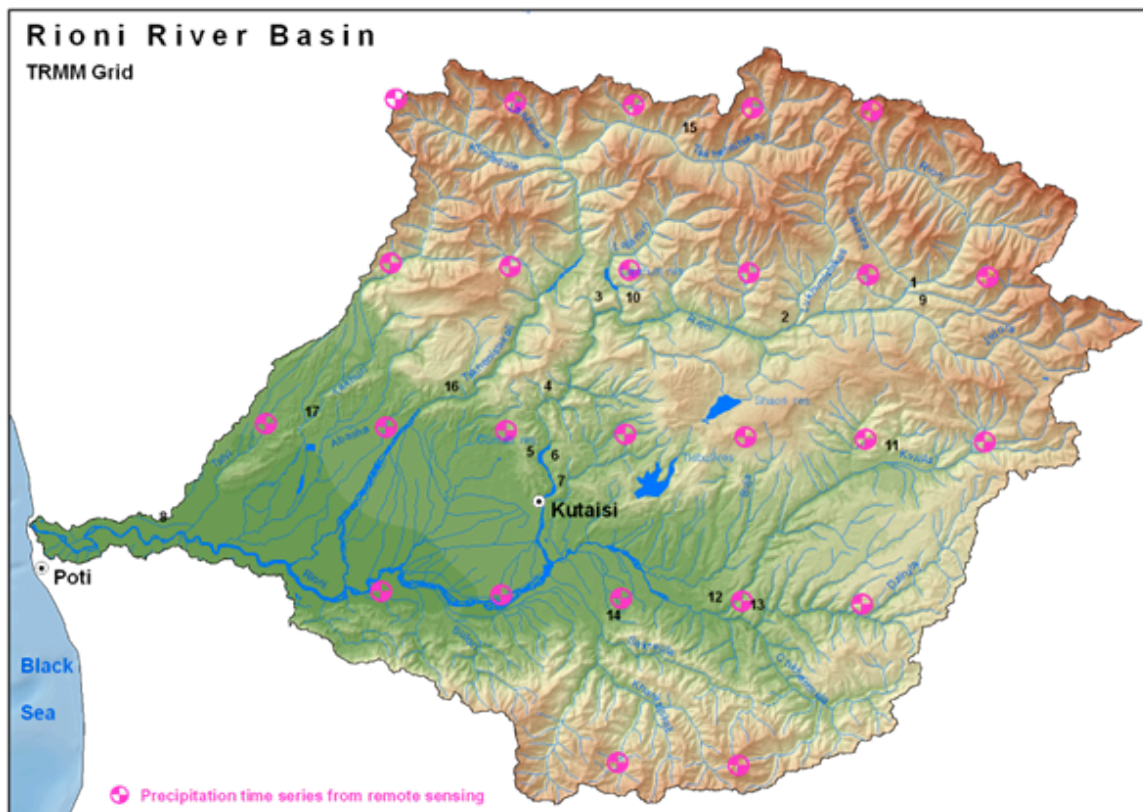


Figure 3.12 TRMM Precipitation Time Series 0.25° x 0.25° grid points, falling within the Rioni Basin.

SSM/I sensor is a 7 channel passive microwave radiometer operating at four frequencies onboard the Defence Meteorological Satellite Program (DMSP) series of polar orbiting satellites. SSM/I provides instantaneous **rainfall** rates on a global coverage at a spatial resolution of 0.5° x 0.5° from 1987 to present. The SSM/I data can be downloaded from. (See SSM/I data held with Remote Sensing Systems at <http://www.ssmi.com>, <http://www.ssmi.com/ssmi>, <ftp://ftp.ssmi.com> and <ftp://ftp.ssmi.com/ssmi>).

SMOS (Soil Moisture and Ocean Salinity of ESA), launched on 2 November 2009, low-Earth, polar, sun-synchronous, 23-day repeat cycle, 3-day sub-cycle satellite with microwave imaging radiometer for global and continuous observations of **soil moisture** (every three days within an accuracy of 4% at a spatial resolution of 50 km) and **ocean salinity** (0.1 practical salinity unit, psu, for a 30-day average over an area of 200x200 km), see <http://www.esa.int/SPECIALS/smos> for more.

MODIS (Moderate Resolution Imaging Spectroradiometer of NASA) is aboard the Terra and Aqua Earth Observing Satellites launched in 1999. MODIS has a temporal resolution of 1 to 2 days and a spatial resolution of 250m-1km. MODIS provides measurements in 36 spectral bands. The generated data can be used to examine global processes on the land, in the oceans, and in the lower atmosphere. MODIS products can be downloaded from <http://modis.gsfc.nasa.gov>. Binary files can be accessed freely through ftp feeds. See section 3.3 for description of MODIS **land cover** products useful for gridified web services.

MERIS (Medium Resolution Imaging Spectrometer, ESA, <http://envisat.esa.int/instruments/meris>) is the earth orbiting ENVISAT mission satellite with 15 spectral band imaging spectrometer onboard, designed and operated for applications in ocean and coast (ocean colour/biology), land (vegetation) atmosphere (clouds/precipitation). Products include freely available 300m resolution land cover mapping produced by GlobCover project through automatic and regionally-tuned classification of a MERIS time series, with 22 land cover classes defined with the UN **Land Cover** Classification System. Another related GlobCorine

project is also based on Envisat MERIS fine resolution (300m) mode data and was executed for two periods, producing two Pan-European land cover/use maps for 2005 and 2009, latter completing the coverage of last remaining gap for the European continent and BSC (countries of Caucasus).

QuikSCAT microwave radar polar orbiting satellite provides high resolution measurements of **near-surface winds** over global oceans approximately twice daily. The spatial resolution is 25km. The data can be downloaded at <http://manati.orbit.nesdis.noaa.gov/quikscat>. (See also QuikSCAT data held with Remote Sensing Systems at <http://www.ssmi.com>, <http://www.ssmi.com/qscat>, <ftp://ftp.ssmi.com> and <ftp://ftp.ssmi.com/qscat>)

Aster G-DEM (NASA and Japan's ERSDAC). Global **DEM** is acquired by Aster satellite mount sensor covering all the land on Earth at a resolution of 30 m. The distribution started in June 2009 (upgraded to version 2 in October 2011) and is freely accessible via electronic download from the ERSDAC of Japan <http://www.gdem.aster.ersdac.or.jp> or from DAAC of NASA. Files are provided in 1° x 1° degree tiles in GeoTIFF format, with 100 tiles per one time download. (See enviroGRIDS D2.4 and enviroGRIDS D2.6.)

3.2.3.2 Local (hydrometeorology, soils)

The basic local in-situ data resources for hydrological modelling application (see enviroGRIDS D4.1 for SWAT data requirements) available in the project partner (GeoGraphic) files include the following tabular time series datasets (source Namakhvani ESIA, 2010, with data from National Environmental Agency of Georgia):

Hydrology: Excel files with daily flow rates for Alpana (1996-2009), and Namakhvani (1996-2005) hydrological stations available for last 10 years, as well as monthly multi-year average data for Chaladidi hydrological station (1990-1993, 1995-1998, 2008-2009).

Note: daily flow rates data for R. Rioni – Alpana hydrologic station of 2000-2005 and 2007-2008 are approximate (estimated discharge rates based on water level measurements recalculated through level/discharge curves).

See Figure 3.13 depicting map with currently operational and historic hydrological stations of R. Rioni Basin.

Meteorology: Excel files with 3-hourly meteorology data for Oni (1998-2000), Kutaisi (1998-2009) and Poti (1998-2009).

Pollution: Monthly average values for Rioni River water quality for the period of 2007 (January)-2010 (July) for four locations: Oni, Kutaisi upstream, Kutaisi downstream, Chaladidi, and the following parameters: temperature; rigidity; transparency; suspended particulate matter; pH; dissolved oxygen; saturation; biological oxygen demand; nitrite nitrogen; nitrate nitrogen; ammonium nitrogen; ammonia; total nitrogen in mineral compounds; orthophosphate; sulphate; chloride; hydro carbonate; Na+K; calcium; magnesium; iron.

In addition, following ad hoc pollution measurement is available for middle section of Rioni: 11 GPS measurement points along Rioni with in-situ measurement (Dissolved Oxygen, pH, Conductivity, Temperature), out of which 7 GPS measurement points were subjected to ex-post analysis for more parameters (Total Suspended Solids, COD, Nitrite, Nitrate, pH, Hardness, Total Phosphorus, Total Kjeldahl Nitrogen, Total Coliform). These measurements were performed in the context of the planned Namakhvani HPP Cascade Project Environmental and Social Impact Assessment (source Namakhvani ESIA, 2010).

Above listed local in-situ data (available through different project) were contributed by GeoGraphic for the benefit of enviroGRIDS project objectives, but for hydrometeorology data with wider spatial and temporal coverage additional enviroGRIDS resources would be required for application to data holder, the National Environmental Agency of Georgia.

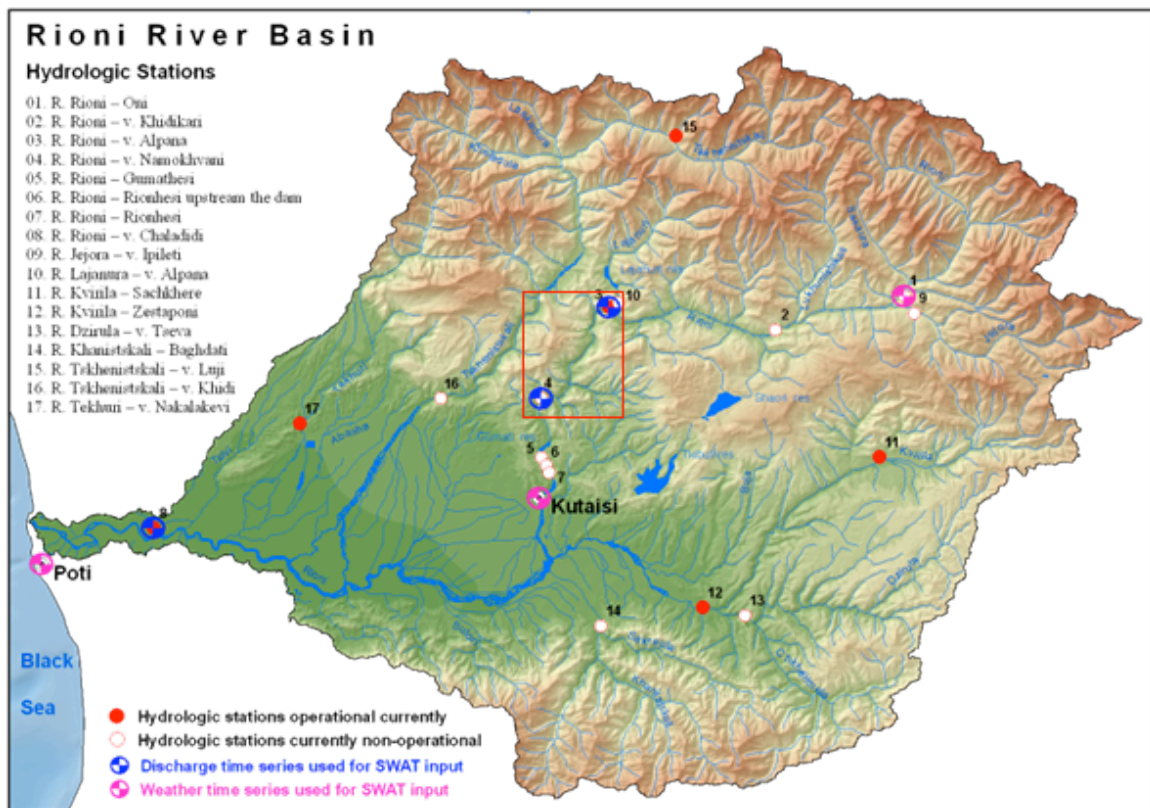


Figure 3.13 Hydrological and meteorological stations in the Rioni Basin (box indicates location of Namakhvani hydropower cascade with 3 planned reservoirs on R. Rioni)

Soils: National soil coverage map 1/500,000 scale produced by Urushadze, 1999 (see map on Figure 3.13), is available in shapefile format. More detailed soils maps are accessible as well, though on site specific basis.



Figure 3.14 Soils of Georgia and in the Rioni Basin.

3.2.4 Results (SWAT)

In this section data sources and methodologies outlined above are applied to handle SWAT hydrological modelling in case of Rioni catchment, based on manual extraction of requisite data. Basic datasets and parameters used are first explained and then initial results of the model run are presented, demonstrating the utility and compatibility of the combination of global and local data sources to the case study catchment processes. But first we reiterated here some study objectives and basic figures of the case area.

