

# Politecnico di Milano

---

Dipartimento di Ingegneria Idraulica,  
Ambientale, Infrastrutture Viarie,  
Rilevamento (DIIAR)

*Dottorato di ricerca in Geodesia e Geomatica*



## **A GIS embedded approach for Free & Open Source Hydrological Modelling**

PhD dissertation  
Massimiliano Cannata

Coordinatore:  
Prof. Fernando Sansò  
Advisor:  
Prof. Maria Antonia Brovelli  
Co-advisor:  
Prof. Helena Mitasova



# Abstract

The aim of this research is to demonstrate how watershed management can be achieved by using Free and Open Source Software for Geomatic (FOSS4G) through hydrological modelling.

This work was carried on in the frame of an interreg IIIA project between Italy and Switzerland named “Sviluppo di un sistema di gestione dei rischi idrogeologici nell’area del lago Maggiore” (“Development of a management system for the hydro-geological risk in the Lake Maggiore area”) partially funded by the Europe Union.

For this purpose a new hydrological model called “HydroFOSS” was developed.

HydroFOSS is:

- distributed – the hydrological variables are continuously described in the space.
- Physically based – all the involved variables have a physical meaning.
- Continuous – operates over an extended period of time, determining flow rates and conditions during both runoff periods and periods of no surface runoff.
- Modular – is a combination of different modules describing the processes involved in the rainfall-runoff process.
- GIS embedded – is fully developed into a GIS system by using the GIS's commands and library functions.
- Open Source – developed by using exclusively Free and Open Source Software.

Due to the specific area of study (the Lake Maggiore is situated across Italy and Switzerland, in the Alpine region) the processes considered in the hydrological model development are the solar radiation, the evapotranspiration, the snowmelt and accumulation, the canopy interception, and the runoff.

The heterogeneity of the needed data, in terms of formats, topology, coordinate systems, and time-spatial resolutions bring us to the

development of a geodatabase that considers the time component. For this task a methodology to handle raster series has been developed.

Once the data have been organized in the geodatabase, the further required step is the data processing for model input generation. This task involves either the spatialization of numerous variables and the validation of different data, two cases were deeply investigated (the validation of the meteorological radar rainfall observations, and the best interpolation technique for temperature meteorological station observations) while, due to time restrictions, a standard approach has been followed in other cases.

The chosen processes were then simulated by developing specific new commands in the GIS GRASS and the overall model was setted up by means of a general script that automatically executes all the required operations.

Finally a link between the HydroFOSS model and the automatic inverse calibration model UCODE-2005 (Poeter et al., 2005) was generated and a case study application was successfully carried on.

This solution has shown how a fully open access, both in term of cost and in term of control, to all the modelled processes and data can be achieved by using Open Source Software and a modular approach. Moreover the usage of a GIS allows the management of heterogeneous data and helps models integration because of its intrinsic data exchange, analysis and visualization capabilities.

# TABLE OF CONTENTS

<b>1. Introduction</b>	p. 1.1
1.1 Problem Identification	p. 1.1
1.2 Objective and main steps of the research	p. 1.2
<b>2. GIS and hydrological models: an overview</b>	p. 2.1
2.1 Geographical Information Systems	p. 2.1
2.1.1 The evolution of GIS	p. 2.1
2.1.2 GIS softwares	p. 2.6
2.2 Hydrological models	p. 2.10
2.2.1 Hydrological model classification	p. 2.10
2.2.2 Historical review of Hydrological Models	p. 2.15
2.2.3 Principal characteristics of some of the currently used distributed watershed models	p. 2.18
2.3 GIS for water resources	p. 2.21
2.3.1 GIS applications in hydrology	p. 2.22
2.3.2 Linking GIS and hydrological model	p. 2.25
<b>3. Methods</b>	p. 3.1
3.1 Chosen approaches	p. 3.1
3.1.1 GIS-model integration level	p. 3.1
3.1.2 GIS package	p. 3.2
3.1.3 Hydrological Model type	p. 3.3
3.2 Rainfall-runoff conceptual model	p. 3.4
3.2.1 Hydrological physical model	p. 3.4
3.2.2 Hydrological logic model	p. 3.5
3.3 Rainfall-runoff mathematical model	p. 3.7
3.3.1 Evapotranspiration	p. 3.7
3.3.2 Canopy interception	p. 3.9
3.3.3 Snow melt and accumulation	p. 3.10
3.3.4 Flow component	p. 3.11

## Table of contents

3.4 Data structure _____	p. 3.16
3.4.1 Basic input data data _____	p. 3.16
3.4.2 Geo-database logic model _____	p. 3.19
3.4.3 Geo-database conceptual model _____	p. 3.21
3.4.4 Geo-database physical model _____	p. 3.23
<b>4. System implementation _____</b>	<b>p. 4.1</b>
4.1 Catchment maps generation _____	p. 4.1
4.1.1 Selected approach _____	p. 4.1
4.1.2 Climatic map generation _____	p. 4.2
4.1.3 Watershed characterization maps generation _____	p. 4.6
4.2 Net radiation map generation _____	p. 4.8
4.2.1 Net shortwave radiation (Rns) _____	p. 4.8
4.2.2 Net longwave radiation (Rnl) _____	p. 4.11
4.2.3 Global radiation (Rn) _____	p. 4.12
4.3 Snow accumulation and melt maps generation _____	p. 4.14
4.4 Potential evapotranspiration map generation _____	p. 4.16
4.5 Canopy interception maps generation _____	p. 4.19
4.6 Rainfall runoff maps generation _____	p. 4.22
4.6.1 Initialization command: h.hydroFOSS.init _____	p. 4.24
4.6.2 Runoff computation command: h.hydroFOSS.runoff _____	p. 4.27
4.7 Automatic model running _____	p. 4.33
4.8 Sensitivity analysis and model calibration _____	p. 4.37
4.8.1 Advanced calibration method _____	p. 4.37
4.8.2 Calibration procedure implementation _____	p. 4.39
<b>5. Investigation of the radar rainfall estimates and of the temperature interpolation method _____</b>	<b>p. 5.1</b>
5.1 Radar rainfall estimates investigation _____	p. 5.1
5.1.1 State of the art in rainfall estimate _____	p. 5.1
5.1.2 Study area _____	p. 5.3

## Table of contents

5.1.3 Radar data_____	p. 5.3
5.1.4 Rainfall Data_____	p. 5.4
5.1.5 Preliminary analysis_____	p. 5.5
5.1.6 Rainfall detection analysis_____	p. 5.6
5.1.7 Rainfall intensity analysis_____	p. 5.7
5.1.8 Spatial correlation analysis_____	p. 5.7
5.1.9 Linear and functional correlation analysis_	p. 5.8
5.1.10 Z-R relationship adjustment_____	p. 5.9
5.1.11 Conclusions of the validation of radar rainfall estimates_____	p. 5.9
5.2 Investigation of interpolation methods for temperatures maps derivation_____	p. 5.13
<b>6. Case study: the Verzasca basin_____</b>	<b>p. 6.1</b>
6.1 General characteristics_____	p. 6.1
6.2 Projection and region properties_____	p. 6.3
6.3 Topography_____	p. 6.4
6.4 DTM derived maps_____	p. 6.5
6.4.1 Filled DEM_____	p. 6.5
6.4.2 Stream, basin, flow accumulation, and flow direction maps_____	p. 6.6
6.4.3 Slope map_____	p. 6.12
6.5 Landuse_____	p. 6.13
6.6 Soiltype_____	p. 6.16
6.7 Riverwidth_____	p. 6.18
6.8 Model verification and running strategy_____	p. 6.20
6.9 Climatological maps_____	p. 6.23
6.10 Watershed characterization maps_____	p. 6.29
6.11 Net radiation_____	p. 6.37
6.12 Snowmelt_____	p. 6.40
6.13 Evapotranspiration_____	p. 6.45
6.14 Canopy interception_____	p. 6.47
6.15 Runoff calibration_____	p. 6.54
6.15.1 First sensitivity analysis_____	p. 6.55
6.15.2 First model calibration_____	p. 6.65

## Table of contents

<b>7. Conclusion and future works</b> _____	p. 7.1
<b>Acknowledgements</b> _____	p. 8.1
<b>References</b> _____	p. 9.1
<b>Appendix:</b> Meteorological Network Stations Database_	p. A.1

# 1. Introduction

In the first section of this chapter the motivation that bring to the development of this research are discussed. While in second section will be illustrated the identified objective and main step that will be carried on during this thesis.

## **1.1 Problem Identification**

Recently many European countries have been affected by extreme hazard events, therefore the associated risk for human life and property damages rose with increased frequency.

Following a global trend, also the Alpine region shows an increasing rate of natural hazards, that according with a study conducted in Switzerland by WSL (Eidg. Forschungsanstalt für Wald, Schnee und Landschaft) (Schmid et al., 2004) – based on the period 1972-2002 in the Confederation territory – are subdivided as follows:

- floods (annually 60-95% of the total);
- soil slips (annually 5-35% of the total);
- debris flows (annually 0-10% of the total);

with a mean annual “costs” calculated in 2.93 dead and 0.28 billion of Swiss Francs (CHF).

Looking at these values and considering the expected growing number of natural hazards due to changes in climate (different rainfall pattern and higher intensity) and in land-cover (growing urban areas and deforestation) it is clear that institutions have to invest in prevention and mitigation strategies. These procedures are particularly important and should consider transboundary environment management strategies in order to achieve effective improvements.

Floods have been shown to be the most frequent events thus their risk management should be one of the major goals in the alpine regions.

Successful floods management requires:

- the development of common sustainable flood management strategies through land use, power plant and irrigation regulation

all over the basins (without any administrative border limitation) in order not to over-stress the land;

- the building or improving of flood warning systems and emergency management tools to reduce losses of life and property due to “unexpected” extreme hazards.

It is also evident that problems associated with water resources cannot be solved by hydrologists alone: even if an hydrological model could perfectly simulate the real circulation of water this is not enough to guarantee the correct basin management (by means of economic, social and ecological point of view).

Therefore an increasing collaboration between hydrologists, economists, ecologists, water users, and decision makers is a key point towards more sound policies in water resources (including floods) management.

The collaboration needs between different disciplines and interest groups call for more efficient sharing of data sets and methods (which may exist but not be each other compatible).

The main candidate for grouping all this disciplines and methods is certainly the Geo-Information science: its intrinsic interdisciplinary and its strong linkage with the territory makes the Geographical Information System (herein referred to as GIS) an ideal instrument for environmental management.

## ***1.2 Objective and main steps of the research.***

Nowadays GIS is assuming an important role in Management Information System (MIS), its ability in spatially describe and analyse events is a key point in successful decision making processes. It is clear that effective MIS requires both mathematical and geographical components. While mathematical component itself is most suited to detailed investigation of the particular environmental processes it is not suited to collocate these processes into their wider environmental, ecological and economic context.

Believing that GIS-based modelling allows scientists to develop complex, interactive, and flexible applications using geospatial data

to simulate and predict real-world events, the objective of this thesis is to develop a GIS-based hydrological model, able to simulate the effects of environmental policies, estimate all the hydrological variables for further analysis and evaluate/forecast the discharges.

I also believe that such a kind of model can constitute an effective tool for other water related hazard management systems.

The main steps of this research are:

1. Appropriate hydrological/GIS approach choice – Different kind of hydrological model approach are available nowadays, each one with its property and applicability. Also GIS are implemented with different software and tools, each one with its specific features and programming environment.
2. System conceptual modelling – A schema describing the considered modules and data (spatial and semantic) with the respective structure and relationships is needed. It helps the understanding of the relationship between the diverse inputs and the watershed system.
3. System implementation – How the system has to be realized and how each single block of the conceptual model has to be developed.
4. Sensitivity analysis and model calibration procedure – While sensitivity analysis is used to evaluate how a given model output depends upon the input parameters (this is an important method for checking the quality of a given model, as well as a powerful tool for checking the robustness and reliability of its analysis), model calibration consists in the research of the optimal (or sub-optimal) parameters values to improve the model fit.
5. Case study application – A real case application allows to test all the implemented procedures.

## **2. GIS and hydrological models: an overview**

### **2.1 Geographical Information Systems**

GIS is the combination of data management with content and locational indexing, spatial representation, and analysis techniques to facilitate the understanding of real-world entities in their context and interactions.

A GIS, therefore, records objects (in terms of their location and characteristics), and provides tools (analysis and display) for their modelling (spatial and semantic) and further modelling the effects of characteristics and/or location changes upon the surrounding environment.

#### **2.1.1 The evolution of GIS**

GIS technology evolved through multiple parallel but separate applications across numerous disciplines; this caused the development of a software with interdisciplinary tools. Along its evolution – going from the 1950s to now – we can detect three different phases driven by the type of applications, data and user/supplier interactions (Crain and MacDonald, 1984).

In the first phase – around the 1960s, the 1970s and the early 1980s – GIS was a technology born as an inventory tool able to assemble, organize and determine the extent of existent data. For this purpose it developed a defined data structure with primary data types (point, line, area, raster) and functions for data importing, editing, retrieving, updating, querying and reporting. The relationship between users and suppliers was of a mere customer-client type, where suppliers' duties were to ensure data integrity, long time living systems (software and hardware updates), and on demand reports and outputs. Milestones during this period are:

- 1962: the first GIS system using digital data on a computer was

## Chapter 2 – GIS and hydrological models: an overview

developed in Canada, by Alan Tomlinson, called the Canadian Geographic Information System;

- 1965: the development of the GBF/DIME (Geographic Base File / Dual Independent Map Encoding) files by the U.S. Census Bureau in the 1960s marked the large-scale adoption of digital mapping by the government. This system led to the production of the Census TIGER files, one of the most important socio-economic spatial data sets in use today.
- 1966: a grid-based mapping program called SYMAP, developed at the Laboratory for Computer Graphics and Spatial Analysis at the Harvard Graduate School of Design in, was widely distributed and served as a model for later systems (Mark at al., 1997). These early GIS packages were often written for specific applications and required the mainframe computing systems available usually in government or university settings.
- 1969: Ian McHarg in his book "Design with Nature", popularized the use of georeferenced transparent map overlays for resource planning purposes. The Environmental Systems Research Institute (ESRI) is founded by Jack Dangermond as a privately held consulting firm that specialized in land use analysis projects. Jim Meadlock establishes Intergraph Corporation (originally called M & S Computing Inc).
- 1970s, data formats begin to emerge and private vendors began offering GIS packages. M&S Computing (later Intergraph) and Environmental Systems Research Institute (ESRI) emerged as the leading vendors of GIS software (Antenucci at al., 1991).
- 1978, ESRI releases the first version of Arc/Info, the current leading GIS software package. ERDAS, provider of geographic imaging solutions, was founded by Lawrie Jordan and Bruce Rado.
- 1982, ESRI's ARC/INFO® 1.0, the first commercially available GIS software package, which ran on mainframe computers ([www.esri.com/company/about/history.html](http://www.esri.com/company/about/history.html)) was released.
- 1982, Army Corps of Engineers Construction Engineering Research Laboratory (CERL) started the development of the GIS GRASS – Geographic Resources Analysis Support System – as a

## Chapter 2 – GIS and hydrological models: an overview

tool for land management at military installation (Mitasova and Neteler, 2002).

- 1986, Laszlo Bardos, Andrew Dressel, John Haller, Mike Marvin and Sean O’Sullivan founded MapInfo. ESRI's PC ARC/INFO® 1.0, the first GIS software available for the personal computer, was released.

In the second phase – during the late 1980s and the early 1990s – the GIS was pushed to evolve towards analysis because of the users need of more complex retrievals and queries for hypotheses exploring, theory confirming or data providing for research and modelling. In this stage more functions for user interaction was developed mainly in a graphical way by a user friendly interface (GUI, Graphical User Interface). It gave to the user the ability to sort, select, extract, reclassify, reproject and display data on the basis of complex geographical, topological and statistical criteria. The suppliers increased their knowledge on existing and growing data analyses techniques, specific subject matters (e.g.: ecology and hydrology), and data context issues. While data storage were still mainly centralized the user’s access became more decentralized.

Important facts during this phase years are:

- 1987, SPANS GIS (INTRA TYDAC, 1992) and the IDRISI project (<http://www.clarklabs.org/IdrisiSoftware.asp>) both started.
- 1988, the National Centre for Geographic Information and Analysis (NCGIA, <http://www.ncgia.ucsb.edu/>) is established in the USA.
- 1991, GRASS 4.0 is released through Internet - popularity grows with universities, business and government agencies. Vector Library changed significantly from the library used in previous versions.
- 1992, ESRI released ArcView® 1.0, a desktop mapping system with a graphical user interface that marked a major improvement in usability over Arc/Info’s command-line interface (<http://www.esri.com/company/about/history.html>).
- 1992, OSIRIS system corporation released GrassWare®: the first

## Chapter 2 – GIS and hydrological models: an overview

graphical interface to the GRASS GIS software.

- 1993, Steve Putz developed PARC, the first Web-based interactive map viewer (<http://mapweb.parc.xerox.com/map>);
- 1994, with the release of its ArcView 2.0 software, ESRI took desktop mapping to an entirely new level: desktop GIS. The Open GIS Consortium aiming at developing publicly available geoprocessing specifications was founded.
- 1995, ESRI released Spatial Database Engine (SDE®), an innovative tool for storing and managing GIS data in a commercially available database management system (DBMS).

Starting from the late 1990s GIS entered in a new era. As computing power increased and hardware prices plummeted the GIS became a viable technology for state and municipal planning (Harris et al., 1993). In this third phase of evolution GIS is asked to become a real Management Information System (MIS), and thus able to support decision making processes.

Users look for answers to the “What if?” question, thus new modelling and planning facilities have to be developed inside the GIS itself. Apart from technical staff (programmers) users and suppliers have no more distinction and are both entities of the management system process. Pushed by the growing diffusion of the Internet/Intranet technology the data become decentralized and integrated control procedures are requested. In the 2000s GIS moved to Web and become more popular for geographical information dissemination.

This enhancement came together with growing resources requirement , in fact:

- software development involved multiple sciences;
- data management and quality control become more complicated due to decentralization;
- new technologies have to be tested to improve performance and usability,
- continuity during changes should be ensured in order to let existent application running.

## Chapter 2 – GIS and hydrological models: an overview

The main facts characterizing GIS in this third phase are:

- 1997, the University of Minnesota (UMN) released MapServer 1.0, an open source development environment for building spatially-enabled Internet applications (<http://mapserver.gis.umn.edu/>, 2006). ESRI released ArcView Internet Map Server (IMS), a commercial tool for publishing GIS data over the Internet.
- 1999, GRASS 5.0 is released under GNU GPL by the GRASS Development Team (<http://grass.itc.it/>, 2006). The first major change in years, this version incorporates floating point calculations and NULL support into GRASS raster engine.
- 2001, Refraction Research released PostGIS 0.1 (<http://postgis.refrations.net/>, 2006), an open source "spatially enabler" adding support for geographic objects to the PostgreSQL object-relational database.
- 2002, ESRI began offering a wide selection of GIS software compatible with the Linux operating system. In its continuing commitment to open systems, ArcIMS 4, ArcSDE 8.2, MapObjects--Java Standard Edition, and ArcExplorer 4 software are all supported on Linux.
- 2005, GRASS 6.0.0 is released with new interface, vector engine, and database support.
- 2005, Google® lunches two services: the GoogleMaps® (<http://maps.google.com>, 2006) that uses a new technology for fast navigation of Web-GIS applications (AJAX, Asynchronous JavaScript and XML) and the GoogleEarth® (<http://earth.google.com/>, 2006) that adds new appeal to the map navigation (3D fly view). They highly contribute to the diffusion of GIS concepts and capabilities across society, through and

Figure n. 1, summarize the three phases of the GIS evolution and the main activities as described above.

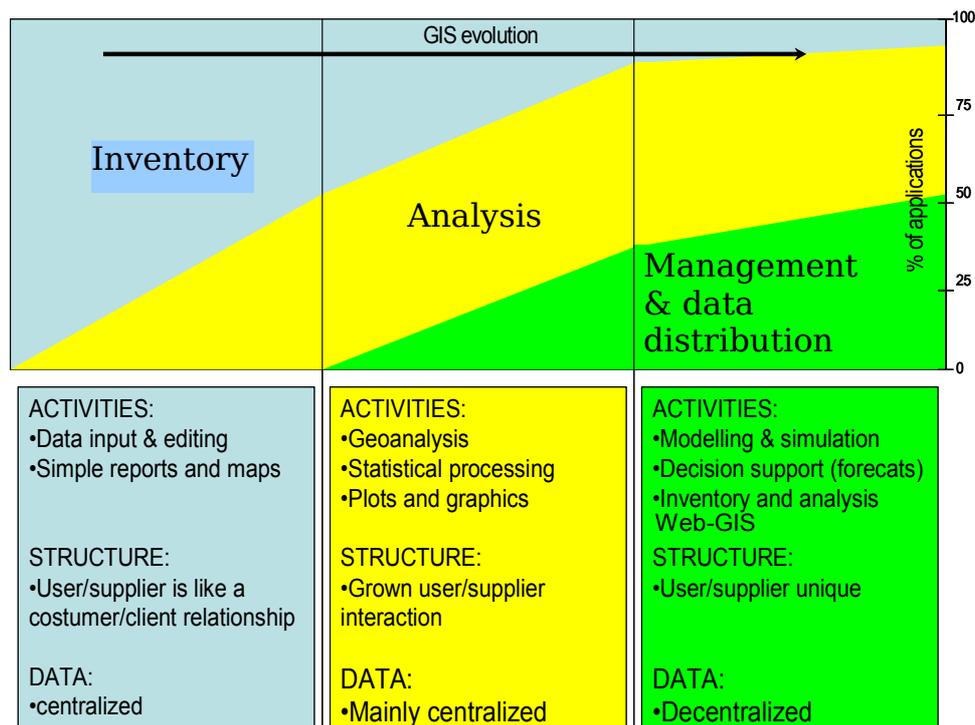


Figure 1 – GIS evolution representation after Crain and MacDonald, 1984.

Nowadays a 50% of activity are in management applications, 40% in analyses and 10% in inventory.

Today, GIS is a user friendly software and is no more used only by experts (geomatics): also geologists, ecologists, doctors and economists make extensive use of this technology (user/supplier uniqueness). It can access data held in different place, managed by different agencies (decentralization), and stored in different format. Applications are mostly in the hazard/impact assessment for sustainable management (management applications) and modelling is one of the key issues for efficient Spatial Decision Support Systems (SDSS) development.

### 2.1.2 GIS software

As illustrated in the previous paragraph, a large number of GIS software has been developed during the last decade. During their evolution often they took inspiration each others and now as a

## Chapter 2 – GIS and hydrological models: an overview

result they are quite similar in capabilities: generally they all (i) manage numerous data format, (ii) produce customizable map visualization, (iii) handle attributes, vector, raster and image data types, (iv) perform general and spatial query/analysis.

We can individuate two classes of software: the commercial or proprietary software and the Free and Open Sources Software (herein referred to as FOSS).

The first class describes software offered for sale or license, where the users are not allowed to see the source code. Nor are they able to modify the code for their own use or to distribute to others. The vendor is the only proprietary of the software and the user just acquire a licence for use it as it is.

On the contrary FOSS denote some user freedom rights such as to run, copy, distribute, study, change and improve the software. As described in the Web-Site of the the Free Software Foundation (FSF) ([www.fsf.org](http://www.fsf.org)) the word “free” refers to four kinds of user freedom:

- The freedom to run the program, for any purpose.
- The freedom to study how the program works, and adapt it to any needs.
- The freedom to redistribute copies.
- The freedom to improve the program, and release the improvements to the public, so that the whole community benefits.

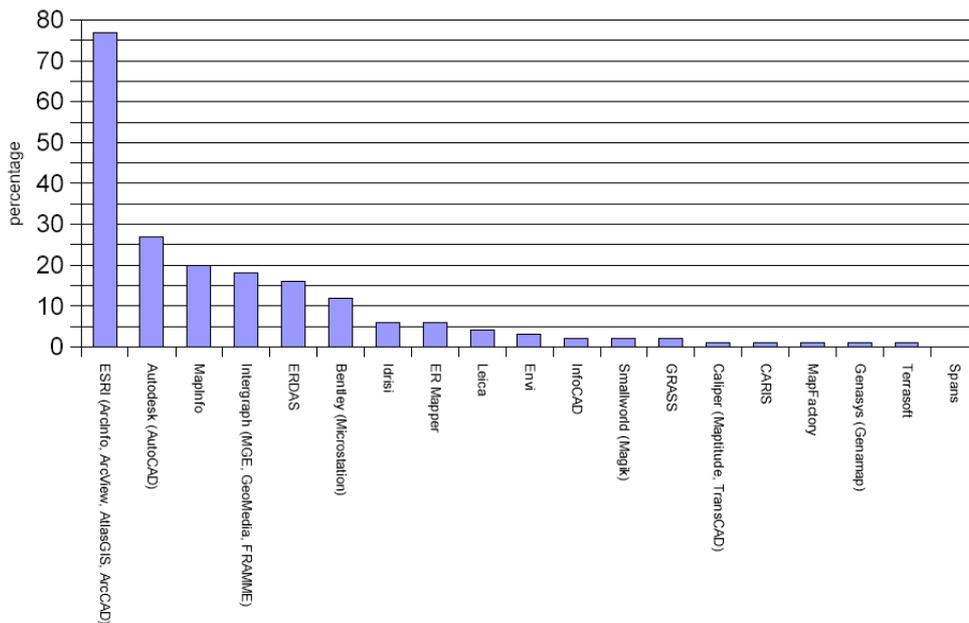
For the second and fourth points access to the source code is a precondition.

In order to guarantee this freedom, FOSS are usually licensed under the GNU General Public License (GPL): GPL licence makes restrictions that forbid anyone to deny these rights. Distributed copies of a GPL licensed program, whether gratis or for a fee, must give the recipients all the original rights of freedom; thus the distributor must also make sure that users, too, receive or can get the source code.

A survey currently carried on by the “*GISjob*” Web Site (<http://www.gisjobs.com/survey/countries.jsp>, 2006) based on 27438

## Chapter 2 – GIS and hydrological models: an overview

users (March 2006) shows how ESRI is the GIS market leader, being used by the 76% of the GIS users and GRASS is the only FOSS used (see Graph 1). This survey also shows how a part of the GIS users take advantage of more than just one software.



Graph 1, percentage of user using different GIS software

According to graph 1 the first four GIS companies providers and the relative software are:

- ESRI (Redlands, CA), sell ArcGis, a suite of three different products (ArcMap, ArcCatalog and ArcTools) with a series of extension adding features to the products. The base package is ArcView, a desktop with visualization, query, analysis, integration, and data automation capabilities (1'500 US\$ in the 2006). Some extension – costing 2'500 US\$ each – are: 3D analyst (three-dimensional visualization and analysis); geo-statistical analyst (statistical tools for modelling and advanced surface generation); network analyst (network-based spatial analysis for ArcGIS); spatial analyst (advanced raster GIS spatial analysis).
- Autodesk (San Rafael, CA), sell Autodesk Map® 3D 2006 and

## Chapter 2 – GIS and hydrological models: an overview

other products that interface with its flagship AutoCAD software package. This version of Autodesk Map software, can create, maintain, and produce maps; integrate data from various sources and formats, including Ordnance Survey MasterMap, ESRI Shape files, MapInfo TAB and Oracle Spatial; perform data analysis; and produce thematic maps. The 3D functionality includes the ability to import survey point data, build Digital Terrain Models, create contours and perform height analysis. Its price in the 2006 is 6775 US\$.