Objective of this work is to explore the approaches for utilizing remote sensing and hydrological modelling tools (such as SWAT, Arnold et al., 1998, and ESIP, Gorgan et al., 2009) promoted through the enviroGRIDS BSC-OS Portal (see enviroGRIDS, D6.1) and testing related methodologies through application to the Rioni River Basin (defined in enviroGRIDS, D2.4). After short characterization of the river basin, case study requirements are defined and potential for grid-enabled remote sensing applications are outlined, specified with regard to potential user profiles, data sources, applications, functionalities, and use scenarios, formulated earlier.

Aims of the case study are therefore to:

- apply enviroGRIDS tools (SWAT, ESIP)
- utilise and test applicability of global data sources
- set up hydrological model for Rioni R. Basin
- assess relevance of local data sources, and recommendations for model refinement
- specifying applications for ESIP grid platform

The basic parameters of the Rioni Basin are as follows:

- Length ~330 km, catchment area ~13,400 km²
- 3 main right tributaries (Greater Caucasus), 3 main left tributaries (Lesser Caucasus)
- Population in the basin >500,000 (including Kutaisi ~250,000 and Poti ~50,000)
- 5 reservoirs and 9 HPPs operating, at least 3 more reservoirs / HPPs planned

3.2.4.1 Land (hard) data

As indicated on the diagram below (see Figure 3.15), the following datasets were used to define the land related part of the ArcSWAT input:

- GlobCover land classification (layer based on 300 m Envisat MERIS)
- Food and Agriculture Organization (FAO) global soil cover
- SRTM (90 m) was used as DEM grid

These datasets, characterizing the land parameters within the catchment basin, form the 'hardware' part of the SWAT model, and can be considered as constant frame within which the climate parameters are at play as variables. Land datasets can be further improved and refined, as suggested further below.

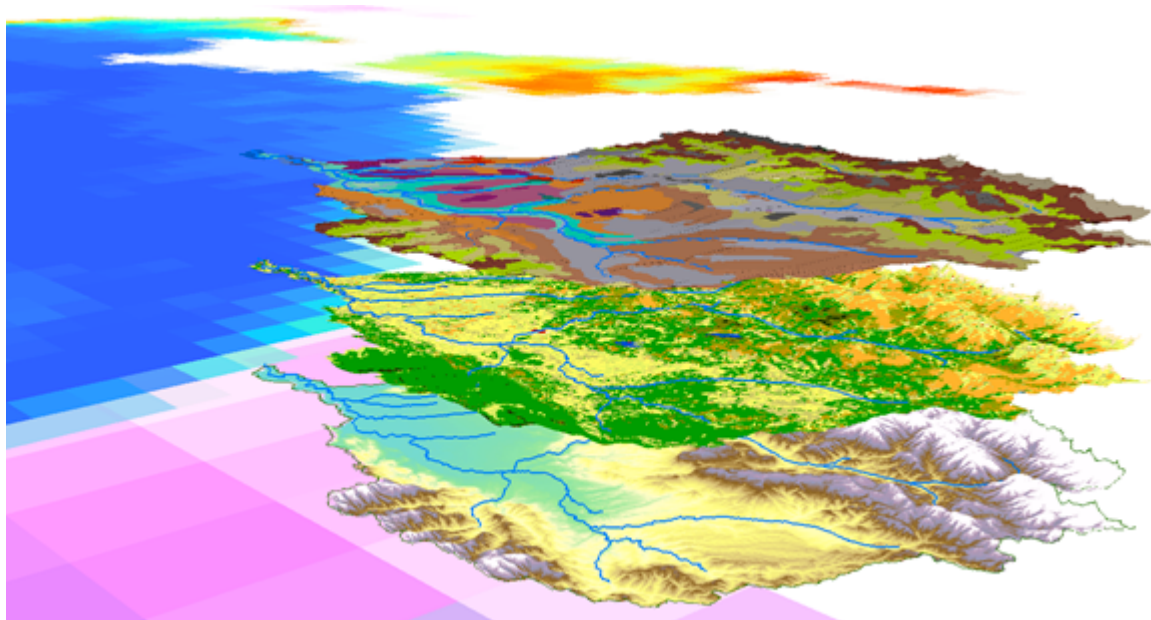


Figure 3.15 Global sources used in hydrological model setup for Rioni (soils, land cover, DEM). (Background shows summer chlorophyll in the Black Sea and sample TRMM grid on landward side.)

3.2.4.2 Climate (soft) data

Section 3.2.3 provides references for global as well as available local sources of rainfall and other climate data required to populate and animate the SWAT model. Data from global sources is typically available as zipped binary and/or as ASCII files, downloadable through FTP sites, but significant pre-processing and transformation steps are required to generate location bound time series in the format required for SWAT input. In case of relatively smaller catchments easiest way to generate required precipitation time series (such as Daily Accumulated Rainfall in this example) is to utilise existing web instruments such as NASA's Giovanni accessible through TRMM Online Visualization and Analysis System, abbreviated as TOVAS: http://disc2.nascom.nasa.gov/Giovanni/tovas/TRMM_V6.3B42_daily.2.shtml. We employed GIOVANNI to generate rainfall data in ASCII time series format for a location in the Rioni catchment. Example of such a dataset for Rioni basin can be derived by executing the following link http://disc2.nascom.nasa.gov/daac-bin/Giovanni/tovas/Giovanni.cgi.pl?west=42.25&north=42.5&east=42.25&south=42.5¶ms=0|3B42_V6&plot_type=Time+Plot&byr=2000&bmo=01&bdy=1&eyr=2011&emo=01&edy=31&begin_date=2000/01/01&end_date=2011/01/31&cbar=cdyn&cmin=&cmax=&yaxis=ydyn&ymin=&ymax=&yint=&ascres=0.25x0.25&global_cfg=tovas.global_cfg.pl&instance_id=TRMM_V6&prod_id=3B42_daily&action=ASCII+Output for the following sample location in the catchment: lat=42.50° lon=42.25° (some time is needed for ASCII time series table to generate in the browser). One could also generate same dataset through TOVAS Daily Rainfall by opening webpage at http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=TRMM_3B42_Daily, inserting extents west=42.375° east=42.625° north=42.375° south=42.125° (corresponds to same sample location as above), selecting Time Series option, and pressing Generate Visualization to initiate processing.

Some extra effort using MS Office tools would be required to generate data in the format accepted as SWAT input, i.e. as ASCII file prefixed with start date and continuing with precipitation series in mm/day of as many days as determined by the time interval. The typical result should look as follows in the table below (selected parameter: Daily TRMM 3B42(V6) Accumulated Rainfall; selected point: lat=[42.50N], lon=[42.25E]; undefined/missing value: -99; unit: (mm); start date: 20000101 in year-month-day):

20000101	0	0	0	0	18.3059	0	0	6.1306	0.7481
----------	---	---	---	---	---------	---	---	--------	--------

0	-99	0	-99	0	0	-99	-99	-99	0
-99	-99	-99	-99	-99	-99	-99	-99	0	0
-99	-99	0	0	-99	-99	-99	0	-99	-99
-99	0	0	-99	-99	-99	0	-99	-99	-99
-99	-99	-99	-99	-99	0	-99	0	0	0
3.96	0	0	0	0	0	0	0	0	0
0	0	0	0	0	39	0	0	0	0
8.16	0	0	0	0	0	0	43.98	0	0
0	0	0	0	0	0	0	0	0	...

3.2.4.3 Hydrological modelling results

Assumptions for setting up the model input data include the following: *Precipitation* data for 4 grid points were used as inputs into the model, covering the TRMM data period starting 01.01.1998 till 31.12.2009 matching available hydrological *discharge* daily time series datasets at two locations (vv. Alpana and Namakhvani). *Temperature* climate of Rhone River (Kazi Rahman, UNIGE, personal communication) for the same time span as precipitation and discharge was used to emulate mountainous region of Rioni Basin, which worked quite well because watersheds share similar characteristics (research is underway to investigate the utility of satellite based Land Surface Temperature, LST, from global sources, such as MODIS or METEOSAT space products).

Figure 3.16 displays screenshot map of the SWAT model setup with data inputs described in this section:

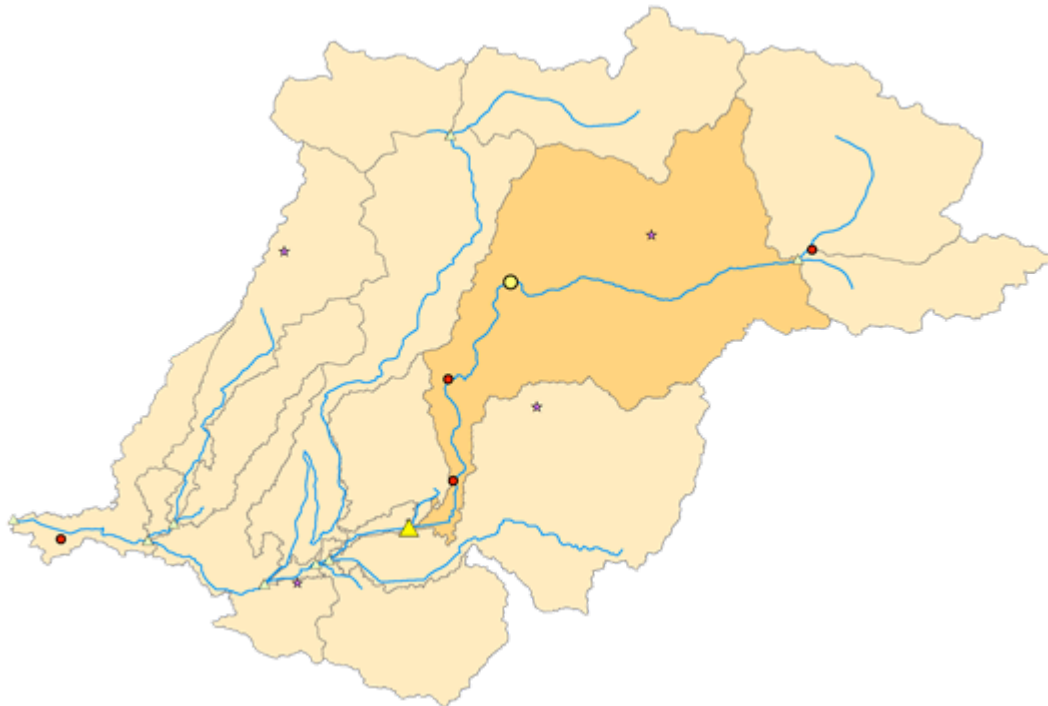


Figure 3.16 SWAT setup for Rioni River collating local hydromet & global TRMM precipitation data.
Note: observed data several km upstream the simulated data (Alpana versus Outlet 8).

The following parameters and tools were applied to setup the model in addition to ArcSWAT module to prepare input and run SWAT 2009 (<http://swatmodel.tamu.edu/software/arcsbat>) in ArcGIS environment:

- Model interface: ArcSWAT2009; year of study: 1998-2009; warm up period: 1998-2000, calibration period: 2001-2006; validation period: 2007-2009; time step: daily average.
- pcpSTAT program (see http://www.brc.tamus.edu/swat/manual_pcpSTAT.pdf) was used to prepare weather generator to populate SWAT database2009.mdb with userwgn.dbf table for the case study area.
- R project statistical computing code developed by Kazi Rahman (UNIGE, personal communication) was applied to calibrate simulated discharge against measured discharge time series.

The first initial results of the hydrological simulation are presented on Figure 3.17 below. Despite numerous assumptions and approximations made, and lack of detailed local data, visual check results are surprisingly positive, demonstrating that further improvement, to achieve acceptable statistical measures (such as Nash-Sutcliffe Efficiency), are feasible and should be pursued as outlined in the recommendations sub-section.

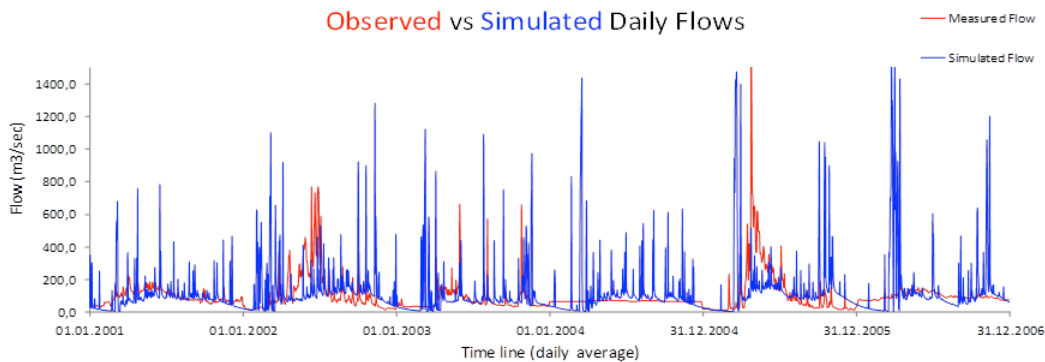


Figure 3.17 Observed versus SWAT simulated daily flow calibration

3.2.5 Recommendations for model improvements

As SWAT is most sensitive to climate parameters (precipitation in particular), quite acceptable first approximation matching of the simulated and measured discharge rates indicates, that TRMM can be considered as acceptable source for model data population. Nevertheless, further significant improvement of statistical parameters can be achieved with fine tuning of at least the following components:

- Apply 30 m per grid point G-DEM instead of 90 m SRTM, with correction of potential inaccuracies in standard product both upstream (mountainous areas) and downstream (lowland areas).
- Utilise gridified pre-processing routines for geometric and atmospheric correction of satellite imagery (developed within ESIP/GreenLand environment for Istanbul and Rioni cases, see sections 3.1 and 3.2).
- Apply Landsat land use / land cover classification at 30m resolution instead of GlobCover with one order of magnitude coarser resolution. GlobCover could be used for definition of land cover classes and as classification training seeds.
- Enable more TRMM time series grid gauge points by delineating smaller area sub-watersheds with DEM-defined streams (in SWAT each type weather data in each sub-watershed is linked to one gauge).
- Employ available national soil database (such as Urushadze, 1999) by developing suitable look up table for SWAT input, complementing it with soil field sampling at certain groundtruthing locations.
- Implement hydropower stations & dams utilising reservoir definition function available in the SWAT and reservoir hydrology data (or investigate usability of radar altimetry, see <http://www.altimetry.info>).

- Acquire land surface temperature (sources could include MODIS LST product, or Meteosat Second Generation MSG available at 15 min or daily timestep and 1 km spatial resolution at nadir, see e.g. <http://landsaf.meteo.pt>), evapotranspiration (MSG, 30 min or daily, 3 km at nadir), soil moisture (12.5 km, see <http://www.eumetsat.int/Home/Main/DataProducts/Land/index.htm?l=en>), wind speed, solar radiation, relative humidity and other climate data, obtained either from global remotely sensed and/or local in-situ sources.

Latter is feasible as described in Mannaerts et al. (2011), employing extraction tools in ILWIS open source (see <http://52north.org/download/ILWIS>, sample data at ftp://ftp.itc.nl/pub/52n/gnc_exercisedata/gnc_data) and GEONETcast ([http://52north.org/download/Earth Observation/GEONETCast/GEONETCast Toolbox](http://52north.org/download/Earth%20Observation/GEONETCast/GEONETCast%20Toolbox)) Toolbox. These instruments could also be consulted when developing download and processing routines within ESIP environment.

3.2.5.1 Specifications for ESIP and GreenLand Based Processing

Hydrological modelling application efficiency and range of capabilities can be advanced with the further development of the GRID-enabled ESIP platform (Gorgan et al., 2009), as GreenLand 2 implementation (see <http://cgis.utcluj.ro/applications/greenland>) is now equipped with the ability to include various GRASS GIS and RS image processing modules (<http://grass.osgeo.org>) as well as file management functions (such as ftp upload and OGC feature extraction services). Figure 3.18 reproduces the basic ESIP flow diagram with essential elements of the platform, defining functional requirements for utilizing grid-enabled remote sensing data input/transformation/output and for preparing and extracting data from global or local sources. Subsequent processing through SWAT model and hydrological output presentation and dissemination could then be accomplished. The following implementation sequence could be suggested in this regard:

- Develop ESIP processing workflow to derive 30m accuracy land cover / land use from Landsat imagery (with approaches similar to those implemented for the Istanbul case study application).
- Elaborate ESIP processing workflow to generate rainfall, land surface temperature, evapotranspiration and soil moisture time series from global data sources and in SWAT compatible format.
- Run gridified SWAT model with generated input data products.
- Reporting SWAT results with BASHYT.

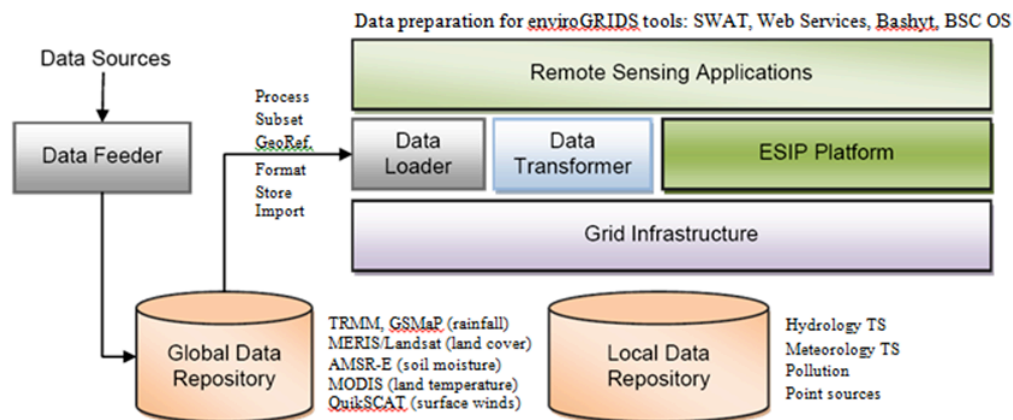


Figure 3.18 Input sources, data flow and processing suggestions for ESIP based platform

3.2.6 Synergy with PEGASO (Georgia CASE Site)

Application of the remote sensing tools to aid coastal & watershed assessments and planning is the stated objective of one of the so called CASE Sites under the FP7 PEGASO project, see map (left) on Figure 3.19 and brief description of the Georgia CASE at http://www.pegasoproject.eu/wiki/Guria_Coastal_Region. In

synergy with tools developed under enviroGRIDS project and applied to its WP2 case studies, the SWAT, ESIP and other enviroGRIDS instruments will be used in coastal watershed of R. Supsa, south of R. Rioni.

Objectives for enviroGRIDS input into Georgia CASE Site are formulated, as follows to:

- Create a web processing tool to assess beach erosion using various models
- Evaluate impacts of sea level rise and/or changes in sediment loads
- Quantify economic impacts of coastal erosion.

Anticipated data input needs include DEM and bathymetry, river discharges and sediment loads, soil cover and vegetation maps, beach profiles, satellite images.

Tools and instruments to use include SWAT, beach erosion models written in MATLAB, R, and GRASS.

Together with enviroGRIDS WP2 (UNIGE and GeoGraphic) partners involved from PEGASO WP3 would include University of Aegean (Dept. of Marine Sciences), as well as the Black Sea Commission Permanent Secretariat (through PEGASO task manager and case coordinator, both commissioned by BSC PS).

Other related case tasks include testing the mutual compatibility of watershed and coastal sustainability spatial indicators at regional and local scale (similar to one shown on the right map of Figure 3.19, see Gvilava et al., 2011) and through integrated application of PEGASO toolbox and enviroGRIDS instruments, as advocated in Lehmann et al. (2009).

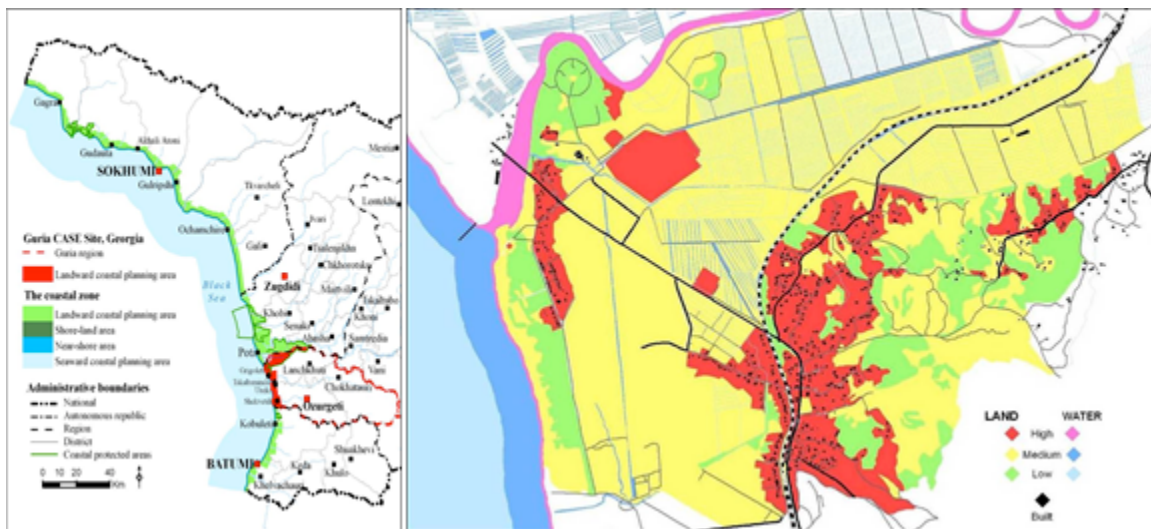


Figure 3.19 Pegaso/eG Georgia Case Site (left) & indicator map for Tskaltsminda coastal unit (right)

3.3 Remote Sensing products for SWAT outputs validation

3.3.1 Introduction

Black Sea catchment covers more than 2 million of square kilometers spread over 24 countries; such a politically fractured huge area presents a great challenge to manage on a sustainable way.

Therefore one of the objectives of EnviroGRIDS is to provide past, present and future state on the environment across hydrological modelling using SWAT. Great work have already been done to collect, match formats of large datasets collected by all countries to be able to run SWAT on the entire BSC. Another source of information should be better exploited i.e. large collections of global datasets of derived remote sensing products (RSP) which are easily available through web services.

Indeed, the combined use of remote sensing products with in-situ data could significantly improve modelling complex environments, such as the hydrology of the Black Sea watershed.

Remote sensing observations seem promising in hydrologic research as they provide a mean to observe hydrological state variables repeatedly over large areas (Schmugge et al., 2002).

Their integration in hydrological modelling processes could facilitate the calibration step, by complementing input data where reliable data are scarcely available, alternatively they could also serve for model validation.

3.3.2 Remote sensing products

3.3.2.1 RSP commonly used in hydrologic research

Even if remote sensing products (RSP) could not replace ground based methods to provide high quality information, their advantages include the provision of homogenous datasets over broad area with large temporal resolution. Therefore RSP could be a valuable source of information for many applications related to earth observations but also to better understand the water cycle, and could also efficiently complement data gaps.

For more details on useful datasets specifically for SWAT modelling, at local or Black Sea catchment extent, refer to chapter 3.2.3.1.

There is abundant literature on the use of RSPs to calibrate hydrological models and/or to validate results.

Below a non-exhaustive list of the most commonly used RSP in hydrologic research for model calibration or result validation:

- Vegetation indices
- Crop mapping, phenological cycle
- Leaf Area Index (LAI)
- Land use/cover
- Evapotranspiration (ET)
- Soil moisture
- Snow cover
- Land surface temperature
- Climatic data, precipitation and temperature

Generally, hydrological models simulate interactions between soil, vegetation and atmosphere and therefore parameters related to these three categories will be of great importance.

Land-derived remote sensing products are essential for land-use/cover classification allowing change detection analysis or for ameliorating vegetation categories and improving crop type determination (Boegh et al., 2004; Sánchez et al., 2010; Wardlow and Egbert, 2008).

Land uses coupled to vegetation parameters such as vegetation indices, LAI, soil moisture, land surface temperature (LST) are particularly important to evaluate soil water balance and evapotranspiration (ET) rates (Boegh et al., 2004; Immerzeel and Droogers, 2008; Immerzeel et al., 2008; Schmugge et al., 2002; Stefanova, 2011; Xie and Zhang, 2010).

Finally, radar observations provide estimated climatic data such as rainfall, precipitation, over large areas on high temporal range (Krajewski and Smith, 2002; Milewski et al., 2009b; Yan et al., 2010).

3.2.2.3 Quality assessment

Global long-term series of RSP, generated in a routine manner, may contain errors or artefacts that need to be rectified in order to allow the identification of real earth changes processes.

Existing RSP of high processed level (level 3 or 4 of products issue from EOS instrument) are specially designated for multidisciplinary science, since these products are already processed and can be directly interpreted without requiring expert knowledge in remote sensing. An assessment of the quality of the data provided is usually included in the final product and should always be inspected before use (Morissette et al., 2002; Roy et al., 2002).

Literature is very abundant on the assessment of RSP quality. The following references are readings suggestions for soil moisture index (Brocca et al., 2010; C. Albergel et al., 2009; Dorigo et al., 2010; Li et al., 2010; Loew and Schlenz, 2011), for evapotranspiration (Gao and Long, 2008; Kalma et al., 2008; Q. Mu et al., 2011), and for LAI (Bavec et al., 2007).

As the main final purpose is to facilitate the access to existing RSP for the entire Black Sea catchment that will serve at the same time local or large scale studies, emphasis has been put on available products of higher resolution covering large areas on high temporal ranges.

3.3.3 Evaluation on Vit river basin, Bulgaria

Large amount of existing RSP are freely available through the web. A challenge was to find freely-available products of scientific quality, covering the entire Black Sea catchment, with large temporal availability, necessary especially for improving crop's type monitoring but also more generally for helping SWAT's calibration and results validation.