- MapInfo (Troy, NY) offers MapInfo Professional version 8.0, relatively user-friendly desktop mapping tools favoured in demography, geology, and business analysis. In the 2006 it costs about 1'500 US\$. Its features are: analysis, geocoding, 2D and 3D displaying, different data formats managing, geoprocessing, spatial querying, graph producing, reprojecting, editing and printing.
- Intergraph (Huntsville, AL) offers GeoMedia professional 6.0, its geospatial product suite. As a high-end GIS, it provides connectivity to industry-standard databases and file formats, analysis tools, and tools for data creation and maintenance (price available only by commercial sales contact).

While proprietary software grown pushed by commercial goals and concentrated their efforts in developing “easy to use” software, FOSS grew to solve problems and develop projects – often governmental or research projects – resulting in high effective but often not friendly software and therefore delegated to a small niche of GIS specialists.

The most extensively used FOSS GIS is GRASS, an application with geospatial data management, processing and analysis (vector 2D/3D, raster, volxel, images), graphics and maps production (2D/3D), spatial modelling and visualization. GRASS is currently used in academic and commercial settings around the world, as well as by many governmental agencies and environmental consulting companies.

Many other FOSS GIS and GIS-related software (more than 250) can be found at the FreeGIS Project Web site (<http://freegis.org>).

## **2.2 Hydrological models**

Models are a mere simplification or reduction in complexity of real world that express our understanding of the way the phenomenon works with sufficient precision and accuracy to allow prediction and confident decision-making. In the specific field of surface water resource rainfall-runoff models attempt to establish a link between precipitation measured over the catchment and runoff (streamflow) observed at the considered outlet. The meteorological input can be either rainfall or snowfall depending by the air temperature. Although precipitation and runoff can be considered as surface waters, the transformation of the first into the second at the catchment scale involves many hydrological processes (interception, evapotranspiration, subsurface flow, groundwater recharge, channel routing, etc.) of which the most important take place underground (Beven, 2001).

### **2.2.1 Hydrological model classification**

A possible general classification of models (Chorley and Haggett, 1967) shows four classes:

1. Natural analogues - Both historical and space analogues are used to explain events on the basis of similar occurrence happened in the past or somewhere else.
2. Physical scale models - Reproduce the physics of the phenomenon with a miniaturization of reality, here is where general behaviours, influence of variables, change in condition are studied.
3. Mathematical - Use equation, function and statistic to describe the physics of the real world, if a unique solution is given then the model is labelled as deterministic else if probabilistic solution is given then is labelled as stochastic.

4. Computational - A computer is used for symbolically represent the phenomena by mean of software code.

The first three kind of models converge in the fourth one, in fact nowadays we use all of this three types of models for develop a good understanding of the world that is then transformed in computer programs, able to produce fast, accurate and reusable estimations.

Computer models main importance reside in their ability to run scenarios more efficiently assessing the environmental implications (physical, biotic, socio-economic) that may result from the application of policies, plans or development project.

Another prerogative of computer models are their aptitude in geo-hazard disaster management by suggesting planning measures (zoning, resettlement, strengthening), giving alarms (forecasting) and assisting in post-disaster emergency (real-time monitoring).

In the field of hydrology, one of the main reason for the need of modelling the hydrological processes is due to the lack of measurement technique able to observe all the wanted information. Having just a limited capability in monitoring hydrological variables in time and space we therefore need an extrapolation of these quantities. This also gives us the ability to forecast and to evaluate different scenarios to assess the likely impact of future hydrological events or land changes.

The main applications of hydrological models are intended for managing purpose (e.g. flood mitigation; storm drainage; bridges, sewerage and culverts design; water supply; irrigation; hydro power; navigation; pollution control) by using model-extrapolated information (e.g. how much water is needed; how much water can be expected; the minimum flow; the annual flow; the floods peaks; the floods volume; the groundwater level).

To identify the different characteristics of models a series of attributes is used.

#### *Metric, Conceptual and Physically Based*

Following Beck (1991) a model can belong to one of these three general classes based on the level of conceptualization of the

physical processes:

- Metric models - are strongly observation-oriented, and they are constructed with little or no consideration of the features and processes associated with the hydrological system. A metric model omits the general laws and is in reality a representation of data.
- Conceptual models - Describe all of the hydrological processes component perceived to be of importance as a simplified conceptualizations. This usually leads to a system of interconnected stores, which are recharged and depleted by appropriate component processes of the hydrological system.
- Physically based models - Attempt to mimic the hydrological behaviour of a catchment by using the concepts of classical continuum mechanics. However, has to be noted that all existing theoretical models simplify the physical system and often include obviously empirical components, (e.g., the conservation of momentum equation used to describe surface flow includes an empirical hydraulic resistance term and the Darcy equation used in subsurface problems is an empirical equation).

#### Event and Continuous

Based on the time wise application it is suitable for it a model could be defined as:

- Event models - Represent a single runoff event occurring over a period of time ranging from about an hour to several days. The initial conditions in the watershed for each event must be assumed or determined by other means and supplied as input data. The accuracy of the model output may depend on the reliability of these initial conditions.
- Continuous models - Operate over an extended period of time, determining flow rates and conditions during both, runoff periods and periods of no surface runoff. Thus the model keeps a continuous account of the basin moisture condition and therefore determines the initial conditions applicable to runoff events. At the beginning of the run, the initial conditions must be known or

## Chapter 2 – GIS and hydrological models: an overview

assumed. However the effect of the selection of those initial conditions decreases rapidly as the simulation advances. Most continuous watershed models utilize three runoff components: direct or surface runoff (overland flow), shallow surface flow (subsurface flow) and groundwater flow, while an event model may omit one or both of the subsurface components and also evapotranspiration.

### Complete and Partial

Based on the completeness of the represented hydrologic processes:

- Complete models - Account for all the known hydrological processes significantly affecting runoff and maintain the water balance, by solving the continuity equation of precipitation, evapotranspiration and runoff:  $\text{Precipitation} - \text{Actual Evapotranspiration} \pm \text{Change in Storage} = \text{Runoff}$
- Partial models - Represent only a part of the overall hydrological process. Not to solve the water balance equation decrease the accuracy of the model and therefore it constitutes one of the most important disadvantage of partial models compared to complete models.

### Calibrated Parameters and Measured Parameters

All the models deal with some parameters, based on the way their values are estimated they could be defined as:

- Calibrated parameter models - Have one or more parameters that have to be evaluated by fitting computed hydrographs to the observed hydrographs. Calibrated parameters are usually necessary if the watershed component has any conceptual component models, which is true for most presently used watershed models. Thus, with a calibrated parameter model, a period of recorded flow is needed, usually several years, for determining the parameter values for a particular watershed.
- Measured parameter models - Have all the parameters determined satisfactorily from known watershed characteristics, either by measurement or by estimation. For example, watershed

## Chapter 2 – GIS and hydrological models: an overview

area and channel length can be determined from existing maps, channel cross sections can be measured in the field, and soil characteristics can be determined in a lab. Characteristics like channel roughness are often estimated. A measured parameter model can be applied to totally ungauged watersheds and is therefore highly desirable.

### General and Special Purpose

Based on the usability of the model in heterogenic basins:

- General models - Are acceptable, without modifications, to watersheds of various types and sizes. The models have parameters, either measured or calibrated, that adequately represent the effects of a wide variety of watershed characteristics. In order to achieve this, it is generally necessary to use conceptual models which have parameters that require calibration.
- Special purpose models - Are applicable to a particular type of watershed in terms of topography, geology or land use, e.g. an urban runoff model. Usually, such models can be applied to watersheds of different sizes, as long as the characters of the watersheds are the same.

### Lumped and Distributed:

From a spatial point of view models can be subdivided in:

- Lumped models – Also known as "spatially averaged" models, do not explicitly take into account the spatial variability of inputs, outputs, or parameters. They are usually structured to utilize average values of the watershed characteristics affecting runoff volume. Averaging a certain parameter also implicitly averages the process being represented. Because of non-linearity and threshold values, this can lead to significant error.
- Distributed models - Include spatial variation in inputs, outputs, and parameters. In general, the watershed area is divided into a number of elements (e.g. cells) and runoff volumes are first calculated separately for each element.

*Steady and Unsteady:*

Based on the time dependence of the flow equations:

- Steady models – also known as static models, consider the depth of the accounted flows constant within a given time interval.
- Unsteady models – also known as dynamic models, consider the depth of the accounted flows as a function of time.

### **2.2.2 Historical review of Hydrological Models**

The development of a rainfall-runoff model for the simulation of continuous streamflow time-series is a rather recent issue in water research. Until the middle of the 1960s, hydrologic modelling involved the development of models and theories of individual components of the hydrological cycle, such as overland flow, infiltration, evapotranspiration, interception, base flow, subsurface flow. For example, one of the first attempts to develop a theory for infiltration was by Green and Ampt (1911): they used simplified principles of physics to derive a formula that is still popular for computing the infiltration capacity rate. Later on, in 1933, Horton developed a theory of infiltration to estimate rainfall excess and improve hydrograph separation techniques.

The first simulation rainfall-runoff models as integration of the different components summarized above were proposed at the very beginning of the 1960s when modern computer appeared (Singh, 2002).

In the following 50 years an abundance of different mathematical models has been proposed to simulate the transformation of precipitation into stream flow (see paragraph 2.2.1). Their diversity comes from different sources: the large number of point of view existing between hydrologists, the wide range of physical and climate conditions and the wide range of hydrological issues that a rainfall-runoff model can help to solve.

The first attempt to model virtually the entire hydrologic cycle was by the Stanford Watershed Model-SWM (now HSPF) by Crawford

## Chapter 2 – GIS and hydrological models: an overview

and Linsley (1966). Immediately after another model that became really popular was developed: HEC-1 (Hydrologic Engineering Center 1968) by Dawdy and O'Donnell (1965). These models represented in different “conceptual” ways the response mechanism of the various phenomena and interconnections of the systems. In theory, if the structural description was exact, the parameters of the models, such as the water storage capacities, roughness coefficients and threshold effects could be set in relation to the actual physical quantities of the basin. Unfortunately, in many cases, due to the large number of parameters used in the models, calibration is needed anyway because it is impossible to have measurements for all the parameters and sometimes it is easy to obtain a set of unrealistic parameters.

This lack of one to one relationship between model and reality induced Freeze and Harlan (1969) to propose a new mathematical model based on distributed physical knowledge of the phenomena by means of the numerical integration of the differential equations systems describing surface flow, base flow, flow in the unsaturated area. Unfortunately, because of the still limited computational resources, the approach was directed only at small, almost impermeable mountain basins.

Anyway, the Freeze and Harlan approach was later on considered by several research institutes such as Danish Hydraulic Institute (Denmark), the Institute of Hydrology at Wallingford (United Kingdom) and SOGREAH (France) with the aim to develop a physically based, distributed catchment model that would express the phenomena by integrating the differential equations with partial derivatives based on continuity of mass and momentum and based on assigned boundary conditions.

The first model developed based on this philosophy is the SHE (Système Hydrologique Européen) (Abbot et al., 1986a-b) model. The characteristics of the SHE model will be presented in the following paragraph.

Still of relevant importance in the 1970s is the theory of Dunne and Black (1970) for the runoff generation. They suggested the idea of a

## Chapter 2 – GIS and hydrological models: an overview

“saturation from below” (saturated excess) mechanism for runoff production. Now on, we can have one more fundamental characteristic for hydrological model: they can be based on the Hortonian (hortonian excess, saturation from above) or Dunne (saturated excess, saturation from below) approach.

The first definition of distributed model is from Freeze and Harlan (1969). This is based on the idea that it is possible to remove some of the subjectivity of the conceptual models using equations describing all the surface and subsurface processes in the catchment. In their paper, Freeze and Harlan wrote down the equations for different surface and subsurface flow processes and showed how they would be linked by means of common boundary conditions into a single modelling framework.

In theory, this class of models is the best one because physical parameters used in the model can be measured in the field and the parameters can be defined for every element in the solution mesh. With so many parameters the calibration can be difficult, but, at the same time, it also impossible to measure all the parameters in the field. Furthermore, most measurements techniques can only be used to derive values at scales much smaller than the element mesh used in the approximate solution. In principal, parameters should be strongly related to the physical characteristic of the surface, soil and rock, but in this way, we loose “cell by cell” definitions and we build something like zones with different characteristics.

Despite these problems, the last decade saw a wide use of distributed models, mainly thanks to the increased computer power.

The most important model included in this class is the SHE model.

The fully distributed models described here are complex and both computationally and parametrically demanding. A separate group of models based anyway on the fully distributed ones have been developed: they attempt to maintain a distributed description of the catchment responses but they do that in a simpler way, without all the detailed process representations of the fully distributed models. In this class we can for example include TOPMODEL, the topographic-based model of Beven and Kirkby (1979).

### **2.2.3 Principal characteristics of some of the currently used distributed watershed models**

#### *SHE model*

The SHE model is physically-based: the hydrological processes of water flow are modelled either by finite difference discretization of the partial differential equations of mass, momentum, and energy conservation, or by empirical equations derived from independent experimental research. It uses one-dimensional finite-difference solutions for channel reaches and the unsaturated zone in each grid element, and two-dimensional solutions in plan for the saturated zone and overland flow processes.

The components of the hydrological cycle are modelled as follows: interception, by the Rutter accounting procedure; evapotranspiration, by Penman-Monteith equation; overland and channel flow, by simplifications of the Saint Venant equations; unsaturated zone flow, by the one dimensional Richard equations; saturated zone flow, by the two-dimensional Boussinesq equation; snowmelt, by energy budget method.

In the SHE model there is the potential to have different parameters for every grid element and within each element of the different vertical layers. This requires the specification of thousands of parameters.

The computational requirements of these models are extreme and, for this reason, most current physically based models introduced suitable simplified mathematical formulations. For example, a great deal for the computational burden occurs in any description of a partially saturated soil water systems because of high non-linearity of the process. The SHE model, for example, treats unsaturated soil water flow as a principally vertical process that links surface and saturated subsurface hydrologic components. Such an approximation makes physically based simulations possible for very large catchments, although the validity of effective parameter values that must be used with large scale catchments has been

questioned (Beven, 1989).

In terms of prediction, the SHE model has been proved to be dependent on the grid scale by Refsgaard (1997).

For the SHE model impressive pre- and post-processing packages have been developed by the different development teams.

### TOPMODEL

TOPMODEL (TOPography MODEL) is a variable contributing area conceptual rainfall-runoff model in which predominant factors determining the formation of runoff are represented by the topography of the basin and a negative exponential law linking the soil transmissivity with the vertical distance from the ground level. The spatial variability of topography is reflected in the topographic index,  $k=a/\tan\beta$ , where  $a$  is the area draining through a point from upslope and  $\tan\beta$  is the local slope angle (Beven, 1997). This index is used as an index of hydrological similarity. The simplicity of the model comes from the use of this index.

TOPMODEL represents a first attempt to combine the computational and parametric efficiency of a lumped approach with the link to physical theory and possibilities for more rigorous evaluation offered by a distributed model (Beven et al., 1995).

Though a conceptual model, i.e. a model in which the physical reality is represented as simplified, TOPMODEL is frequently described as being “physically based”, in the sense that its parameters can be measured directly “in situ” (Beven and Kirkby, 1979). However, this definition is optimistic, considering the uncertainties linked in the definition of parameter values even for physically based models. Franchini et al, 1996, studied the TOPMODEL parametrisation with the aim to understand the correspondence between the assumptions underpinning the model and the physical reality, and in particular the role that topographic index information and the nature of the soil have within the model itself. Surprisingly, even though the topography plays an important role in the formation of runoff schematization, the model

## Chapter 2 – GIS and hydrological models: an overview

was found to have little sensitivity to the basin's actual index curve, whereas a close connection exists between the hydraulic conductivity at the ground surface and the size of the grid of the DEM used to calculate the index curve. In general, the hydraulic conductivity increases with the increasing of the grid size, more precisely, it takes values without physical meaning and this reduces the physical credibility of the model. For another analysis of the TOPMODEL limitations, an interesting analysis is presented by Beven (1979).

### WaSiM

WaSiM – Water balance Simulation Model – has been developed at the Institute of Geography at ETH in Zurich (Zollmann, 2000; Schulla, 1997, 1999).

The first version of the model is based on a TOPMODEL approach, later on Richards equation for the unsaturated zone and an integrated multi-layer-2D groundwater model were introduced.

The use of Richard's equations for the description of water flow in the unsaturated zone is the most interesting change with respect to the old conceptual TOPMODEL approach. Furthermore the introduction of the transport algorithms into the model makes the model useful to simulate tracer and salt transport.

WaSiM has already been used successfully in the Swiss Alpine area, running with a grid resolution of 100 m x 100 m. The hydropower system has been considered externally. It is fully distributed and considers all components of the water cycle.

### TOPKAPI

Topkapi (TOPographic Kinematic Approximation and Integration) (Todini and Ciarapica, 2001) is a distributed and physically-based rainfall-runoff model. It is driven by a saturation excess approach (Dunnie et al., 1978) and uses three non-linear reservoir differential equations, for the drainage in the soil, the overland flow on saturated or impervious soil, and the channel flow along the drainage network, respectively. The geometry of the catchment is

described by a lattice of cells (the pixels of a DEM and their slope) over which the equations are integrated to lead to a cascade of non-linear reservoirs. The parametrisation relies on the digital thematic maps of soil, geology and land use. Initially the Topkapi model (Todini and Ciarapica, 2001, Liu and Todini, 2002) was structured around five modules: evapotranspiration, snowmelt, soil water, surface water and channel water.

### **2.3 GIS for water resources**

As discussed in the previous paragraphs, different models have been developed in the past, aiming at reproducing measured variables or, more often, to predict hydrological processes in ungauged catchments and into the future. However, the final aim of using models is very often the improvement of decision-making about a hydrological problem, such as flood protection, water resources allocation, etc.

In recent years, given the necessity to evaluate the consequences of environmental changes exacerbated by the effects of human intervention and climate changes, researchers have turned their attention to an integrated approach to all physical events.

Hydrological modelling is also moving in this direction, developing physically based and spatially distributed models capable of interacting and/or integrating with other interconnected processes.

This kind of modelling calls for the handling of a vast quantity of data and parameters, often coming from different sources and methodologies of measurement corresponding to different spatial and temporal scales.

Because of its advantages of data collection, storage, management, analysis, format-conversion, and display GIS is a very powerful and promising tool in water resource assessment and management.

#### **2.3.1 GIS applications in hydrology**

The first kind of GIS application for hydrological purpose is related

## Chapter 2 – GIS and hydrological models: an overview

to data pre- and post-processing. This type of application makes an extensive use of the following GIS capabilities:

Geostatistic – Hydrologists use geostatistical tools to estimate the distributed field of point-wise observations such as elevation, precipitation, and chemical concentrations (Mitas and Mitasova, 1999). This functions are mainly used to extract input data for hydrological models;

Map algebra and map statistic – Map algebra allows the calculation between maps by means of arithmetic and logic operators. This is particularly useful in parameter estimation. In fact, for instance, it is possible to generate conductivity maps based on soil type classes or CN maps based on landuse type (Pandey and Sahu, 2002). Also same statistic functions for raster data (i.e. the GRASS “*r.stats*” command and the ESRI “*zonal statistics*” command of the *spatial analyst* tool) can help in determining values to be used as an input in lumped models: they can detect for different sub-basins the areas, the mean slope, the mean height, the percentage area for each landuse class, and so on;

Visualization – the presentation of the results is very important because it can strongly affect the decision-making understanding of the modeled phenomena. Once the model outputs are imported into the GIS, it can be used to produce 2D images and large format maps, rendered 3D views, and animations showing the evolution of the phenomena.

Hydrological tools – These are new hydrological procedures developed into the GIS. Among the existing ones, we can note the *Arc Hydro Toolset*, the *TauDEM*, the *RiverTools*, the hydrological package of JGrass and the GRASS existing commands.

The *Arc Hydro Toolset* of the ESRI's ArcGIS suite ([www.crwr.utexas.edu/gis/archydrobook/ArcHydro.htm](http://www.crwr.utexas.edu/gis/archydrobook/ArcHydro.htm), 2006) is a suite of tools which facilitate the creation, manipulation, analysis and display of features and objects within a given hydrological data model (ArcHydro Data Model). Some of the included operational capabilities are for example terrain pre-processing, watershed processing, attribute manipulation, and network utility.

## Chapter 2 – GIS and hydrological models: an overview

The *TauDEM* (Terrain Analysis Using Digital Elevation Models) is an interesting extendible component (plug-in) to ESRI ArcGIS (Tarboton, 1997): a package for the analysis of digital elevation data and mapping of channel networks and watersheds.

The *RiverTools* (<http://www.rivertools.com/>, 2006) is a user-friendly GIS application for analysis and visualization of digital terrain, watersheds and river networks. This software from Research Systems Inc. provides a set of analysis tools to process and manipulate digital terrain as well as create watershed boundaries, flow models, stream delineations and hydrological analyses. *RiverTools* is written in IDL language and is designed to work with other GIS applications such as ESRI's ArcView (via support for shapefiles, BIL, FLT, GeoTIFF) and with remote sensing, image processing systems such as ENVI (Environment for Visualizing Images).

JGrass (<http://www.hydrologis.com/jgrass.html>, 2006) is a Java based framework for the GRASS Geographic Information System. It has been developed to generate a system independent GRASS interface in order to facilitate its use among a larger number of GIS professionals using a wider range of computer systems. Developed at the International Centre for Environmental and Nuclear Sciences of Kingston, Jamaica and at the Department of Civil and Environmental Engineering of Trento, Italy, it is distributed with an hydrological package by HydroloGIS (a new company that focuses it's efforts on the environmental topic) Web Site. The hydrological package include river extraction, flow direction delineation, fill pit procedure, and many other hydrological commands.

Also into the GRASS GIS a series of hydrological commands exist, some of them are:

- *r.fill.dir* – filters and generates a depressionless elevation map and a flow direction map from a given elevation layer;
- *r.drain* – races a flow through an elevation model on a raster map layer;
- *r.slope.aspect* – generates raster map layers of slope, aspect, curvatures and partial derivatives from a raster map layer of true

elevation values;

- r.terraflow – flow computation for massive grids;
- r.watershed – watershed basin analysis program;
- r.water.outlet – watershed basin creation program.

### **2.3.2 Linking GIS and hydrological model**

In order to take advantage of all the above described capabilities a data exchanges/sharing procedure has to be generated.

This interoperability requirement between hydrological models and Geographic Information Systems led to the establishment of links at three different levels (Wesseling et al., 1996; Mitasova and Mitas, 2001):

1. a simple exchange of data (loose coupling);
2. an interface with GIS (tight coupling);
3. an integrated model (embedded coupling).

#### Loose coupling approach

This level of integration involves an external conversion program that allows data exchanges between the two stand alone packages. In this case the model is run independently of GIS which is used for process and export the model input data, and import the model results for analysis and visualization. This approach requires low programming efforts and thus results to be the simplest of the three. The downside is that this approach requires a big operator work that can drives to errors .

Some examples are:

- *TOPKAPI* – a distributed and physically based model (Ciarapica and Todini, 2002) whose input and output in ASCII format can be imported and exported by ArcView and GRASS GIS;
- *WaSiM* – a distributed and physically based hydrological model developed at the Swiss Federal Institute of Technology (ETH) in Zurich has commands for importing and exporting data from and to ArcInfo (<http://iacweb.ethz.ch/staff/verbunt/Down/WaSiM.pdf>, 2006)

## Chapter 2 – GIS and hydrological models: an overview

- *DHSVM* (Distributed Hydrology Soil Vegetation Model) – originally developed in the early 1990s by Wigmosta at the University of Washington (Wigmosta et al, 1994), freely available on the web ([www.hydro.washington.edu/Lettenmaier/Models/DHSVM/](http://www.hydro.washington.edu/Lettenmaier/Models/DHSVM/), 2006). By using a suitable function (*myconvert.c*) it is possible to convert an ArcInfo exported ASCII file into the binary format of the hydrological model.

### Tight coupling approach

This approach generally embeds an hydrological model into the GIS by means of a common interface driving users in input data processing, model running, and results analysis and presentation. In this case the model is still run independently from the GIS, but the common interface allows automatic procedures for data export/import between the two systems with the effect of easier to use model and limited human errors. This type of link is the most diffuse because does not imply any dependence of the model from the GIS software (possible changes in data structure, libraries or programming tools with different versions).

Same examples are:

- *SWAT* (Soil & Water Assessment Tool) – developed to predict the effect of alternative management decisions on water, sediment, and chemical yields with reasonable accuracy for ungauged rural basins (Neitsch et al., 2002). The SWAT/GRASS input and output interface tools were developed by R. Srinivasan at Blackland Research Center, Texas Agricultural Experiment Station (TAES), Temple, Texas.
- *GLEAMS* (Groundwater Loading Effects of Agricultural Management Systems) – is a field-scale model used to evaluate the effects of management practices on movement of chemicals in the plant root zone (Knisel, 1993).
- *WEPP* (Water Erosion Prediction Project) – a process-based, distributed parameter, continuous simulation, erosion prediction model firstly used with GRASS by Sabivi et al. (1995).

## Chapter 2 – GIS and hydrological models: an overview

- GeoWEPP (Geo-spatial interface for the Water Erosion Prediction Project) – an enhanced version of the WEPP model interfaced with ArcView 3.x and with ArcGIS 9 (release of May 2006) (Renschler, 2003). It allows the import of DEMs, the channel and watershed delineation, the model input parameter setting and the model running.

### Embedded coupling

Into this methodology of integration two opposite subtypes of approach can be recognized. The first one involves the development of GIS-like capabilities into the hydrological model: this leads to big programming efforts and usually limited GIS functions, the advantage is that the development is with no data structure constraints (it is GIS software independent).

Some examples are:

- *RiverCAD* – is a hydraulic model with CAD embedded capabilities with support for HEC-2 e HEC-RAS.
- MIKE-SHE – a fully modular system for mathematical description of the land phase of the hydrological cycle derived from the SHE model (Abbot et al., 1986a, 1986b); it integrates modelling of surface water and ground water and has a sophisticated GUI (Graphical User Interface) for data manipulating and visualization.

The other approach embeds hydrological model routines into the GIS software: this leads to the development of the model by using GIS libraries, and thus taking full advantage of all GIS capabilities and avoiding any needs of data import/export and duplication (data in the GIS format and data in the model format).

- *TOPMODEL* – a widely used semi-distributed topographically based model, (Beven and Kirkby, 1979) of which a GRASS version exists (*r.topmodel*).
- *CASC2D* – is a fully-unsteady, physically-based, distributed-parameter, raster (square-grid), two-dimensional, infiltration-excess (Hortonian) water surface model for simulating the

hydrologic response of a watershed subject to an input rainfall field (Ogden, 1998). In addition to a loose coupling level of integration with ArcInfo, a completely GRASS integrated version of the model is available (*r.hydro.CASC2D*).