In order to evaluate the most useful products, a selection of existing RSP have been collected and provided to UNESCO-IHE for validation of SWAT results in the Vit River basin (Bulgaria) (Stefanova, 2011).

According to their assessment, a selection of products (bolded in Table 3.3) has been chosen to feed GreenLand's application for the entire Black Sea catchment. Note that unselected products are not necessarily useless, but may have not been evaluated due to lack of time.

Table 3.3 List of provided products and their selection for SWAT results validation

#	Name	Product	Since	To	Resolution	Frequency	Retain
1	EVI	MOD13	18.02.2000	Present	250m	16-day	No
2	NDVI						No
3	FPAR	MCD15	05.03.2000	Present	1km	8-day	No
4	LAI						Yes
5	GPP	MOD17	18.02.2000	Present	1 km	8-day	No

6	PsnNet						No
7	Annual NPP		01.01.2000	31.12.2010		yearly	Yes
8	Snow8d	MOD10	24.02.2000	Present	500m	8-day	No
9	MaxExtent						No
10	LST	MOD11	05.03.2000	Present	0.05 Deg	daily	No
11	ET	MOD16	01.01.2000	31.12.2010	1 km	8-day (available monthly or annual)	Yes
12	LE						No
13	PLE						No
14	PET						Yes
15	SoilMoist	AMSR-E	18.06.2002	Present	25 km	Daily (available weekly)	No
16	SurfTemp						No
17	VegWat						No

3.3.3.1 Data Storage

In order to cover the entire Bulgarian case site, only one tile of RSP was necessary. However, even within this small coverage area, large data storage is needed. Indeed, the total memory size of the provided RSP, containing original downloaded product and processed images (file type conversion, band extraction), was more than 26 Gb, as shown in Table 3.4 below.

Table 3.4 List of provided products, covered period, number of different dates supplied, total folder size

Product	Period	Dates	Band	Folder Size
AMSR-E	2002-2009	91	SoilMoist SurfTemp VegWat	6 Gb
MOD17	2002-2009	73	GPP PsnNet Annual NPP	700 Mb
MCD15	2002-2009	67	FPAR LAI	500 Mb
MOD11	2002-2009	64	LST	720 Mb
MOD16	2005-2009	20	ET LE PLE PET	625 Mb
MOD10	2000-2009	117	Snow8d MaxExtent	690 Mb
MOD13	2002-2009	55	EVI NDVI	17.4 Gb
				26.6 Gb

3.3.4 Processes

Usually RSP are distributed via an FTP server and are provided in a compilation of multiple band in *.hdf format. For this testing phase, NDVI images have been selected manually using USGS Global Visualization Viewer, by choosing cloud-free dates, whereas time series of the remaining provided datasets have been randomly selected.

The Process workflow (Figure 3.20) consists in downloading images from an FTP server, then extracting the bands separately and clipping them to the case site area.

Finally the result is uploaded in another FTP from where UNESCO–IHE could retrieve the useful dataset for their analysis.

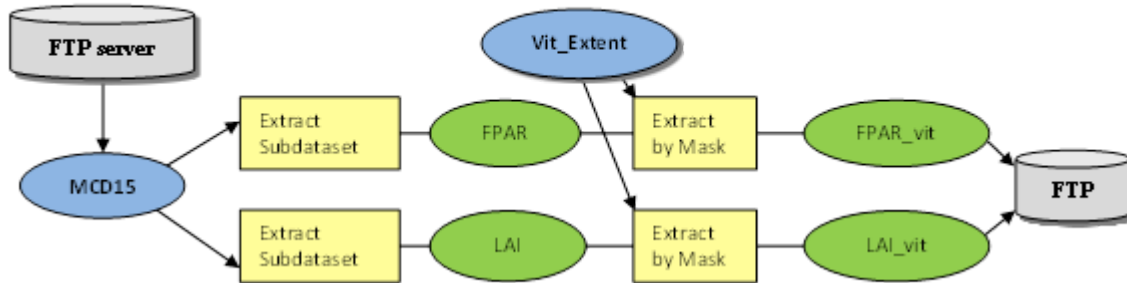


Figure 3.20 Data download and preprocessing workflow, example for MCD15 product

There are several disadvantages in retrieving large datasets without any automation help.

On top of requiring specific software for the analysis, repetitive processes are very unexciting and time consuming and as previously mentioned require powerful computing and storage capabilities. Moreover the retrieved datasets are not made visible to others, and therefore their reuse by other users remains unlikely.

For all these reasons, the GreenLand application will offer among other advantages an effective solution to improve data retrieval efficiency.

3.4 GreenLand Black Sea catchment application

The use of large temporal extents of RSP provides great opportunities for earth modelling. However, as expressed previously, challenges may arise in data retrieval, storage and analysis. Computer science are able to provide smart solutions, as has been demonstrated by the remote sensing data extraction model (RESDEM) developed by Western Michigan University (Milewski et al., 2009a). However similar tools should be even more effective, when directly accessible on internet.

Initially, the main aim of GreenLand's Black Sea catchment application will be therefore to facilitate MODIS land products retrieval for enviroGRIDS's consortium and users at the entire Black Sea catchment.

Only basic functionalities provided by Greenland application will be used to collect and store existing RSP (Figure 3.21), and will offer considerable advantages to facilitate large datasets retrieval.

Time allowed for data searching, downloading, uploading will be considerably reduced with the development of a batch processing option, which automatizes the direct access to existing data repositories.

No particular remote sensing skills, software, or computational performance will be required, as all calculation and data storage will depend on powerful servers as well as on grid computing when necessary. Basically only an internet connection and a user account will be mandatory to process the desired products.

Moreover, an automatized export of any data processed in GreenLand into enviroGRIDS URM geo portal will complete the enviroGRIDS Black Sea data repositories; therefore dataset will be visible and accessible for any users at any time, even after the end of the project. This initiative will encourage the reuse of data that have already been searched and prepared once, therefore avoiding waste of time and work duplication.

3.4.1 Process description

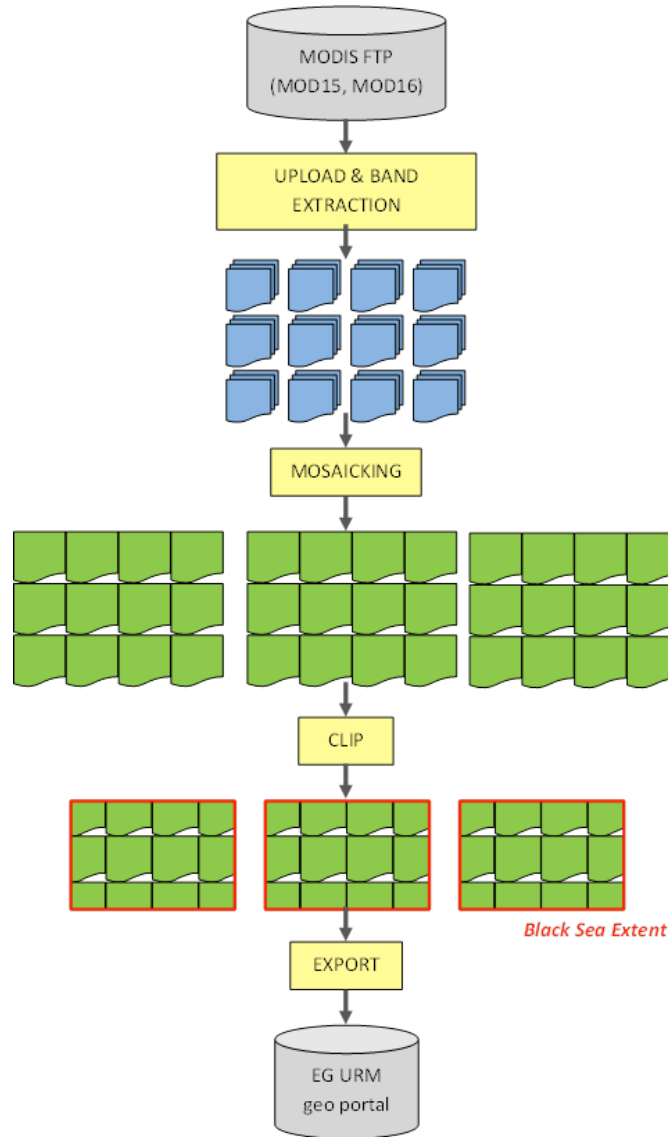


Figure 3.21 General Process Workflow

3.4.1.1 Products

At this initial testing phase, only two MODIS level 4 products will be uploaded within the application; MOD15 and MOD16. These products are multilayers stacks of 1 km resolution provided on 8-day basis, of a maximum file size 5.8 Mb each.

The following data definitions are extracted from EOS data products Handbook (Parkinson, C. L. 2000) (Table 3.5):

“MOD16 consists of surface resistance and evapotranspiration which are essential parameters to global modeling of climate, water balance, and trace gases. Practical applications of this product are for monitoring wildfire danger and crop/range drought.

MOD15 Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) absorbed by vegetation are biophysical variables that describe canopy structure and are essential in calculating terrestrial energy, carbon, water-cycle processes, and biogeochemistry of vegetation. LAI defines the one-sided leaf area per unit ground area (value between 0 and 8), when FPAR measures the proportion of available radiation (400 to 700 nm) that a canopy can absorb (value between 0 and 1).”

Table 3.5 Tested product description

PRODUCT	SHORT NAME	NAME	SPATIAL RESOLUTION	TEMPORAL RESOLUTION	LAYERS	FORMAT
Leaf Area Index	LAI	MOD15A2	1km	8-day	Fpar_1km Lai_1km FparLai_QC FparExtra_QC FparStdDev_1k LaiStdDev_1km	*.hdf
Evapotranspiration	ET	MOD16	1km	8-day	ET_1km LE_1km PET_1km PLE_1km ET_QC_1km	*.hdf

3.4.1.2 Upload

MODIS products are uploadable into GreenLand in their original format *.hdf from our local machine or directly from an FTP server. Files are directly converted into GeoTIFF, band by band separately.

However, at this stage, the user could only upload single files, one by one, or by selecting a short list of them. Therefore an automated process allowing the selection of the entire dataset in a single operation needs to be developed.

A batch processing script, in which a range of dates and specific tiles could be specified, still needs to be conceived. The following information, describing product organization on FTP servers, will serve to create the script.

MOD16

Evapotranspiration products (MOD16) are available on the following FTP server:

ftp://ftp.ntsg.umd.edu/pub/MODIS/Mirror/MOD16/MOD16A2.105_MERRAGMAO/

Products are organized in different folders since 2000 to 2010:

Year (Y2000, Y2001...Y2010),

Day (every each 8-day, D001, D009...D361)

Tile¹ is indicated in the file name (*MOD16A2.A2000001.h00v08.105.2010355153835.hdf*)

Example:

ftp://ftp.ntsg.umd.edu/pub/MODIS/Mirror/MOD16/MOD16A2.105_MERRAGMAO/Y2000/D001/MOD16A2.A2000001.h00v08.105.2010355153835.hdf

MOD15

LAI-FPAR products are available on both Terra (1) and Aqua (2) spacecraft or as a combined (3) product.

Third FTP servers are organized on the same way:

Product (different FTP addresses)

1. MOD15A2.005 (TERRA product) <ftp://e4ftl01.cr.usgs.gov/MOLT/MOD15A2.005/>
2. MYD15A2.005 (AQUA product) <ftp://e4ftl01.cr.usgs.gov/MOLA/MYD15A2.005/>
3. MCD15A3.005 (Combined products) <ftp://e4ftl01.cr.usgs.gov/MOTA/MCD15A3.005/>

Date (every each 8-day since February 2000 (1), since July 2002 (2, 3) to present)

2000.02.18, 2000.02.26...2012.02.10

Tiles are indicated in the file name (*MYD15A2.A2002185.h00v08.005.2007175021713.hdf*)

Examples:

1. <ftp://e4ftl01.cr.usgs.gov/MOLT/MOD15A2.005/2000.02.18/MOD15A2.A2000049.h00v08.005.2006268222954.hdf>
2. <ftp://e4ftl01.cr.usgs.gov/MOLA/MYD15A2.005/2002.07.04/MYD15A2.A2002185.h00v08.005.2007175021713.hdf>
3. <ftp://e4ftl01.cr.usgs.gov/MOTA/MCD15A3.005/2002.07.04/MCD15A3.A2002185.h00v08.005.2007166031220.hdf>

3.4.1.3 Processes (band extraction, conversion, mosaic, subset)

GreenLand application can read directly MODIS original format (*.hdf). During the uploading process, each band is extracted separately and converted into GeoTIFF format.

Twelve MODIS tiles² are necessary to cover the entire Black Sea watershed (see Figure 3.22).