- GSSHA - is a physically based, distributed-parameter, structured grid, hydrologic model that simulates the hydrologic response of a watershed subject to given hydro-meteorological inputs ([http://chl.erd.c.usace.army.mil/software/gssha/Primer\\_20/gssha\\_primer\\_20.htm](http://chl.erd.c.usace.army.mil/software/gssha/Primer_20/gssha_primer_20.htm), 2006). It is an evolution of the CASC2d model. It is fully embedded in the Watershed Modeling System (WMS), a GIS-based data processing framework for watershed hydrology and hydraulics modelling developed by the Scientific Software Group company ([http://www.scisoft.com/products/wms\\_overview/wms\\_overview.html](http://www.scisoft.com/products/wms_overview/wms_overview.html), 2006).
- *LISFLOOD* – a physically based river basin model (De Roo et al., 2000), written using the PCRaster GIS environment; the PCRaster language is a computer language for construction of spatio-temporal environmental models and supports immediate pre- or post-modelling visualisation of spatio-temporal data (Van Deursen, 1995).
- *ANSWERS* (Areal Nonpoint Source Watershed Environmental Response Simulation) – is an event oriented, distributed parameter model that was developed to simulate the behaviour of watersheds having agriculture as their primary use (Beasley et al., 1980). It is present into the GRASS GIS with the command *r.answers*.

## 3. Methods

Recent development of GIS and remote sensing technology make possible to capture and manage a vast amount of spatial data relative to hydrological parameters and variables. Linking GIS and hydrological models is therefore of rapidly increasing importance to test and validate these data in watershed management.

### 3.1 Chosen approaches

In this paragraph I present the choices of the GIS-model integration level approach, the GIS package and the hydrological model type. Motivation and advantages are also discussed.

#### *3.1.1 GIS-model integration level*

As discussed in paragraph 2.3.5 three integration level between models and GIS exist.

With the present research I intend to develop a system able to be used as a Watershed Management Information System (WMIS). This involves the ability to link geospatial data (stored into GISs) with a model (describing the evolution or state of phenomena, in this case the water flow) and the GIS's analysis tools (to evaluate the phenomena with respect to a further analysis of the relationship with other factors like for example landuse changes).

The GIS embedded approach allows the development of such a system in a single environment avoiding the time consuming operations like data import/export and format conversion that otherwise are needed.

Furthermore this approach allows an efficient and not redundant development: existing algorithms (i.e interpolation techniques, map algebra, basin extraction) and existing optimized libraries (i.e. compressed format files, optimized data access and writing) can be used avoiding code duplication.

### 3.1.2 GIS package

Among the existing GIS packages described in the literature review (chapter 2), GRASS is the software chosen for this research.

A short description of this GIS software given by Dr. Helena Mitsova in one of the GRASS mailing lists is: "Geographic Resources Analysis Support System. Commonly referred to as GRASS, this is a Geographic Information System (GIS) used for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization. GRASS is currently used in academic and commercial settings around the world, as well as many governmental agencies and environmental consulting companies". Further information about the GRASS GIS can be found on the official GRASS web page <http://grass.itc.it/> (2006).

Five relevant points motivate this choice:

- data format – it handles all the three existing data structures: raster, vector and database. It is able to use a large number of format inside each data structure by using OGR-GDAL library (for raster and vector format conversions) and PostgreSQL, DBF, SQLite, MySQL, and ODBC drivers (for database connection).
- Code access – it is a public domain free and open source GIS.
- Documentation and support – it is supported by an extensive documentation with programming manuals, commands manuals, books, tutorials, and a huge community with numerous users groups, mailing lists, and a Wiki site (a collaborative Web site including the perpetual collective work of many authors: a Wiki allows anyone to edit, delete or modify content that has been placed on the Web site using a browser interface)
- Programming language – it is written in standard ANSI C, a language robust, fast, and widely used in modelling.
- Specific existing functions – it includes a large number of hydrological modules that are used in the development of this research.

*3.1.3 Hydrological Model type*

The chosen model type is: continuous, distributed, physically based, and modular.

The continuous and distributed approach has been selected because a Watershed Management Information System (WMIS) should be able to estimate the state of the hydrological quantities over time and space. For example it is useful to understand the influence of landuse changes, to implement a flood warning system, to monitor the catchment status, to predict the state of ungauged catchment, and to estimate the movement of pollutants and sediments.

The physically based approach has been selected because it attempts to link catchment behaviour with measurable physical landscape properties (i.e. soil moisture); this can be a valuable characteristic in coupling with other interconnected models (i.e. landslide hazards model) and in model validation.

I used a modularity approach(simulation of the different processes by using separate functions) because:

- it can be helpful in future improvements or changes of the code;
- it allows a full control of the rainfall-runoff process in terms of freedom, in combining and choosing the desired processes;
- it permits the use of single modules for whatever purpose (i.e. evaporation for irrigation planning);

### 3.2 Rainfall-runoff conceptual model

In this paragraph the physical model, the logic model and the mathematical model of the rainfall-runoff process are described.

#### *3.2.1 Hydrological physical model*

In the rainfall-runoff modelling the aim is to simulate the water movement of a sub-system of the water cycle that is related to the watershed area (herein referred to as watershed system). That sub-system is therefore linked to the global water circulation and constitutes a process with exchanges of mass and energy. As represented in figure 3.2.1, the watershed system includes a series of sub-processes.

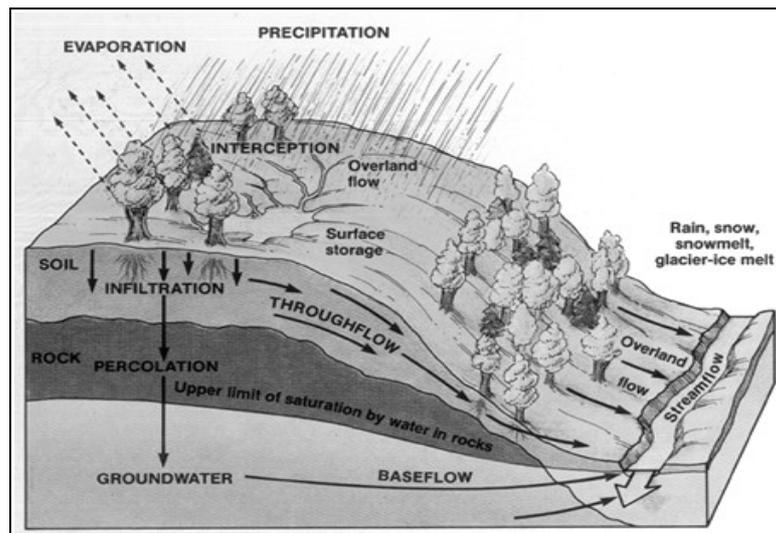


Figure 3.1 – watershed physical model

At first, the precipitation can be solid (snow) or liquid (rain), depending on the air temperature. It can fall over the canopy and a part of it can be intercepted while the rest of the rain reaches the ground. Then the energy of the sun induces the water evapotranspiration and/or the snow melting. Later on the water on the ground flows following two ways: along the surface (faster process) or along the deeper layers (slower process).

The first way involves the flows over and into the first thin upper

soil layer; the second way involves the percolation and the flows into the deeper soil layer.

### 3.2.2 Hydrological logic model

Schematizing the physical model of the watershed system the logic model can be derived (see figure 3.2).

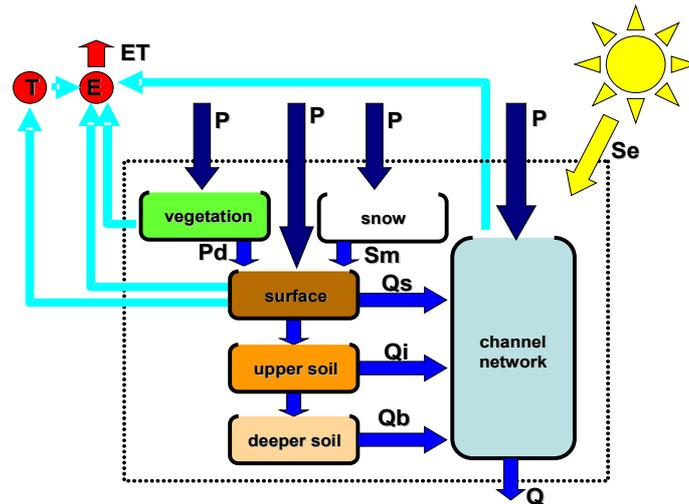


Figure 3.2 – watershed logic model ( $P$ =precipitation,  $Pd$ =drained precipitation,  $Sm$ =snowmelt,  $Qs$ =surface discharge,  $Qi$ =subsurface discharges,  $Qb$ =groundwater discharges,  $E$ =evaporation,  $T$ =transpiration,  $ET$ =evapotranspiration,  $Se$ =solar energy)

In this schematization of the watershed system the precipitation  $P$  (mass input) and the solar energy  $Se$  (energy input) are the two inputs, while the river discharge  $Q$  (mass output) and the evapotranspiration (mass output) are the two outputs.

The water during its transformation from the input (rainfall) to the output (runoff) pass through different stages that accumulate, delay, route and lose part of the mass.

The processes that mainly describe these stages are:

- the vegetation module – it mimics the canopy interception process causing mass losses;
- the snow module – it mimics the snow accumulation and melt process causing accumulation and therefore delay;
- the surface module – it mimics the routing of water over the

surface causing low delay;

- the upper soil module – it mimics the routing of water into the first thick soil layer causing medium delay;
- the deeper soil module – it mimics the routing of water into the deep soil causing high delay;
- the channel module – it mimics the routing of the water in the channels causing very low delay and output generation.
- The solar energy affects different stages and it is used to evaporate the water intercepted from the canopy; to transform the accumulated snow to water; and to evaporate the water that resides in the upper soil.

For model implementation all these modules have to be formally defined by mathematical equation that represents the single process: the next section will describe the equations governing the different processes.

### **3.3 Rainfall-runoff mathematical model**

In the following paragraphs the chosen approach for each component of the logic model is described and the relative mathematical formulation is reported.

Looking at the existing solutions I used different formulation found in literature that shown proven accuracy. Following this idea the SHE approach was used to simulate the snow melt, the Penman-Monteith approach to simulate the evapotranspiration, the TopKapi approach to simulate the soil flux component, and the Rutter approach to simulate the canopy interception

#### 3.3.1 Evapotranspiration

This physical process is controlled by both the energy availability on the evaporating surface and the capability of which the water vapour to be diffused in the atmosphere. Methods, that differ in the specific data availability (temperature, global radiation, wind velocity, etc.) and the temporal resolution (hourly up to several days), have been developed and can be found in literature (Shuttleworth, 1992; Xu and Chen, 2005).

The approach used in this work is the hourly Penman-Monteith method as presented in Allen et al. (1998) for land surfaces and the Penman method (Penman, 1948) for water surfaces.

This choice is due to:

- no calibration needs – unlike other methods, neither a local calibration nor the definition of the wind field on the basis of the measurements are needed: it allows to comparing data from different zones;
- it is physically based – it explicitly contains the parameters describing both physiological and aerodynamic components;
- it is recommended by FAO – the results of the studies by many authors summarized in the document of the experts of FAO (Allen et al., 1998), suggest to use the Penman-Monteith method for the determination of the potential evapotranspiration.

Being:

$\lambda$	=	latent heat of vaporisation [MJ·Kg <sup>-1</sup> ];
$\Delta$	=	slope of vapour pressure curve [kPa·°C <sup>-1</sup> ];
$\gamma$	=	psychrometrics constant [KPa·°C <sup>-1</sup> ];
$\gamma^*$	=	modified psychrometric constant ( $r_s$ and $r_a$ dependent);
$\rho$	=	atmospheric density [Kg·m <sup>-3</sup> ];
$c_p$	=	specific heat of moist air [KJ·Kg <sup>-1</sup> ·°C <sup>-1</sup> ];
$r_s$	=	bulk surface resistance [s·m <sup>-1</sup> ];
$r_a$	=	aerodynamic resistance [s·m <sup>-1</sup> ];
$e_s$	=	mean saturation vapour pressure [KPa];
$e_a$	=	actual vapour pressure [KPa];
$R_n$	=	net radiation [W·m <sup>-2</sup> ];
$G$	=	soil heat flux [MJ·m <sup>-2</sup> ·d <sup>-1</sup> ];
$u_2$	=	wind speed at 2 meters height [km·h <sup>-1</sup> ];
$ET_p$	=	potential evapotranspiration [mm·h <sup>-1</sup> ];
$ET_{aero}$	=	aerodynamic term of the evapotranspiration [mm·h <sup>-1</sup> ];
$ET_{rad}$	=	radiative term of the evapotranspiration [mm·h <sup>-1</sup> ];
$f(u)$	=	mass transfer coefficient [mm·mb <sup>-1</sup> ]

The applied equations for Penman-Monteith are:

$$ET_p = ET_{rad} + ET_{aero};$$

$$ET_{aero} = (0.001/\lambda) \cdot [1/(\Delta + \gamma^*)] \cdot (\rho \cdot c_p / r_a) \cdot (e_a - e_s); \quad [\text{eq. 3.1}]$$

$$ET_{rad} = [\Delta/(\Delta + \gamma)] \cdot [(R_n - G)/\lambda \cdot 10^6];$$

and for Penman are:

$$ET_p = [ (R_n \cdot \lambda \cdot \Delta) + (\gamma \cdot ET_{aero}) ] / (\Delta + \gamma);$$

$$ET_{aero} = f(u) \cdot (e_a - e_s); \quad [\text{eq. 3.2}]$$

$$f(u) = (0.35/24) \cdot (0.5 + (0.621375 \cdot u_2/100)) \cdot 7.500638$$

where  $f(u)$  is expressed following the AWSET software approach ([www.silsoe.cranfield.ac.uk/iwe/\\_private/software/awset/awset3.pdf](http://www.silsoe.cranfield.ac.uk/iwe/_private/software/awset/awset3.pdf), 2006).