¹ Total tiles covering the globe, 286 tiles

² Total column 4800, row 3600 (1 tile 1200*1200)

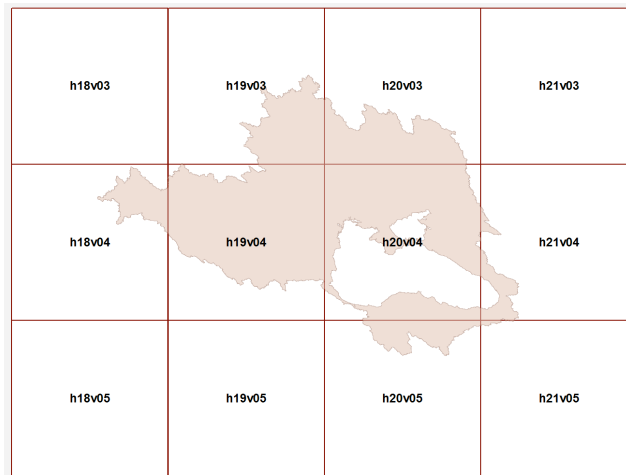


Figure 3.22 MODIS files products to cover the extended Black Sea watershed, horizontal (h) and vertical (v) position

Presently, only two tiles have been successfully merged together. Tiles need to be selected manually from a drop-down list, which is very efficient when a limited number of images are used but unpractical when large amount of data are available (Figure 3.23).

Consequently the next steps will have to improve the available 12 tiles mosaic tool (Mosaic12), so as to be able to merge at once twelve tiles of identical product at the same date, and thus cover the whole region of interest. Also, a smarter solution should be thought to find a more suitable way to select input files from a very long list.

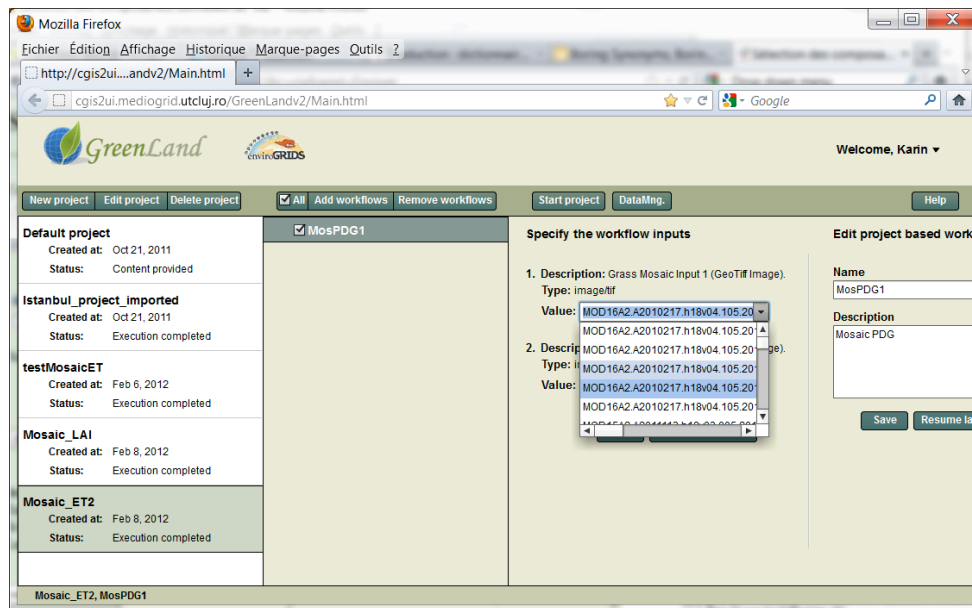


Figure 3.23 Drop-down list of accessible products for GreenLand's tools

In order to reduce file size, the image could be clipped to a smaller extent that better fits the Black Sea watershed.

Therefore, product subsetting is another useful tool that could be implemented in Greenland using either a vector file (*.shp) or geographic coordinates that define the area of interest.

3.4.1.4 Export

All results should be exported into the EG URM data portal. This makes available datasets visible to any potential user, and allows them to be easily retrieved from the enviroGRIDS BSC-OS Portal. The result should be exported into an illustration (*.png) but also as a geo-referenced GRID file (GeoTIFF) when datasets have authorization to be redistributed publicly.

MODIS products are fully described in embedded HDF metadata associated file and in external ECS metadata. A solution should be envisaged to import directly this information, even partially, into EG URM geo portal as a Metadata.

This functionality could be afforded by Greenland application only if the EG URM data portal provides an API for such features. In this case, it will remain to define what kind of fields will be selected to be stored in these metadata files.

3.4.2 Iterations

In order to evaluate SWAT results, or calibrate a model, products of large temporal availability are more suitable.

Both tested products (MOD16, MOD15) are available every each 8-days since February 2000 until December 2010, which represent about 45 products per year.

If we assume that we would like to get full datasets (since February 2000 until December 2010), about 453 dates should be processed for both products, multiplied by 12 tiles each to cover the entire area of interest (and so a total of ~ **10'872** downloaded files will be processed). If all layers (6 bands for MOD15 and 5 bands for MOD16) need to be exported to EG URM geo portal, then (453 dates multiplied by 11 layers) 4'983 files will be transferred (Figure 3.24).

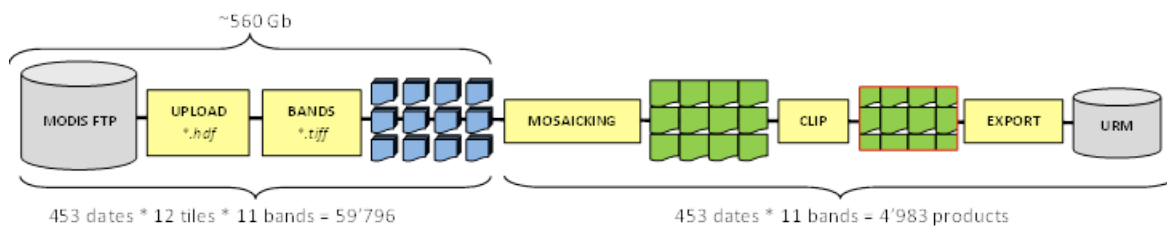


Figure 3.24 Number of expected processed products according workflow steps

3.4.3 File size

Note that the following estimated weights are approximated using one date of MOD16 or MOD15 (Figure 3.25) product and ERDAS software, since the work has not been done yet on the platform.

One tile of MOD15 is approximately 3 Mb for each multilayers *.hdf file, whereas a tile of MOD16 represents approx. 5 Mb (so a total of ~160 Gb for 10'872 downloaded files).

Data conversion into GeoTIFF and band extraction increase considerably the weight of one tile, reaching about 14 Mb for the five separated layers of a MOD16 tile (~3-4 times heavier). The total uploaded file could therefore reach a total of ~ 560 Gb within GreenLand application, without counting the output processed products (mosaics).

Consequently retrieving large temporal global remote sensing datasets necessitates rapidly very large storage requirements.

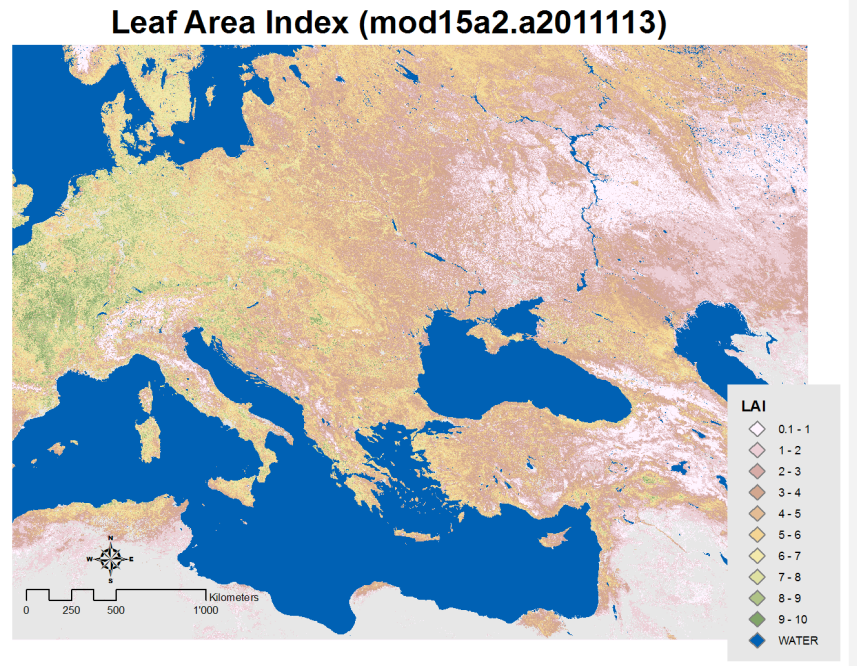


Figure 3.25 Mosaic of 12 tiles of LAI band (MOD15) on 23 April 2011

3.4.4 ESIP application at Black Sea catchment extent

3.4.4.1 ESIP and GreenLand Based Processing

The Mosaic algorithm is one of the most intensive used operations when projecting the Black Sea catchment, based on different tiles represented by satellite images of the central and western part of Europe. That is the main reason why this section describes the behavior of this operator and the required steps in order to execute and to manage the output results. The Mosaic is implemented as a standalone workflow within the ESIP database. It operates on single-banded satellite images, converted in the external data upload processes. In its simplest form this workflow requires two inputs, but an extension of it (called Mosaic12) allows the usage of 12 simultaneous inputs.

Following the steps described in the GreenLand application section, the user should first create a new project and attach to it one or multiple Mosaic instances. In order to obtain a valid result the two inputs (required by the Mosaic workflow) should represent adjacent geographical areas. Then the Grid execution phase will transfer all these inputs to the Grid worker nodes that will process the operations encapsulated within the Mosaic workflow. At any time the user has knowledge about the execution status by periodically interrogating the ESIP and GreenLand backend services.

An example of using the Mosaic workflow is highlighted in Figure 3.26. The two satellite images used as inputs represent adjacent geographical areas of the western continental part of Europe. The result is generated as a single image, where the common areas of the two inputs are overlapped into this output result. The output generated by the Mosaic workflow could be downloaded by the user for further analysis.

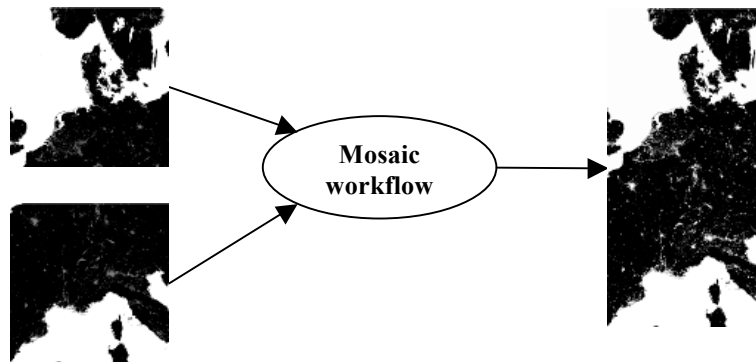


Figure 3.26 Mosaic workflow usage result

3.4.5 Future improvements

Greenland application is constantly changing, by actively offering solutions that meet user needs. The assessment of the usefulness of the platform, for large scale remote sensing datasets retrieval, will be evaluated after testing several improvements that will be developed in the near future.

- Automate the upload of large datasets available on a FTP server, using for instance a batch processing script in which a range of date and specific tiles could be specified
- Merge at once twelve tiles of same product and date.
- Improve the selection of similar products to be merged together among large datasets.
- Subset mosaics to an extended Black Sea catchment using a vector layer (*.shp) or coordinates
- Export results and metadata into enviroGRIDS URM geo portal

4 Conclusion and Recommendations

This document presents a detailed explanation of ESIP and GreenLand platform by using three different case studies.

The GreenLand application offers a friendly user interface for viewing and editing workflows and operators. The description involves the basic operators provided by the GRASS library as well as many other image related operators supported by the ESIP platform. The processing workflows are represented as directed graphs, giving the user a fast and easy way to describe complex parallel algorithms, without having any prior knowledge of any programming language or application commands. Also this Web application does not require any kind of install, because it is a remote application that could be accessed over the Internet.

Despite some improvements that will be implemented in Greenland application to allow fully analyze and process of each scenario, the application seems already very promising to facilitate processing and analysis of satellite images. This powerful tool will be surely very useful for sustainable management of the Black



Sea watershed; moreover the tool will survive to enviroGRIDS project and could become a reference to process images in the Black Sea region by serving other projects (like for instance FP7 PEGASO project).

Greenland will mainly facilitate:

- Large datasets retrieval (allowing establishing a complete database of the region, which will be accessible via the data EG URM geo portal)
- Remotely sensed data pre-processing
- Land use/cover classifications
- Remote Sensing Vegetation indices calculations
- SWAT input data collections

Following improvements and further steps are proposed for local applications of enviroGRIDS tools, like in case of Rioni River Basin study:

- Apply higher resolution 30 m G-DEM in addition to 90 m SRTM.
- Use GreenLand for radiometric, atmospheric correction & classification of Landsat as SWAT LULC input.
- Substitute FAO global dataset with local/national soil cover as SWAT soil input.
- Implement hydropower stations & dams utilising SWAT reservoir function.
- Similar to TRMM precipitation, test LST, ET, soil moisture, climate, other RS datasets as input to SWAT.
- Run gridified SWAT model with tested RS data products and report results through BASHYT.
- Replicate the application of enviroGRIDS tools to Pegaso CASE Site in Georgia.