### 3.3.2 Canopy interception

Canopy interception plays an important role in determining the amount of rainfall reaching the soil surface. The amount of rainfall intercepted by a forest canopy depends on the storm size, intensity, and duration, the rainfall frequency, the forest structure, and the tree species, age, and density. Depending on all these conditions, canopy interception can account for 15 to 35% of annual rainfall, although the influence during intense storm events is obviously reduced (Blake, 1975).

The chosen method is the one presented in Zengh et al. (2000), a Rutter approach (Shuttleworth, 1988) modified to take into account the temporal variability of rain. This approach has been selected because it is physically based and permits to estimate different interception losses for drizzles (higher) and intensive showers (lower).

The canopy interception is modelled by means of a storage described by a maximum storage capacity ( $C_m$  [mm]), filled by precipitation ( $P$  [mm]) and dried by evaporation ( $IL$  [mm]) and drainage ( $D$  [mm]), both in function of the current storage level ( $C_a$  [mm]).

The water balance for the canopy storage is evaluated by:

$$\frac{\partial C_a}{\partial t} = P - IL - D \quad [\text{eq. 3.3}]$$

The intercepted water exceeding  $C_m$  is assumed to drain instantaneously to the ground and therefore the drainage  $D$  is modelled by the following equation:

$$D = \begin{cases} \infty & C_a > C_m \\ 0 & C_a \leq C_m \end{cases} \quad [\text{eq. 3.4}]$$

and the  $IL$  (called interception loss) is:

$$IL = \left( \frac{C_a}{C_m} \right) \cdot ETp \quad [\text{eq. 3.5}]$$

with  $ETp$  = evaporation from a water surface.

### 3.3.3 Snow melt and accumulation

As shown in Harr (1981), snow melt in mountainous region is an important factor in surface water balance.

Despite the numerous complex models proposed in literature (i.e. Morris, 1991; Rango and Martinec, 1995; Tarboton & Luce, 1996), the World Meteorological Organization (WMO, 1986) shows how no clear advantage in using high complex model can be detected.

Thus the SHE model approach (Abbott at al., 1986a,b) has been chosen because it is quite simple but still physically based.

The snow pack dynamic is driven with a simplified energy budget.

By using the hypothesis of zero snowmelt the water equivalent mass ( $Z^*$ ) and the energy ( $E^*$ ) are estimated with the equations 3.6:

$$\begin{aligned} Z^*(t + \Delta t) &= Z(t) + Ps \\ E^*(t + \Delta t) &= E(t) + Rad + Qp; \end{aligned} \quad [\text{eq. 3.6}]$$

where:  $Ps$  is the estimated solid precipitation,  $Rad$  is the available net solar radiation, and  $Qp$  is the energy advected by precipitation (solid and liquid) into the snow.

If the total available energy is bigger than that required to maintain the total mass in a solid phase at the temperature  $T_0$  of  $0^\circ\text{C}$ , then the snowmelt, the updated energy ( $E$ ) and mass content ( $Z$ ) of the snowpack are calculated by means of the equations 3.7:

$$Sm = \frac{E^*(t + \Delta t) - C_{si} \cdot T_0 \cdot Z^*(t + \Delta t)}{C_{lf}} \quad [\text{eq. 3.7}]$$

$$Z(t + \Delta t) = Z^*(t + \Delta t) - Sm$$

$$E(t + \Delta t) = E^*(t + \Delta t) - (C_{si} \cdot T_0 + C_{lf}) \cdot Sm;$$

where:  $C_{si}$  is the specific heat of ice,  $C_{lf}$  is the specific latent heat of fusion of water, and  $Sm$  is the snowmelt.

Otherwise no snowmelt occurs and the snowpack energy content remains available for further steps.

### 3.3.4 Flow component

As described in the logical model of the rainfall-runoff process the water that reaches the ground, once depleted of the evapotranspiration ratio, is available for the runoff.

The runoff may occur on the ground surface (surface flow), in the first thick soil layer (subsurface flow), in the deeper soil (groundwater flow), in the rivers (channel flow), through the lakes (lake), and through the artificial basins (dam).

As already mentioned the Topkapi approach is limited in parameters number and it is scale independent.

For the sake of simplicity, in this first version, only the three major flow component involved in floods generation are considered: the subsurface flow, the surface flow, and the channel flow.

#### *Subsurface flow*

There are five fundamental basic assumptions for the description of this component:

1. precipitation is assumed to be constant over the integration domain (namely the single cell), by means of suitable interpolation of the local rainfall data, such as Thiessen polygons techniques, Block Kriging or others.
2. all the precipitation infiltrates into the soil, unless the soil is already saturated; this is equivalent to adopting, at the pixel scale, the saturation excess or Dunne mechanism (Dunne, 1978),

instead of the possible activation of the Hortonian mechanism due to infiltration excess;

3. the slope of the water table is assumed to be equal to the slope of the ground, unless the latter is very small (less than 0.01%); this is the fundamental assumption of the approximation of the kinematic wave in the De Saint Venant equations, and it implies the use of a kinematic wave propagation model with regard to horizontal flow, or drainage, in the unsaturated area (Henderson and Wooding, 1964; Beven, 1981, 1982);
4. the integral of the hydraulic conductivity over the vertical in the unsaturated zone, an extension of the concept of transmissivity, together with the horizontal flow, can be reasonably expressed as a function of the total water content of the soil, i.e. on the basis of the integral of the water content profile in the vertical direction ;
5. saturated hydraulic conductivity is constant with depth in a surface soil layer but much larger than that of deeper layers; this is the basis for the vertical aggregation of the transmissivity, and therefore of the horizontal flow.

The kinematic wave formulation for sub-surface flow is then based on the approximation on horizontal transmissivity, the continuity and momentum equations for soil water and is expressed as follows:

$$\begin{cases} (\vartheta_s - \vartheta_r)L \frac{\partial \tilde{\theta}}{\partial t} + \frac{\partial q}{\partial x} = p \\ q = \tan(\beta) k_s L \tilde{\theta}^\alpha \end{cases} \quad [\text{eq. 3.8}]$$

where  $x$  is the main direction of flow along a cell (dimension of the cell) [m],  $t$  is the time [s],  $q$  is the flow in the soil due to drainage [m<sup>3</sup>/s], corresponding to a discharge per unit of width,  $p$  is the intensity of precipitation [mm/h],  $\theta_s$  is the saturated soil water content [m<sup>3</sup>/m<sup>3</sup>],  $\theta_r$  is the residual soil water content [m<sup>3</sup>/m<sup>3</sup>],  $\beta$  is the soil slope [deg],  $k_s$  is the soil conductivity [m/s],  $L$  is the soil thickness [m],  $\tilde{\theta}$  is the mean reduced water content along the

vertical profile [-], and  $\alpha$  is a soil type dependent parameter [-].

The model is written in just one direction since it is assumed that the flow along the slopes is characterised by a preferential direction, which can be described as the direction of maximum slope.

From the combination of the two equations (3.8), considering the problem in terms of actual total water content in the soil,  $\eta = (\theta_s - \theta_r)L$  along the vertical profile, and making the following substitution:

$$C = \frac{Lk_s \tan(\beta)}{(\vartheta_s - \vartheta_r)^\alpha L^\alpha} \quad [\text{eq. 3.9}]$$

we obtain the following kinematic equation:

$$\frac{\partial \eta}{\partial t} = p - \frac{\partial q}{\partial x} = p - \frac{\partial (C\eta^\alpha)}{\partial x} \quad [\text{eq. 3.10}]$$

Where, from the physical point of view, the term C representing a local conductivity coefficient depends on soil parameters and is directly proportional to physically meaningful quantities such as hydraulic conductivity and slope, and inversely proportional to the storage capacity.

#### *Surface and channel flow*

Remembering the assumption for which all the precipitation infiltrates into the ground, the input to the surface water model is the precipitation excess resulting from the saturation of the soil layer. Furthermore, the flow in the soil can exfiltrate, thus also feeds the overland flow.

The description of overland flow is similar to the soil component, according to the kinematic approach (Wooding, 1965), in which the momentum equation is approximated by means of the Manning's formula. For a general cell, the kinematic wave approximations of

overland flow and channel flow are described in equations 3.11 and 3.12.

$$\begin{cases} \frac{\partial h_o}{\partial t} = r_o - \frac{\partial q_o}{\partial x} \\ q_o = \frac{1}{n_o} (\tan \beta)^{\frac{1}{2}} h_o^{\frac{5}{3}} = C_o h_o^{\alpha_o} \end{cases} \quad [\text{eq. 3.11}]$$

where  $h_o$  is the water depth over ground surface in m,  $r_o$  is the saturation excess resulting from the solution of the soil water balance equation, either as the precipitation excess or the ex-filtration from the soil in absence of rainfall, in m/s,  $n_o$  is the Manning friction coefficients for the surface roughness in  $\text{m}^{-1/3}\text{s}^{-1}$ ,  $C_o = (\tan \beta)^{1/2} / n_o$  is the coefficient of the Manning formula for overland flow,  $\alpha_o = 5/3$  is the exponent which derives from the use of Manning formula.

$$\begin{cases} \frac{\partial h_c}{\partial t} = r_c - \frac{\partial q_c}{\partial x} \\ q_c = \frac{1}{n_c} s_o^{\frac{1}{2}} h_c^{\frac{5}{3}} = C_c h_c^{\alpha_c} \end{cases} \quad [\text{eq. 3.12}]$$

where  $h_c$  is the water depth in the channel reach in m,  $r_c$  is the lateral drainage input, including the overland runoff reaching the channel reach and the soil drainage reaching the channel reach in m/s,  $s_o$  is the bed slope, assumed to be equal to the ground surface slope  $\tan(\beta)$ ,  $n_c$  is the Manning friction coefficients for the channel roughness in  $\text{m}^{-1/3}\text{s}^{-1}$ ,  $C_c = s_o^{1/2} / n_c$  is the coefficient relevant to the Manning formula for channel flow,  $\alpha_c = 5/3$  is the exponent which derives from the use of Manning formula.

### **3.4 Data structure**

According to the formulation described in the paragraph 3.3 different input data are required for each module.

In the paragraph 3.4.1 I will just shortly list them in order to understand how to design the logic model of the geo-database (described in the paragraph 3.4.2) that will be used by the hydrological model.

#### ***3.4.1 Basic input data***

The basic data, from which all the other needed input data can be derived, can be classified in three different groups: territorial data, meteorological data, and soil-land semantic data (i.e. hydrological information like literature Manning coefficient values for different landuses, vegetation parameter like the canopy height at given reference dates) herein referred to as “watershed characterizing data”.

##### *Territorial data*

They are the digital elevation model (DTM), the landuse, and the soiltype raster maps. These three maps will be used to derive other informations such as the water flow directions or the soil conductivities.

##### *Meteorological data*

This data set includes maps of air temperature, air humidity, wind speed, and rain intensity.

These information are usually acquired by using a network of meteorological stations and the relative maps are generated by means of spatial interpolation of the point-wise observations. Some of these information can also be extrapolated by remote sensing observation: such as, for instance, the thermal infrared satellites sensors for temperature like the MODIS Land Surface Temperature product (<http://modis-land.gsfc.nasa.gov/temp.htm>, March 8, 2006).

*Watershed characterizing data*

These data set describe the basin through indexes representing the characteristics of the system.

These indexes are:

- Leaf Area Index (LAI) – the one sided green leaf area for unit ground area in broadleaf canopies. It is generally used to predict the photosynthetic primary production of vegetation that drives the transpiration process: a higher LAI corresponds to a higher leaf mass and therefore a higher respiration rate of the crop. It's values, that can be directly observed, are generally derived from the literature (obtained from a statistically significant sample of plants from a crop) as a function of the vegetation type (landuse category) and of the crop growth (function of time and quota). Alternative indirect methods to obtain the LAI values are algorithms that derive this index from the satellite data, one example is the MODIS “Leaf Area Index/FPAR 8-Day L4 Global 1km” product (Knyazikhin et al., 1999).
- Vegetation cover – the percentage of a unit area covered by vegetation; this index is used to detect the areas influenced by vegetation dependent processes, like evapotranspiration or rainfall interception. It's values can be either be extracted by literature as a function of the vegetation type (landuse category) and of the crop growth (function of time and quota) or by algorithms that computes it from the normalized differential vegetation index (NDVI) derived from satellite data ([www.essc.psu.edu/frac\\_veg/methodology/](http://www.essc.psu.edu/frac_veg/methodology/), 2006).
- Vegetation height – the mean height of the vegetation for unit areas; this index is used to derive the “zero plane displacement height” (the level of zero wind) in vegetated areas that allows the determination of the aerodynamic resistance value (the transfer of heat and water vapour from the evaporating surface into the air above the canopy). It's values can be derived from literature values as a function of the vegetation type (landuse category) and of the crop growth (function of time quota) or from LIDAR (Light Detection and Ranging) observations (St-Onge and Achaichia,

2001).

- Albedo – a (unitless) measure of reflectivity of a surface or body. It is the ratio of total electromagnetic radiation (EM radiation) reflected to the total amount incident upon it, thus it represents the fraction of solar energy (shortwave radiation) reflected back into the space; Its values can be derived either from literature values as a function of the land typology (landuse category) or from satellite observations such as the MODIS “BRDF/Albedo Product” (Strahaler et al., 1999)
- Linke – an index of the atmospheric absorption and scattering of the solar radiation under clear skies. It describes the optical thickness of the atmosphere due to both the absorption by the water vapor and the absorption and scattering by the aerosol particles relative to a dry and clean atmosphere. It summarizes the turbidity of the atmosphere, and hence the attenuation of the direct beam solar radiation (WMO, 1981; Kasten, 1996). The larger TL, the larger the attenuation of the radiation by the clear sky atmosphere. Its values can be derived from literature as a function of lanuses and months, or derived from the SoDa “Service for Knowledge in Solar Radiation: Data, Databases, Applications, Education (SoDa: Solar Data)” databases ([http://www.soda-is.com/eng/services/climat\\_eng.html](http://www.soda-is.com/eng/services/climat_eng.html), 2006).
- Hydrological indexes (Manning, soil water contents, transmissivity exponent, horizontal permeability, and soil thickness) – they characterize from an hydrological point of view the soil. Theirs values can be measured “in situ” or derived from literature as a function of the soiltype category. Even if these indexes are physically based, their values can be different from the directly observed, in fact when applied for gridded areas, they assume a mean value over the cell, thus a calibration of these values is often suggested. Further informations on standard values can be found in section 6.8 dedicated to the case study.

As described above most of these data at a given time (date and hour) can be derived combining landuse and/or soiltype and/or elevation maps in conjunction with literature values of the

parameters (both monthly or seasonally known).

Other common sources for these parameters are remotely sensed data (e.g. Terra & Aqua MODIS Data Products, <http://edcdaac.usgs.gov/dataproducts.asp>, January 09, 2006). Those sources have the advantage of measuring, with a correct spatial distribution, the current values and therefore considering all the anomalies that can occur (like as an example the delay in crops growing due to a particularly dry season). The drawback is that they aren't direct measures and often their validity has still to be proven.

Once the needed basic data has been detected, the Geo-database logic model can be modelled.

### 3.4.2 Geo-database logic model

The GRASS geo-database is structured in three subsequent levels:

- GISDBASE – the directory housing all the geo-database;
- LOCATION – the directory housing a spatial domain;
- MAPSET – the directory housing a user project;

In each LOCATION a special MAPSET called PERMANENT is automatically generated; this is where the maps available to all users should be placed. Every user will store his own map in his own MAPSET, where other users have no access in writing mode.

Into each MAPSET, there are two main folder sets, one related to the raster maps, and the other related to the vectors maps geometry. All the attributes can be stored into an external database table or into an internal dbase (*dbf*) file.

From a logic point of view, hydrological modelling involves the usage of three kind of data sets: point-wise (climatic station observations), raster (distributed observations, extrapolations, or simulations), and attribute (literature parameters and general informations).

The adopted logic model of the data is represented in figure 3.3.

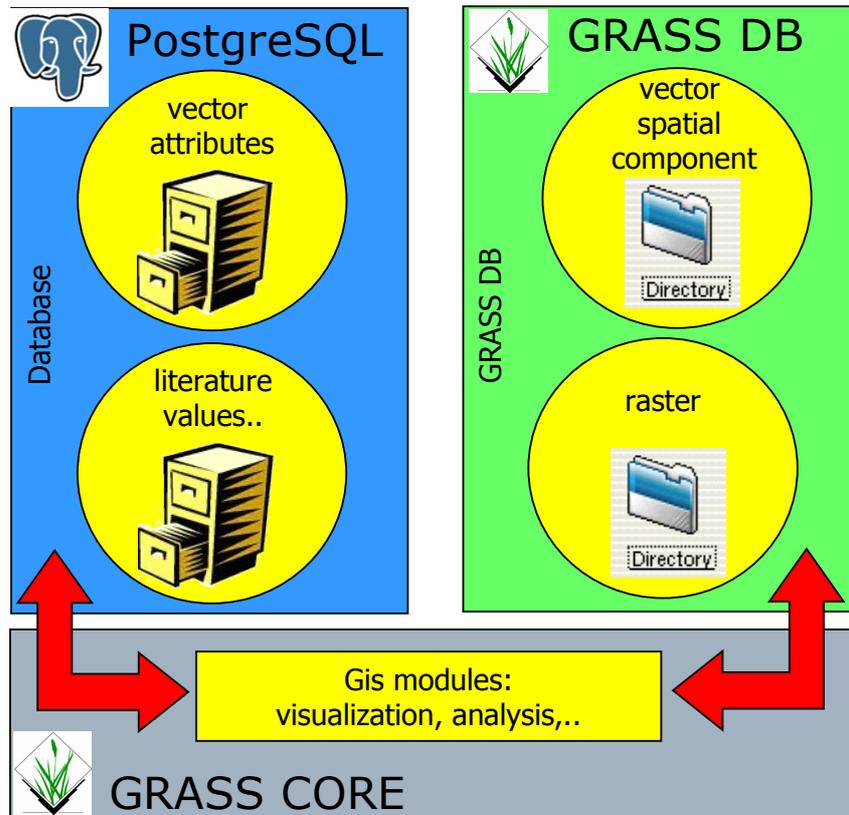


Figure 3.3 – logic data structure.

### 3.4.3 Geo-database conceptual model

The conceptual model of the geo-database is represented in figure 3.4 where three different areas can be detected: point data series, raster series and static rasters.

#### *Point data series:*

Each climatic station, owned by an organization, collects, at different time interval, observations of various parameter types. Maintenance informations are also considered by storing involved people and their operations.

#### *Raster series:*

Some distributed values are dynamic and therefore have different representation for different time: for example the rainfall field

typically change over space and over time.

*Static rasters*

Some distributed values are “static” (DEM, landuse, soiltype,etc.) and some of their classes could be related to monthly or seasonal parameters.

From a spatial point of view the climatic stations represented with the point primitive and the distributed values (series and static) are represented with rasters.

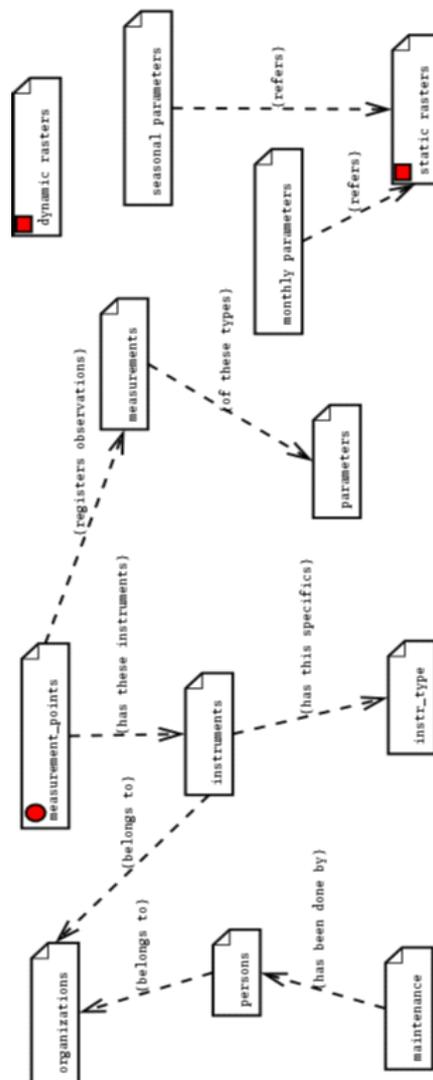


Figure 3.4 – Logic data model, red squares indicate raster data and red point indicate point data

3.4.4 Geo-database physical model

The final aim of an hydrological model is to represent water fluxes over a spatio-temporal space: this leads to the establishment of procedures to correctly represent this spatial and temporal dimensions.

*Point data series*

Vector data in GRASS are represented by a geometry native file linked with a table of attributes through a unique feature identification number (cat) present both in the file and in the table.

The spatial component of point data series (measurements stations) is fixed and therefore it should be stored as a static geometry native file in GRASS while their attributes (observed values) at a given time should be dynamically linked to the geometry.

Unfortunately at this time this option can not be realized: in fact GRASS data structure allows to establish a link only between a geometry file and an attribute table, but not with a selection of it.

This lack leads, for each considered time instant, to the generation of a new vector files which will be used for the interpolation process and then erased: obviously this is not the optimal procedure and a modification of the GRASS vector structure is desirable in order to handle time series data.

*Raster series*

Also the distributed values change and the surface describing variables is different time to time (i.e. temperature field).

Conceptually a three dimensional (2D in space + 1D in time) distributed information could be well represented by voxels; the drawback of this possible choice is that the file dimension could become huge (around 8700 levels a years are needed) and numerous operations of cutting (computationally expensive), aimed to extract information at a time instant, could be required.

Usually in hydrological models time is represented by consecutive steps: like photograms in a movie, ASCII files are used to store the status of the system at consecutive instants.

Transferring this solution into GISs, series of rasters should be used by find out a method to relate the raster files with the respective time instants.

Even if GRASS handle time information of a layer (vector or raster) by writing/reading a timestamp in a specific file (Neteler M., 2005) I decided to adopt a more direct and fast approach by including the time information into the raster name suffix.

The suffix is composed by a 12 digit number representing:

- year (1<sup>st</sup> to 4<sup>th</sup> digit);
- month (5<sup>th</sup> to 6<sup>th</sup> digit);
- day (7<sup>th</sup> to 8<sup>th</sup> digit);
- hour (9<sup>th</sup> to 10<sup>th</sup> digit);
- minute (11<sup>th</sup> to 12<sup>th</sup> digit).

Therefore the raster representing the rainfall on the 12<sup>th</sup> of August 2005 at the 17:00 will be named "rain.20050812170".

### *Static maps*

Not dynamic features can easily be handled by common GIS vector or raster files and the monthly or seasonally known parameters values can be stored in database tables. The relationship between rasters and parameters can be done by using raster category values as a key field of the attributes tables.

## 4. System implementation

In this chapter I present the solution adopted and the modules development that bring to the model realization.

The programming languages used are the GRASS native languages: the C ANSI for implementing new modules and the Bash Shell for scripting.

The resulting model, aggregation of different procedures and modules called in the right sequence, will be herein referred to as HydroFOSS (Hydrological Free and Open Source Software).

### 4.1 Catchment maps generation

A distribute model need distribute inputs.

Distribute inputs can be derived from distributed observations (i.e. satellite images) or from spazialization of point wise observations (i.e. climatic stations) or literature values (i.e. leaf area index).

Spazialization of point wise observations and literature values can be achieved applying different methods.

#### *4.1.1 Selected approach*

In the present work I adopted the classical approach (the most diffuse and tested approach in hydrology modelling) that uses point wise observations mainly regionalized by means of Thiessen tassellation and literature values distributed on the basis of landuse and soiltype classes.

I performed a more detailed investigation, presented in the next chapter 5, only for the analysis of distributed rainfall estimates obtained from the meteorological radar, and for the investigation of the best method for temperature interpolation.

These analyses reveal that nowadays radar rainfall estimates in alpine regions seem to be still too inaccurate to be used with confidence, and that the best method for the temperature

interpolation is a simple linear gradient.

Further studies should be carried on in future works.

A general consideration is that even if the catchment distributed maps can be both derived from remote sensing, this technology presents at the moment the following limitation:

- it is not yet commonly validated, being a quite new object in hydrology research;
- it is hardly available at hourly time step;
- it is often expensive for high resolutions;

Anyway, I consider that distributed hydrological models can take full advantage of remote sensing observations that are able to well describe the spatial distribution of variables. Thus, with the aim to validate remote sensing data in future works, I designed a model that allows their usage.

The adopted modular approach, in combination with the already existent GRASS commands for remote sensing data, guarantees the sufficient flexibility to use remote sensed data within the HydrFOSS model.

Some of the existent GRASS command for remote sensing are:

- `r.in.gdal`, `r.in.asci`, `r.in.bin`, `r.in.erdas`, for importing;
- `r.proj`, `i.group`, `i.target`, `i.points`, `i.rectify` for geocoding;
- `i.tm.dehaze`, `r.mapcalc`, `i.cluster`, `i.maxlik`, `i.class`, `i.gensig`, `i.smap`, for radiometric processing and classification.

### 4.1.2 Climatic map generation

In order to proceed with the estimation of the evapotranspiration and of the snowmelt, a number of distributed map (representing climatic variables fields) has to be derived. Two steps have to be done in order to proceed with the implementation:

1. generation of a meteorological database to archive meteorological stations observations;
2. application of spatial interpolations.

*Meteorological network stations database (MNSD)*

In the chosen approach the sampled data are point-wise observations of climatic stations. The database schema that has been adopted (following the logical model shown in paragraph 3.4.3) is represented by eight tables as represented in the relational model of figure 4.1. It has to be noted that this database is not only functional to the model development but also to the network maintenance.

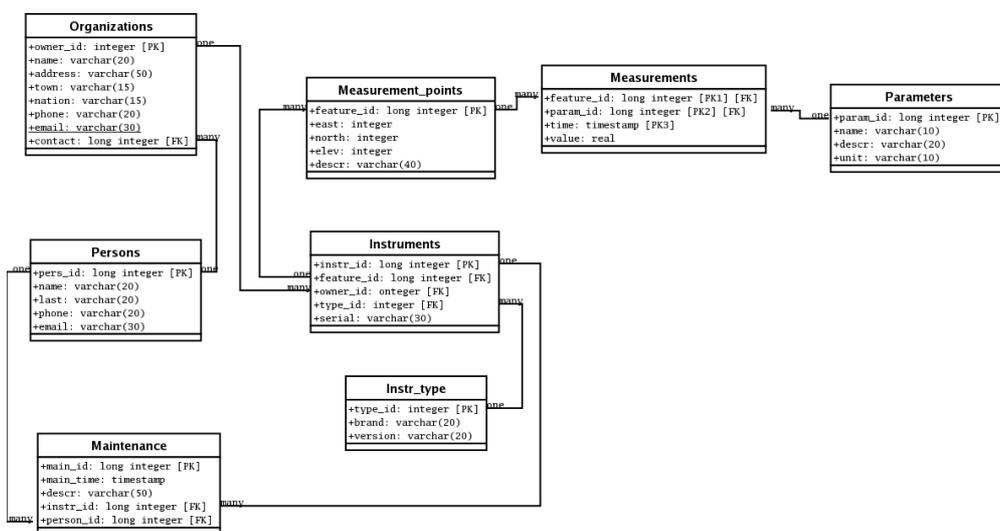


Figure 4.1 – Relational model of the point wise observations

In Appendix A.0 a detailed description of each table is presented.