An update of this deliverable will be provided by the end of the project and will include a detailed description of each fully-processed case study.

5 Abbreviations and Acronyms

AMSR-E	: The Advanced Microwave Scanning Radiometer for EOS
AVHRR	: The Advanced Very High Resolution Radiometer
BASHYT	: Basin Scale Hydrological Tool Decision Support System
BSC	: Black Sea Catchment
DAG	: Direct Acyclic Graph
DEM	: Digital Elevation Model
DN	: Digital Number
DOS	: Dark Object Subtraction
ECS	: Extended Continental Shelf
EG	: EnviroGrids
EOS	: Earth Observation Services
ESIP	: Environment Oriented Satellite Data Processing Platform
ET	: Evapotranspiration



ETM	: Enhanced Thematic Mapper
EVI	: Enhanced Vegetation Index
FAO	: Food and Agriculture Organization
FPAR	: Fraction of Absorbed Photosynthetically Active Radiation
FTP	: File transfer Protocol
G-DEM	: Global Digital Elevation Model
GEMI	: Global Environmental Monitoring Index
GIS	: Geographic Information Systems
GPP	: Gross Primary Production
GPS	: Global Positioning System
GSMaP	: Global Satellite Mapping of Precipitation
GXML	: Guideline XML
HDF	: Hierarchical Data Format
iPDG	: Instantiated Process Description Graph
JDL	: Job Description Language
JSDL	: Job Submission Description Language
LAI	: Leaf Area Index
LE	: Latent heat flux
LST	: Land Surface Temperature
LULC	: Land Use Land Cover
MERIS	: Medium Resolution Imaging Spectrometer
MODIS	: The Moderate Resolution Imaging Spectroradiometer
MS	: Multispectral
NASA	: National Aeronautics and Space Administration
NDVI	: Normalized Difference Vegetation Index
NIR	: Near Infrared
NPP	: Net Primary Production
PDG	: Process Description Graph
PET	: Potential evapotranspiration
PLE	: Potential Latent heat flux
PNG	: Portable Network Graphics
RESDEM	: Remote Sensing Data Extraction Model
RGB	: Red-Green-Blue
RMSE	: Root Mean Square Error

RS	: Remote Sensing
RSP	: Remote Sensing Products
SIS	: Turkish Statistical Institute
SMOS	: The Soil Moisture and Ocean Salinity
SRTM	: Shuttle Radar Topography Mission
SSM/I	: The Special Sensor Microwave Imager
SWAT	: Soil and Water Assessment Tool
SWIR	: Short Wave Infrared
TM	: Thematic Mapper
TM	: Thematic Mapper
TOA	: Top of Atmosphere
TRMM	: The Tropical Rainfall Measuring Mission
URM	: Uniform Resource Management
USGS	: U.S. Geological Survey
UTM	: Universal Transfer Mercator
Vis	: Vegetation Indices
VNIR	: Visible Near Infrared
VO	: Virtual Organization
WIST	: Warehouse Inventory Search Tool
WS	: Web service
XML	: Extensible Markup Language

6 References

- Albergel C., Rüdiger C., Carrer D., Calvet J.-C., Fritz N., Naeimi V., Bartalis Z., and Hasenauer S., 2009: An evaluation of ASCAT surface soil moisture products with in-situ observations in Southwestern France, *Hydrology and Earth System Sciences*, v.13, p. 9.
- Allenbach K. and Gvilava M., 2011: Applying enviroGRIDS Tools to Rioni Basin, in the Extended Abstracts of the OpenWater Symposium, 19th April, 2011, UNESCO-IHE, Delft, The Netherlands (http://hikm.ihe.nl/openwater_eg/Pdf/Abstracts-OpenWater-final-for-print.pdf, accessed (March 31, 2012).
- Arnold J. G., Srinivasan R. Muttiah R. S. and Williams J. R., 1998: Large area hydrologic modelling and assessment: Part I. Model development, *J. American Water Resour. Assoc.*, v. 34(1), p.73-89.
- Băcu V., Rodilă D., Mihon D., Ștefănuț T. and Gorgan D., 2010: Error prevention and recovery mechanisms in the ESIP platform: IEEE 6th International Conference on Intelligent Computer Communication and Processing, p. 411 – 417.
- Băcu V., Ștefănuț T., Rodilă D. and Gorgan D., 2009: Process Description Graph Composition by gProcess Platform: HiPerGRID - 3rd International Workshop on High Performance Grid Middleware, p. 423-430.



- Bavec M., Vukovic K., Grobelnik S., Mlakar Č. R. and Bavec F., 2007: Leaf Area Index in Winter Wheat: Response on seed rate and nitrogen application by different varieties, *Journal of Central European Agriculture*, v. 8 (3), p. 5.
- Boegh E., Thorsen M., Butts M.B., Christiansen J.S., Abrahamsen P., Hasager C.B., Jensen N.O., van der Keur P., Refsgaard J.C., Schelde K., Soegaard H., Thomsen A., 2004: Incorporating remote sensing data in physically based distributed agro-hydrological modelling, *Journal of Hydrology*, v. 287(1–4), p. 279-299.
- Brocca, L., Melone F., Moramarco T., Wagner W., and Hasenauer S. , 2010: ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy, *Remote Sensing of Environment*, v. 114(11), p. 2745-2755.
- Chavez P.S.J., 1996: Image-based Atmospheric Corrections-Revisited and Improved, *Photogrammetric Engineering and Remote Sensing of Environment*, v. 62, p. 1025-1036.
- Cohen W. B. and Goward S. N., 2004: Landsat's role in ecological applications of remote sensing, *BioScience*, v. 54, p. 535-545.
- Curran P. J., 1994: Attempts to drive ecosystem simulation models at local to regional scales. In G. M. Foody, & P. J. Curran (Eds.), *Environmental Remote Sensing of Regional to Global Scales*, Chichester: Wiley, p. 149– 166.
- Dorigo W. A., Scipal K., Parinussa R. M., Liu Y. Y., Wagner W., de Jeu R. A. M., and Naeimi V., 2010: Error characterisation of global active and passive microwave soil moisture datasets, *Hydrology and Earth System Sciences*, 14(12), p. 2605-2616.
- enviroGRIDS 1st Periodic Report (2010).
- enviroGRIDS D2.1 Interoperability Guideline (2009).
- enviroGRIDS D2.2 Data Storage Guidelines (2009).
- enviroGRIDS D2.4 Remote Sensing Data Use and Integration Guideline (2009).
- enviroGRIDS D2.6 Gap Analysis Report (Phase I) (2010).
- enviroGRIDS D2.7 GRID Services Supporting Massive Data Management (2010).
- enviroGRIDS D4.1 Database of useful data for SWAT modelling and report on data availability and quality for hydrological modelling and water quality modelling in the Black Sea Catchments (2010).
- enviroGRIDS D6.1 Requirements and specifications for the development of BSC-OS Portal (2010).
- Foody G. M., 2002: Status of land cover classification accuracy assessment, *Remote Sensing of Environment*, v. 80, p. 185-201.
- Foody G. M., 2008: Harshness in image classification accuracy assessment, *International Journal of Remote Sensing*, V. 29, no. 11, p. 3137-3158.
- Gao, Y., and Long D., 2008: Intercomparison of remote sensing-based models for estimation of evapotranspiration and accuracy assessment based on SWAT, *Hydrological Processes*, v. 22 (25), p. 4850-4869.
- Gavier-Pizarro G., Kuemmerle T., Hoyos L., Stewart S., Huebner C., Keuler N., Radeloff C., 2012: Monitoring the invasion of an exotic tree (*Ligustrum lucidum*) from 1983 to 2006 with Landsat TM/ETM satellite data and Support Vector Machines in Córdoba, Argentina, *Remote Sensing of Environment*, Available online 15 February 2012.
- Giuliani G., Ray N., Lehmann A., 2010: Grid-enabled Spatial Data Infrastructure for environmental sciences: challenges and opportunities, *Future Generation Computer Systems*, v. 27(3), p.292-303.
- Gorgan D., Bacu V., Rodila D., Pop Fl., Petcu D., 2009: Experiments on ESIP - Environment oriented Satellite Data Processing Platform, SEE-GRID-SCI User Forum, Bogazici University, Istanbul, Turkey, December 9-10, ISBN: 978-975-403-510-0, p. 157-166.
- Gorgan D., Ștefănuț T., Băcu V., Mihon D., and Rodilă. D., 2010, Grid based Environment Application DevelopmentMethodology, *Lecture Notes in Computer Science*, p. 499-506.



- Gvilava M., Bakuradze T. and Gigineishvili A., 2011: Pilot Testing the Relevance of the National ICZM Strategy for the Local Coastal Planning in Georgia. Published in the Abstracts of the Black Sea Outlook 2011, 3rd Bi-annual BS Scientific Conference, 1-4 November 2011, Odessa, Ukraine (http://81.8.63.74/Downloads/3BSCConf/3BSC_Conference_Book_of_Abstracts.pdf, see poster at http://81.8.63.74/Downloads/3BSCConf/Posters/20111102_ECBSea_ICZM_Strategy_vs_Tskaltsminda_Plan_BS-Outlook_3_Odessa_Ukraine_MGvilava.pdf, accessed March 31, 2012). Submitted in paper form for publication in the Turkish Journal of Fisheries and Aquatic Sciences (TrJFAS).
- Huete A., Justice C. and Liu H., 1994: Development of vegetation and soil indices for MODIS-EOS, *Remote Sensing of Environment*, v. 49, p. 224– 234.
- Immerzeel W. W. and Droogers P., 2008: Calibration of a distributed hydrological model based on satellite evapotranspiration, *Journal of Hydrology*, v. 349(3–4), p. 411-424.
- Immerzeel W. W., Gaur A., and Zwart S. J., 2008: Integrating remote sensing and a process-based hydrological model to evaluate water use and productivity in a south Indian catchment, *Agricultural Water Management*, v.95(1), p.11-24.
- JV Group (Nurol and SK E&C) and ENCON Environmental Consultancy Co., 2010: Namakhvani HPP Cascade Project Environmental and Social Impact Assessment Report, Draft, October, 2010 (http://aarhus.ge/uploaded_files/f3571b588794da1fe6ed1022a6481a4ed777f9b9c37a07cc2e6f1a1ea0eb561e.pdf, accessed March 31, 2012).
- Kalma J. D., McVicar T. R., and McCabe M. F., 2008: Estimating Land Surface Evaporation: A Review of Methods Using Remotely Sensed Surface Temperature Data, *Surveys in Geophysics*, v. 29(4-5), p. 421-469.
- Krajewski W. F., and Smith J. A., 2002: Radar hydrology: rainfall estimation, *Advances in Water Resources*, 25(8–12), p. 1387-1394.
- Lathrop R. G., 1988: The integration of remote sensing and geographic information systems for Great Lakes water quality monitoring, PhD Thesis, University of Wisconsin, USA.
- Lehmann A., Ray N., Kideys A., Weller P. and Gvilava M., 2009: Connecting the EnviroGRIDS Black Sea Catchment Observation System to Integrated Coastal Zone Management, The Proceedings of the MEDCOAST 2009, 9th International Conference on the Mediterranean Coastal Environment, 10-14 November 2009, Sochi, Russian Federation.
- Li F., Crow W. T., and Kustas W. P., 2010: Towards the estimation root-zone soil moisture via the simultaneous assimilation of thermal and microwave soil moisture retrievals, *Advances in Water Resources*, v. 33(2), p. 201-214.
- Liang S., 2001: Atmospheric Correction of Landsat ETM+ Land Surface Imagery— Part I: Methods, *IEEE Transactions On Geoscience And Remote Sensing*, v. 39 (11), p. 2490-2498.
- Liang S., 2004: *Quantitative Remote Sensing of Land Surfaces*, John Wiley and Sons, New Jersey, USA.
- Loew A. and Schlenz F., 2011: A dynamic approach for evaluating coarse scale satellite soil moisture products, *Hydrology and Earth System Sciences*, v. 15(1), p.75-90.
- Mannaerts C. M., Maathuis B.H.P., Wang L., Schouwenburg M. and Retsios B. V., 2011: Real time satellite observation of large basins using the ILWIS Open GEONETCast Toolbox, in the Extended Abstracts of the OpenWater Symposium, 19th April, 2011, UNESCO-IHE, Delft, The Netherlands (http://hikm.ihe.nl/openwater_eg/Pdf/Abstracts-OpenWater-final-for-print.pdf, accessed (March 31, 2012).
- Mihon D., Băcu V., Ștefănuț T. and Gorgan, D., 2010: Considerations on the Grid Oriented Environmental Application Development Framework, *IEEE 6th International Conference on Intelligent Computer Communication and Processing*, p. 419-426.
- Milewski A., Sultan M., Jayaprakash S. M., Balekai R., and Becker R., 2009a: RESDEM, a tool for integrating temporal remote sensing data for use in hydrogeologic investigations, *Computers & Geosciences*, v. 35(10), p. 2001-2010.

- Milewski A., Sultan M., Yan E., Becker R., Abdeldayem A., Soliman F., and Gelil K. A., 2009b: A remote sensing solution for estimating runoff and recharge in arid environments, *Journal of Hydrology*, 373(1–2), p. 1-14.
- Moran M. S., Jackson R. D., Slater P. N. and Teillet P. M., 1992: Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output, *Remote Sensing of Environment*, v.41, p. 169– 184.
- Morissette J. T., Privette J. L. and Justice C. O., 2002: A framework for the validation of MODIS Land products, *Remote Sensing of Environment*, v. 83(1–2), p. 77-96.
- Mu Q., Zhao M. and Running S. W., 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing of Environment*, v. 115, p.19.
- NASA The Warehouse Inventory Search Tool (WIST): <https://wist-ops.echo.nasa.gov/api/>
- Neteler M., Mitasova. H., 2004: *Open Source GIS: A GRASS GIS Approach*, Kluwer Academic Publishers/Springer, Boston, Dordrecht, London.
- Parkinson C. L., Greenstone R., 2000: *EOS Data Products Handbook Volume 2*, NASA Goddard Space Flight Center, p. 256.
- Rouse J.W., Haas R.H., Schell J.A. and Deering D.W., 1973: Monitoring Vegetation Systems in the Great Plains with ERTS, *Third ERTS Symposium*, NASA SP-351 I: 309-317.
- Roy D. P., Borak J. S., Devadiga S., Wolfe R. E., Zheng M. and Descloitres J., 2002: The MODIS Land product quality assessment approach, *Remote Sensing of Environment*, v. 83(1–2), p. 62-76.
- Roy P.S., Ravan S.A., 1996: Biomass estimation using satellite remote-sensing data – an investigation on possible approaches for natural forest, *J. Bioscience*, v. 21, p535–561.
- Sánchez N., Martínez-Fernández J., Calera A., Torres E. and Pérez-Gutiérrez C., 2010: Combining remote sensing and in situ soil moisture data for the application and validation of a distributed water balance model (HIDROMORE), *Agricultural Water Management*, v. 98(1), p. 69-78.
- Schmidt K. S., Skidmore A. K., 2003: Spectral discrimination of vegetation types in a coastal wetland, *Remote Sensing of Environment*, v. 85 (1), p. 92-108.
- Schmugge T. J., Kustas W. P., Ritchie J. C., Jackson T. J. and Rango A., 2002: Remote sensing in hydrology, *Advances in Water Resources*, v. 25(8–12), p. 1367-1385.
- Song C., Woodcock C. E., Seto K., Lenney M. P., Macomber S., 2001: Classification and change detection using Landsat TM data when and how to correct atmospheric effects, *Remote Sensing of Environment*, v. 75, p. 230–244.
- Stefanova A., 2011: The use of remotely sensed data for validating a SWAT model: Application on the Vit River Basin, Bulgaria, Master thesis, UNESCO-IHE Institute for Water Education, p. 92.
- Townshend J. R. G., Justice C. O., Gurney C., McManus J., 1992: The effect of mis-registration on the detection of vegetation change, *IEEE Transactions on Geosciences and Remote Sensing*, v. 30, p. 1054–1060.
- Turkish Statistical Institute, 2011: <http://www.tuik.gov.tr/> (visited in 20th March, 2012)
- UNDP Project Proposal to Adaptation Fund: Developing Climate Resilient Flood and Flash Flood Management Practices to Protect Vulnerable Communities of Georgia, (PIMS: 4583, ATLAS IDS: GEO10, PROPOSAL ID: 00060698; PROJECT ID: 00076540), USD5,316,500 ([http://adaptation-fund.org/sites/default/files/Georgia Project Proposal final afb16_0.pdf](http://adaptation-fund.org/sites/default/files/Georgia%20Project%20Proposal%20final%20afb16_0.pdf), accessed March 31, 2012)
- Urushadze T., 1999: Editor-Chief, *Map of Georgian Soils (1:500.000)*, Tbilisi, Cartographical Factory, in Georgian (http://www.agromarket.ge/soil_maps/data/georgia_soil_map.pdf, accessed March 31, 2012).
- USAID Project, Integrated Natural Resources Management in Watersheds of Georgia (INRMW): Rioni, Alazani & Iori River Basins, 2010-2014 (<http://www.globalwaters.net/georgia2.htm>, accessed March 31, 2012).

- Wardlow B. D. and Egbert S. L., 2008: Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains, *Remote Sensing of Environment*, v. 112(3), p. 1096-1116.
- Xie X., and Zhang D., 2010: Data assimilation for distributed hydrological catchment modelling via ensemble Kalman filter, *Advances in Water Resources*, v. 33(6), p. 678-690.
- Yan E., Milewski A., Sultan M., Abdeldayem A., Soliman F. and Gelil K. A., 2010: Remote-Sensing Based Approach to Improve Regional Estimation of Renewable Water Resources for Sustainable Development, *US-Egypt Workshop on Space Technology and Geo-information for Sustainable Development*, Cairo, Egypt.

Appendix 1 Rioni Hydro-Meteorology

This part of the document was compiled based of the data and information of the National Environmental Agency of Georgia (see Namakhvani ESIA, 2010).

Hydrology

River Rioni takes its source on southern slopes of the Greater Caucasus from Patsi Mountain at 2620 elevation and joins the Black Sea near city of Poti.

The length of the river is 327 km, the area of catchment basin – 13,400 km², average elevation of the catchment is 1,084 m, average inclination – 7.2 ‰. Main tributaries of the river are: Jejora (length 50 km), Kvirila (length 140 km), Khanistskali (length 57 km), Tskhenistskali (length 176 km), Nogela (length 59 km), Tekhuri (length 101 km), Tsivi (length 60 km), there are 8 tributaries with length 20-25 km, 14 tributaries – 10-25 km, remaining 355 tributaries have length less than 10 km. Combined length of tributaries is 720 km. Hydrographic network of the basin is more developed on the left side of the river (with density of the network 1.04 km/km², while density on the right side is – 0.99 km/km²).

Average multi-annual discharge of the River Rioni at the mouth is – 424 m³/s, module of the discharge – 31.6 l/s km², which increases with the elevation and near the source reaches 46.0 l/s km².

Upstream the city of Kutaisi the river is of mountainous character, flowing in narrow, deep gorge, while downstream Kutaisi, on Kolkheti Lowland, behaves as typical lowland river, meandering, branching and creating numerous oxbow lakes, so called “Narionali” (“former Rioni”).

The regime of the river used to be studied on up to 30 hydrological stations. Currently functioning are only 3 stations (see Figure 3.13 in the main text).

The river is fed by glacial, snow, precipitation and groundwater. The regime is primarily determined by precipitation and snow waters. The river is characterised by high waters in summer period due to snow melting and precipitation, while flood surges can happen anytime during the year.

High water flooding period is prolonged and continues till end of August. In certain yeas (1922) – till end of December. High water period normally starts in the beginning of April, in the mountainous part of the basin (village Alpana) – in the first half of March, in the lowland part (village Chaladidi) – by the end of February. High water reaches its maximum (3-4 meters compared to mezheni level) in May. End of September is the period of second maxima, caused by strong precipitation.

Lowest water levels are recorded in winter period (XII-II), but these patterns are frequently violated in downstream parts of the river (50-70 cm) due to precipitation.

The river is characterized with high turbidity. Near village Namakhvani it varies between 65,000 g/m³ (24.VIII.1937) to 3,100 (17.XII.1961) g/m³ ranges.

At the river section between the source and Kutaisi the elevation difference is 2,201 m and average inclination is 13.0 ‰.

Hydrologic Stations of Rioni River Basin

The map on Figure 3.13 depicts hydrologic stations (see list below) of the Rioni River Basin, showing both currently operational (filled symbol) ones, as well as those operating in the past (outlined symbol). The hydrologic stations are measuring water level, temperature and turbidity. The following is the list of hydrologic stations operational in Rioni River Basin at various periods:

01. R. Rioni – Oni: 1935-38, 1940-92
02. R. Rioni – v. Khidikari: 1932-47, 1950-87
03. R. Rioni – v. Alpana: 1919, 1927-2009
04. R. Rioni – v. Namakhvani: 1961-2005
05. R. Rioni – Gumathesi: 1962-92
06. R. Rioni – Rionhesi upstream the dam:
07. R. Rioni – Rionhesi:
08. R. Rioni – v. Chaladidi: 1928-77, 1978-2009
09. R. Jejora – v. Ipileti: 1947-94
10. R. Lajanuri – v. Alpana: 1930-31, 1934-35, 1938-58
11. R. Kvirila – Sachkhere: 1935, 1937-39, 1967-2009
12. R. Kvirila – Zestaponi: 1930-2009
13. R. Dzirula – v. Tseva: 1932-92
14. R. Khanistskali – Baghdati: 1936-92
15. R. Tskhenistskali – v. Luji: 1932-43, 1947-51, 1953-2009
16. R. Tskhenistskali – v. Khidi: 1960-92
17. R. Tekhuri – v. Nakalakevi: 1937-42, 1944-2009

Description of Hydrologic Stations along Rioni River

In this section are presented hydrologic stations operating along the River Rioni (except Oni).

No. 03. R. Rioni – v. Alpana (coordinates – Lat 42°33'39.94"N, Lon 42°50'57.50"E). The measuring point ("Sagushago" in Georgian) is located 0.3 km downstream the mouth of the River Lajanuri, right tributary of Rioni. The river gorge near the measuring point is wide. Right slope with height of 200-250 m is inclined towards the river with 20-30°, land is covered with shrubs. The road to Kutaisi is passing along the slope. There are dwelling houses on the right bank side, while left bank is utilised for agriculture.

The river bed shape is straight, built by gravel. The banks of the river are low, and represent the continuation of the river gorge.

The measuring station is located on the right bank and is equipped with water measuring level. Zero of the measuring station is 363.33 m (Baltic system). Hydrological cross-section #1 is located 47 m downstream the station and is equipped with remotely controlled hydrometric unit GR-70. The water suspended sediment measuring samples are taken at 44 m distance downstream with constant reference starting single point methodology, using bathometer-bottle. Water temperature is measured near the station – sampling performed from the bank.

The inclination of the water surface is determined with the IV class levelling instrument along the 100 m distance.

No 04. R. Rioni – v. Namakhvani (coordinates – Lat 42°24'57.50"N, Lon 42°41'43.90"E). The measuring point is located 0.3 km downstream of the v. Namakhvani.

The river gorge near the measuring point is wide, with high slopes covered with broadleaf forests and shrubs, and with terraces. Right terrace at 10-12 during low waters serves as the road to Kutaisi-Oni direction. The river is flowing below the right slopes of the gorge.

The river bed shape is straight, mostly built by fine-size gravel, slightly deformable, the left bank is low, 70-80 m wide, covered with scrubs, while right bank is the continuation of river gorge slopes.

The measuring station is located on the right bank and is equipped with water measuring level and automatic water level measuring unit "Valdai". Automatic measuring unit is of island type. Zero of the measuring station is 219.61 m (Baltic system). Hydrological cross-section #1 is located 87 m downstream the station and is equipped with hydrometric unit GR-64.

The water suspended sediment measuring samples are taken at hydrological cross-section #1 in 24 m downstream from constant reference point, with integration method, using bathometer-bottle with loads.

Water temperature is measured near the station – sampling performed from the bank.

The inclination of the water surface is determined with the IV class levelling instrument along the 70 m distance.

No 05. R. Rioni – Gumati HPP (coordinates – 42° 16', 42° 42'). HPP is located 7 km upstream from Kutaisi and 12 km downstream from Namakhvani measuring station of r. Rioni. The measuring point is located 0.3 km downstream of the v. Namakhvani. HPP is located in front of the dam. Dam is gravitational, concrete, right part of it is unregulated with 74.6 length, regulated (71.0 m, with HPP block) and again unregulated left bank 700 m part with, with length 65.9 m. overall length of the dam is 281.5 m. The regulated part of the dam has four discharge points, each with size 14x8 m, collecting lock of "Glagol" type, discharging tunnel with capacity 105 m/s.

The highest discharge level of the dam is 192.00 m (Baltic system). Total volume of the reservoir is 39 mills. m³, useful volume - 19 mills. m³. Average depth is fluctuating in the range of 6-24 m. With water level reaching 200.20 m (Baltic system), flooded area is spreading for 3 km.

Downstream the dam river is flowing in the wide gorge. The gorge is formed by bedrock and gravel. The HPP is equipped with 4 rotating wing type turbines. The HPP is connected to the national grid and maintains maximal power capacity during summer high waters and minimal – during low water. The power generation is engaged continuously, day and night. The turbines are equipped with "VO-2 Ogres" type discharge measuring units, but they are not functioning. The regime measurements are performed with water measuring levels installed upstream and downstream of the dam.

Zeros of both levels are 196.00 m (Baltic system).

No 08. R. Rioni – Chaladidi (coordinates – Lat 42°12'17.68"N, Lon 41°56'43.29"E). Hydrologic station is located near village Sakochakidze 7 km from railway station Kvaloni and 3 km south of Tsivi river mouth.

Near the station river is flowing on wide area of Kolkheti lowland, on the sediment cone deposited by river itself. To protect adjacent lands from flooding both banks of the river are diked. Adjacent areas are intersected by multiple drainage ditches. These areas are either used for agriculture, or covered by wetlands, shrubs and re-grown deciduous forests.

The river bed is here straight, silted, slightly deformed. The banks are elevated 3-5 meters above the river, loosely bound erodible, partly covered with shrubs.

The station is located on the right bank and is equipped with water measuring level. The zero of the station is at 2.08 m (Baltic system). Hydrological cross-section #1 is combined with the station itself. The flow rates are not being measured since 1983 as remotely controlled hydrometric unit was out of order. The water turbidity sampling is performed with single point methodology; from the primitive cable ferry installed

across the river, some 100 meters from the bank and 50 m downstream the cross-section #1. The water temperature is measured at the station – on the right bank of the river.

Existing Reservoirs

Gumati HPP. Located in Tskaltubo District, 6 km north of Kutaisi, in the narrow gorge of the middle reaches of Rioni. The reservoir, which is formed by the 30 m reinforced concrete dam, regulates River Rioni discharges for Gumatesi HPP (I and II). The dam is composed of middle section with discharge regulation and two side parts, with combined length of – 216 m. The maximal designed water depth at the dam is – 24 m. The reservoir has narrow prolonged oval shape (up to 10 km in length). Its width is varying between 60-550 m ranges. The volume of the reservoir for the normal low level – is 2.4 km². The reservoir is spreading along the meridian direction. The topography is mostly of erosive origin, caused by the action of Rioni and its smaller tributaries. Accumulations are formed by terraces of Rioni and debris cones of its tributaries. There are frequent side by side occurrence of landslide bodies, both of old and recent origin.

The reservoir entered into operation in 1956 and is used for energy generation. The reservoir requires constant regulation on a daily basis. In case of necessity weekly regulation is also possible.

Lajanuri HPP. Located in Tsageri District downstream the River Lajanuri (right tributary of Rioni), between vv. Alpna and Orbeli. The reservoir is surrounded by high southern branches of Caucasus Chain Mountains. The reservoir is formed by high (69 m), concrete arch dam, installed on the narrow section of the river near v. Surmushi. The reservoir has prolonged oval shape. Its length is 3.2 km, average width – 0.5 km, maximal width – 1.3 km, average depths – 30 m, length of the dam is 127 m. The volumes of the reservoir under normal low level conditions are: total – 25.0 mills. m³, useful volume – 17.0 m³, surface area (normal low level) is 1.6 m². In the downstream part the reservoir is joined from opposing slopes by two small streams (Kheleshurisghele, Usakheloghele). Near the junction of streams the reservoir width narrows considerably and transforms into canyon. In this section reservoir has the maximal depth of 70 m.

The accumulation of bottom sediments in the reservoir is noteworthy. In the period after its construction the bottom sediment layers are accumulating in the lower reaches of the reservoir and near the dam the depth of the sedimentation layer reaches 10 m. The survey conducted in 60-s show that bed of the reservoir undergone significant changes. The landslide phenomena are apparent in south-western banks of the reservoir and on its southern end between the river/stream Kheleshurisghele and limestone layers.

River Lajanuri near Orbeli approaches River Tskhenistskali to a distance less than 5 km. Since the construction of 69 m arch dam in the narrow limestone canyon near Orbeli, the discharge rates of Rioni were increased by 35 m³/s due to tunnel diverting waters of Tskhenistskali into Lajanuri, and Rioni.

The reservoir entered into operation in 1960. It is regulated on daily basis and it is used for energy generation.

Vartsikhe HPPs Cascade. Vartsikhe I, II, IV & IV HPPs are located on the river Rioni on 27 km long section from village Vartsikhe till the source of river Gubistskhali. The dam of the reservoir is located on the territory near the river Rioni at the tributary of the rivers Kvirila and river Khanistskhali. The normal water level of the reservoir is 86.77 m. Vartsikhe I HPP was put into operation in 1976, Vartsikhe II HPP in 1978, Vartsikhe IV HPP in 1980 and Vartsikhe IV HPP in 1987.

Rioni HPP is a derivation type HPP utilizing hydropower of River Rioni and River Tskhenistskali. It is located near Kutaisi city. The plant is equipped with four 12 MW units and is expected to produce 325 million kWh, annually. The normal water level of the reservoir is 158.0 m with a total volume of 5.0 million m³ and useful volume of 3.0 million m³.

Shaori HPP is located on River Didi Chala. The plant is equipped with four 9.6 MW units resulting in a total installed capacity of 38.4 MW. The expected annual energy generation is 114 million kWh.

Planned Reservoirs

Proposed Namakhvani HPP Cascade (Tvishi, Namakhvani and Zhoneti Reservoirs). Government of Georgia (Ministry of Energy) announced this project as part of the new Investment Program in Greenfield Hydropower Plants aiming at attracting private investments into development of hydropower in Georgia.

Project developed was awarded through BOO scheme to Joint Venture (JV) Group consisting of Nurol Energy Production and Marketing Inc. (Turkey), KEPCO (Korea Electric Power Corp) and SK Engineering & Construction Co. Ltd (Korea). JV Group has mandated STUCKY Ltd. (Switzerland, with its subsidiaries in Turkey and Georgia) as Technical Advisor for a review and update of the feasibility study of the Namakhvani hydropower cascade project with the total installed capacity of 450 MW.

The proposed Namakhvani HPP Cascade is to be composed of three reservoirs: Tvishi, Namakhvani and Zhoneti HPPs that are located on the Rioni River between existing Lajanuri and Gumati HPPs, along the middle and upper reaches of the river. The project is proposed to utilize the potential energy of Rioni River between elevations of 360.0 m and 200.0 m above sea level. The closest city to the project area is Kutaisi. Construction is to be completed in 6 years with starting planned in 2012.

Proposed Alpana HPP. In addition to Namakhvani HPP Cascade reservoirs, Alpana HPP is planned to be constructed in the immediate upstream of Tvishi reservoir. The location of the project is upstream of Namakhvani HPP Cascade Project and approximately 11 km East of Alpana village. The regulator is planned to be 10 m high with a normal water elevation of 438 m ASL. The water will be diverted with an approximately 8.3 km length diversion tunnel from the reservoir to Alpana HPP (at 367 m elevation) near to the Alpana Village. The installed capacity of the plant would be 70.6 MW. The reservoir will inundate an area of approximately 35 ha where 13 ha of this total area is existing river bed. The rest of this total area (22 ha) land is covered with riparian vegetation, agricultural lands and forests.

Metadata

Observation parameters for existing hydrometric stations (see Figure 3.13) include: (i) daily discharge rates (estimated rates where only water level measurement are available); (ii) station descriptions (physical-geographical and climatic characteristics, elevation, latitude, longitude) (iii) turbidity.

Meteorology

Description of Meteorology Stations along Rioni River

1. Oni Meteorological Station

Elevation: 788.9 m
Latitude: 42°34'54.22"N
Longitude: 43°25'59.96"E

Physical-geographical description. Oni is located in the upstream part of the Rioni River and spreads on both of its banks. Rioni gorge near Oni is spreading from north-east to south-west direction. Branches of Lechkumi mountain chain, characterized with high elevations and steep slopes, are descending to right bank of the river, while left bank is formed by Racha mountain range with lower elevation. Mountain slopes are covered with deciduous and partly coniferous forests, with gardens, vineyards and agricultural crops underneath.

Meteorological station is located in the outskirts of Oni, to the south-west, within 500 meters from the river. Rioni changes its flow at the meteorological station from latitudinal to south-western direction. Within the 200-300 meter distance the station is surrounded by smaller houses with adjacent gardens.

Soils on the flat areas along the gorge are of carbonate nature, while on the mountains black and grey soils dominate. Groundwater is below 3 meters.

Climate. The area is characterized by temperate humid climate. Summer is moderately warm, winter is cold. Vertical zoning of the climate and inversion of precipitation is clearly manifested here. Increase of the precipitation curve drops at higher 2,500-2,800 m elevations. Phonic effect is strongly manifested, accompanied by easterly winds. Share of the snow in precipitation increases with elevation and in higher mountains reaches 70-80%. Duration of sustained snow cover reaches 60-180 days. Precipitation distribution rates are higher in autumn-winter period. Total annual precipitation fluctuates in the 900-1200 range. Average January temperature is -1 to -6°, in higher mountains dropping to -12°, July – in the range of 7-20°, average annual – from 2 to 6°. Profile of the climate is mountainous, with cold winter and moderate summer.

2. Kutaisi Meteorological Station

Elevation: 113 m
Latitude: 42°14'45.62"N
Longitude: 42°37'28.91"E

Physical-geographical description. The meteorological station is located on right bank of Rioni River, outside the Kutaisi, near the airport runway. This area is relatively elevated over the Rioni Gorge and entire Kolkheti Lowland. The city of Kutaisi from the south is bordered by the Imereti foothill topography. The Rioni is leaving the mountainous gorge near this city and from longitudinal meridian direction is turning into latitudinal direction. There are not forest areas near the city. The area is mostly occupied by agricultural cultivations, fruit gardens and wine yards.

Soils are highly inhomogeneous. Higher elevations are red spoils, weakly or strongly non-nutritious. At meteorological station soils are of carbonate type. Groundwater is below the 3 m depth.

Climate. This region is characterised by subtropical climate with certain deficit in humidity. Central parts of the region are moderately humid, while foothill areas are strongly humid. The precipitation inversion takes place along the Rioni Gorge. Breezes and sea influences penetrate along the Rioni. Phonic effect is strongly manifested, accompanied by easterly winds, frequently of storm character. Spring and beginning of summer are drier, while high humidity conditions are established in autumn-winter period. Annual precipitation rates are 1,400-2,200 mm, average January temperature is 4-5°, July – 21-23°, average annual – 12-14°. Profile of the climate is low humid subtropical with positive temperature balance. Permanent snow cover is not observed.

3. Poti Meteorological Station

Elevation: 1 m
Latitude: 42°08'35.71"N
Longitude: 41°41'21.08"E

Physical-geographical description. City of Poti is located on the eastern coast of the Black Sea, to the west of the Kolkheti Lowland, along the old mouth of the River Rioni. Poti is bound from the east by Lake Paliastomi. Stream Kaparchina is flowing out of the lake, which outflows first to north-west direction, then turns into south-west and joins the Sea crossing the system of coastal dunes. In 1933 the sand bank lying between the sea and the lake was breached and lake is now connected to the sea with deep and wide (140-160 m) channel. The banks of Paliastomi Lake are low-lying and are covered with vegetation. Sea shoreline is quite uniform near Poti. Narrow sandy beach along the coast is some 100 m wide. Landward elevations are sometimes negative and create wetland conditions. The Rioni gorge here is of the widespread lowland type, parts of which are covered by wetlands. Some lands are used for agriculture production.

Soils are wet/wetland, some adjacent parts are peatlands.

Groundwater is at 0.5-1.0 m depths.

Climate. The region under consideration is humid subtropical, characterized by elevated humidity and high thermal background. Winters are warm, summers are moderately hot. Air circulation in the region is

governed by breezes. Relatively dry period is spring; more humid is autumn-winter. Total annual precipitation varies between 1,700-2,700 mm ranges. In adjacent mountainous annual total reaches 3,500-4,000 mm range. Persistent layer of snow forms only in mountainous areas. January average temperature is 5-7°, July – 22-24°, average annual – 13-15°.

Metadata

Presented below is metadata on observation parameters at Kutaisi and Poti meteorological stations (Oni meteorological station data is available only till 2000, afterwards the station was closed):

Observation dates and times are according to GMT at 15, 18, 21, 00, 03, 06, 09, and 12.

Station Description –	Physical-geographical and climatic characteristics (elevation, latitude, longitude)
Air Temperature –	Temperature at set times every 3 hours (°C); Daily Absolute Maximal Temperature (°C); Daily Absolute Minimal Temperature (°C);
Relative Humidity –	Measurement values (%) at set times every 3 hours;
Wind speed –	Average Wind Speed (m/s); Daily Maximal Wind Speed (m/s); Wind Direction (degrees); Wind rose;
Cloudiness –	Measurement values (points) at set times every 3 hours;
Cloud altitude –	Measurement values (meters) at set times every 3 hours;
Precipitation –	Daily precipitation (mm).