*Data interpolation*

Climatological maps derivation results from the spatial interpolation of measurements (temperature, humidity, solar radiation, wind velocity, etc.) made at control points (meteorological stations). Unfortunately these observation points are few and far from each other while the investigated phenomenon has a high spatial heterogeneity: as a result the correct field interpretation results difficult.

For the precipitation and other meteorological variables interpolation, the Thiessen’s polygons are still applied in many

applications (e.g. Topkapi, WaSiM, Mike-SHE, and CASC2D): Inverse Distance Squared Weighting (IDW) or Thiessen tessellation (which is a particular IDW with number of interpolation points equal to 1) are both available in GRASS through the *s.surf.idw* command. The commonly adopted procedures, for meteorological variables interpolation, have therefore been applied without any deeper investigation except for the temperature (see chapter 5).

I think that, due to the high variability in space of the variables (wind, rain, temperature, humidity, and solar radiation) and the generally low spatial density of sampling, the improvements derived by using more complex interpolation method (splines, kriging, co-krigin, etc.) is often low compared to the increasing computational cost.

The implemented procedure for meteorological raster maps derivation consists in the access of the MNSD, in the extraction of the desired informations and in the interpolation of the observations by means of statistical, physical or mathematical modelling.

The database access and data extraction has been done by using the *db.connect*, *db.select*, and *v.in.ascii* GRASS commands.

- *db.connect* enables the set up of the parameters needed by GRASS for the connection to the databases;
- *db.select* allows to execute SQL query (read from input file or from standard input) printing the result to standard output;
- *v.in.ascii* imports a vector map in ASCII format (result of the SQL selection) into current GRASS database.

For wind [m/s], rain [mm/h] and humidity [%] the Thiessen tassellation is applied while for solar radiation [ $\text{W}\cdot\text{m}^{-2}\cdot\text{h}^{-1}$ ] an IDW interpolation is used, in order to avoid undesirable breakline, and for temperature a linear gradient (chapter 5) is applied.

As an example the command's sequence to generate distributed rain map for the 15<sup>th</sup> March 2005 at 13:00 is illustrated in frame 4.1.

## Chapter 4 - System implementation

```
L1> #setting of connection with MNSD database
L2> db.connect driver=pg database=MNSD schema=""
group="";
L3> #extraction and generation of the vector of stations
and values
L4> echo "SELECT east, north, elev, value FROM
Measurement_points as p, Measurement as v WHERE
p.feature_id=v.feature_id AND v.time='15-03-2005 13:00'
AND v.param_id=3" | db.select | v.in.ascii
out=rain.200503151300 x=1 y=2 z=3 cat=0 format=point;
L5> #generation of the distributed map
L6> v.surf.idw -n input=rain.200503151300
output=rain.200503151300 npoints=1 col=dbl_1;
L7> #removing of the vector no more needed
L8> g.remove vect= rain.200503151300;
```

Frame 4.1 – procedure for rain map derivation.

Advantages of the GIS-embedded approach can be clearly detected, in fact in order to change the interpolation method it will be enough to change line 6 (L6) of frame 4.1 by using different GRASS interpolation commands.

For example, the usage of spline interpolation can be easily achieved by substituting L6 with the following line:

```
L6> v.surf.rst input=rain.200503151300 zcolumn=dbl_1;
```

As discussed in the previous chapter this procedure involves the generation of a new vector file for each processing instant: even if the file is later erased and therefore does not increase the geodatabase size, it is certainly a computationally expensive process.

### 4.1.3 Watershed characterization maps generation

Meteorological distributed data (precipitation, temperature, wind velocity, global radiation, etc.) are able to account for climatological variability inside the catchment. Moreover land use, soil type, and topography are required to completely parametrize the catchment.

In fact to model the hydrological processes of the catchment some watershed characterization maps have to be derived: indexes distribution mainly based on land use, soil properties and topography.

For the radiation exchange between the earth surface and the atmosphere, the albedo ( $a$ ) of different surfaces and the Linke Turbidity factor ( $Ltf$ ) are required.

For the evapotranspiration process the status of the vegetation, in terms of leaf area index ( $LAI$ ), height ( $Vh$ ), and density ( $Vc$ ) must be considered.

The solutions chosen for the creation of the maps of  $LAI$ ,  $Vh$  and  $Vc$ , according to the WaSiM model approach (Schulla, 1997), calculate the raster map on the basis of values associated with four specific days of the year and relative to a reference height of 400 m a.s.l. (from  $[d1,400]$  to  $[d2,400]$  of the growing season and from  $[d3,400]$  to  $[d4,400]$  of the dormant season) and relative to different land uses.

Keeping in mind the existence of a temporal shift as a function of the altitude, for each cell of the model with its elevation it is possible to recalculate the reference days:

$$d_{i,h} = d_{i,400} + 0.025 \cdot (h - 400) \quad [\text{eq. 4.1}]$$

being  $i = [1,2,3,4]$  the index of the four reference days and  $h$  the height above the sea level of the cell.

By means of the linear interpolation through time it is possible to compute the value assumed by  $LAI$ ,  $Vh$  and  $Vc$  cell by cell.

This algorithm has been implemented in the new GRASS command

## Chapter 4 - System implementation

*h.seasonal.par* that – given a date, the four Julian days of reference ( $j1, j2, j3, j4$ ), the elevation model and the table that archives the parameters reference values – computes all the desired maps.

The algorithm, implemented in a shell script language, reads the input values from a seven columns table:

- column 1 – named PAR contains the parameter name;
- column 2 – named RASTER contains the reference raster map name from which different values depends;
- column 3 – named CAT contains the category values of the reference raster map;
- column 4 – named DAY1 contains the values of the parameters at the reference day  $j1$  for the current row PAR, RASTER and CAT values;
- column 5 – named DAY2 is the same of column 4 but referred to the day  $j2$ ;
- column 6 – named DAY3 is the same of column 4 but referred to the day  $j3$ ;
- column 7 – named DAY4 is the same of column 4 but referred to the day  $j4$ .

Then for each different parameter in the table (column PAR) it calculates four raster maps by reclassifying the reference raster at the four reference days, for each of them applies the equation 4.1 at each cells, and finally, based on the input date, calculates the desired parameter map with a linear interpolation through time.

On the other hand,  $a$  and  $Ltf$  do not undergo to any variation with elevation and their mean monthly values are available in the literature according to the different land uses. In this case the maps do not depend on the elevation, but only on the date and on the land use.

The derivation of these maps are computed by means of the new developed GRASS command (Shell script language) *h.monthly.par* that reads the parameters values from a 15 columns table:

- column 1 – named PAR contains the parameter name;
- column 2 – named RASTER contains the reference raster map name from which different values depends;
- column 3 – named CAT contains the category values of the reference raster map;
- column 4 to 15 – named JEN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SET, OCT, NOV, and DEC each one respectively contains the values of the parameters for the twelve month relative to the current row PAR, RASTER and CAT values;

Then for each different parameter in the table (column PAR), it calculates the reclassified map based on the month extracted from the input date.

### **4.2 Net radiation map generation**

The net radiation ( $Rn$ ) is the difference between the incoming net shortwave radiation ( $Rns$ ) emitted by the sun and the outgoing net longwave radiation ( $Rnl$ ) emitted by the earth.

#### 4.2.1 Net shortwave radiation ( $Rns$ )

Following Hofierka and Suri (2002) the computation of the  $Rns$  has been done in four steps:

- calculation of the net shortwave radiation components in clear-sky condition. The existing *r.sun* command of GRASS computes direct (beam), diffuse and reflected solar irradiation raster maps for a given day, latitude, surface and atmospheric conditions. Optionally a local time can be specified to compute solar incidence angle and/or irradiance raster maps and the shadowing effect of the topography can be incorporated (slope and aspect input map can be derived from the *r.slope.aspect* command);
- Calculation of the total net shortwave radiation in clear-sky condition ( $Rns_k$ ). The sum of the direct (beam), diffuse and reflected solar irradiation raster maps computes the total energy

irradiation map;

- calculation of a clear-sky index ( $K$ ) map. This is an index of cloudiness that can be extracted from the interpolation of the ratio between the estimated  $R_{ns_k}$  and the radiation measured at meteorological stations points; here the *v.sample* command is used to extract the  $R_{ns_k}$  values, *db.execute* to calculate the  $K$  index, and the *v.surf.idw* to interpolate  $K$  and obtaining the distributed  $K$  field.
- calculation of the effective  $R_{ns}$  by multiplying the  $R_{ns_k}$  in clear-sky condition with the  $(1-K)$  factor by means of the map-algebra command *r.mapcalc*.

The code to calculate the  $R_{ns}$  map uses already existing GRASS commands as shown in frame 4.2.

## Chapter 4 - System implementation

```
L1> #calculation of Rsnk [W/m2]
L2> r.sun -s elevin=dem aspin=aspect slopein=slope lin=2
albedo=alb_Mar incidout=RADinc.200503151300
beam_rad=RADbeam.200503151300
diff_rad=RADdiff.200503151300
refl_rad=RADrefl.200503151300 day=73 time=13:00
dist=100;
L3> #GLOBAL Rsnk [W/m2]
L4> r.mapcalc
'Rnsk.200503151300=RADbeam.200503151300+RADdiff.200503151
300+RADrefl.200503151300';
L5> #setting of connection with MNSD database
L6> db.connect driver=pg database=MNSD schema=""
group="";
L7> #extraction and generation of the vector of stations
and values
L8> echo "SELECT east, north, elev, value FROM
Measurement_points as p, Measurement as v WHERE
p.feature_id=v.feature_id AND v.time='15-03-2005 13:00'
AND v.param_id=6" | db.select | v.in.ascii
out=RnskV.200503151300 x=1 y=2 z=3 cat=0 format=point;
L9> #sampling of the Rnsk (clear-sky condition)
L10> v.sample input=RnskV.200503151300 column=dbl_1
output=clear_sky.200503151300 rast=Rnk.200503151300;
L11> #removing no more used vector
L12> g.remove vect=RnskV.200503151300;
L13> #calculate the K index at sample points
L14> echo "UPDATE clear_sky.200503151300 SET diff = 0" |
db.execute;
L15> echo "UPDATE clear_sky.200503151300 SET diff =
(pnt_val/rast_val) where rast_val<>0" | db.execute;
L16> #interpolate K points and get K raster map
L17> v.surf.idw -n input=clear_sky.200503151300
output=clear_sky.200503151300 npoints=4 col=diff
```

```
L18> #removing no more used vector
L19> g.remove vect=clear_sky.200503151300
L20> #calculation of Rns [W/m2]
L21> r.mapcalc 'Rns.200503151300 =Rnk.200503151300 * (1-
clear_sky.200503151300)'
```

Frame 4.2 – procedure for net solar radiation map calculation.

#### 4.2.2 Net longwave radiation (Rnl)

Following Allen (1998), the longwave radiation can be computed in relation to the surface temperature by the Stefan-Boltzman law corrected with a humidity and a cloudiness factor:

$$Rnl = 2.043E-10 \cdot [(T+273.16)^4 \cdot (0.34 - \sqrt{ea}) \cdot (1.35 \cdot (1-K) - 0.35)]$$

[eq. 4.2]

with:

Rnl = net longwave radiation [MJ/(m<sup>2</sup>·h)];

T = surface temperature [°C];

ea = actual vapour pressure [KPa];

K = clear-sky index [-];

The actual vapour pressure can be computed as:

$$e0 = 0.618 \cdot \exp[(12.27 \cdot T)/(T+237.3)]$$

$$ea = (e0 \cdot h)/100$$

[eq. 4.3]

where:

e0 = saturation vapour pressure [KPa];

ea = actual vapour pressure [KPa];

T = air temperature [°C];

h = relative humidity [%];

## Chapter 4 - System implementation

The procedure to calculate the net longwave radiation makes extensive use of the map algebra by means of the *r.mapcalc* command as shown in frame 4.3.

```
L1> #saturation vapour pressure [KPa]
L2> r.mapcalc 'e0=0.6108*
exp((17.27*temp.200503151300)/(temp.200503151300
+237.3))';
L3> #actual vapour pressure[KPa]
L4> r.mapcalc 'ea=e0*humid.200503151300/100';
L5> #calculation of Rn1 [MJ/(m2*h)]
L6> r.mapcalc 'Rn1.200503151300 =2.043e-
10*((temp.200503151300+273.16)^4)*(0.34-
0.14*sqrt(ea))*(1.35*(1-clear_sky.200503151300
)-0.35)';
L7>
```

Frame 4.3 – procedure for the long wave radiation map calculation

### *4.2.3 Global radiation (Rn)*

The global net radiation, resulting from the incoming and outgoing balance, can be computed by a simple map algebra operation (see frame 4.4) where Rns is multiplied for 0.0036 in order to convert W in MJ.

```
L1> #actual net radiation [MJ/(m2*h)]
L2> r.mapcalc 'Rn.200503151300 =(Rns.200503151300
*0.0036)-Rn1.200503151300';
```

Frame 4.4 – procedure for the global radiation computation

The entire procedure has been schematized in figure 4.2.

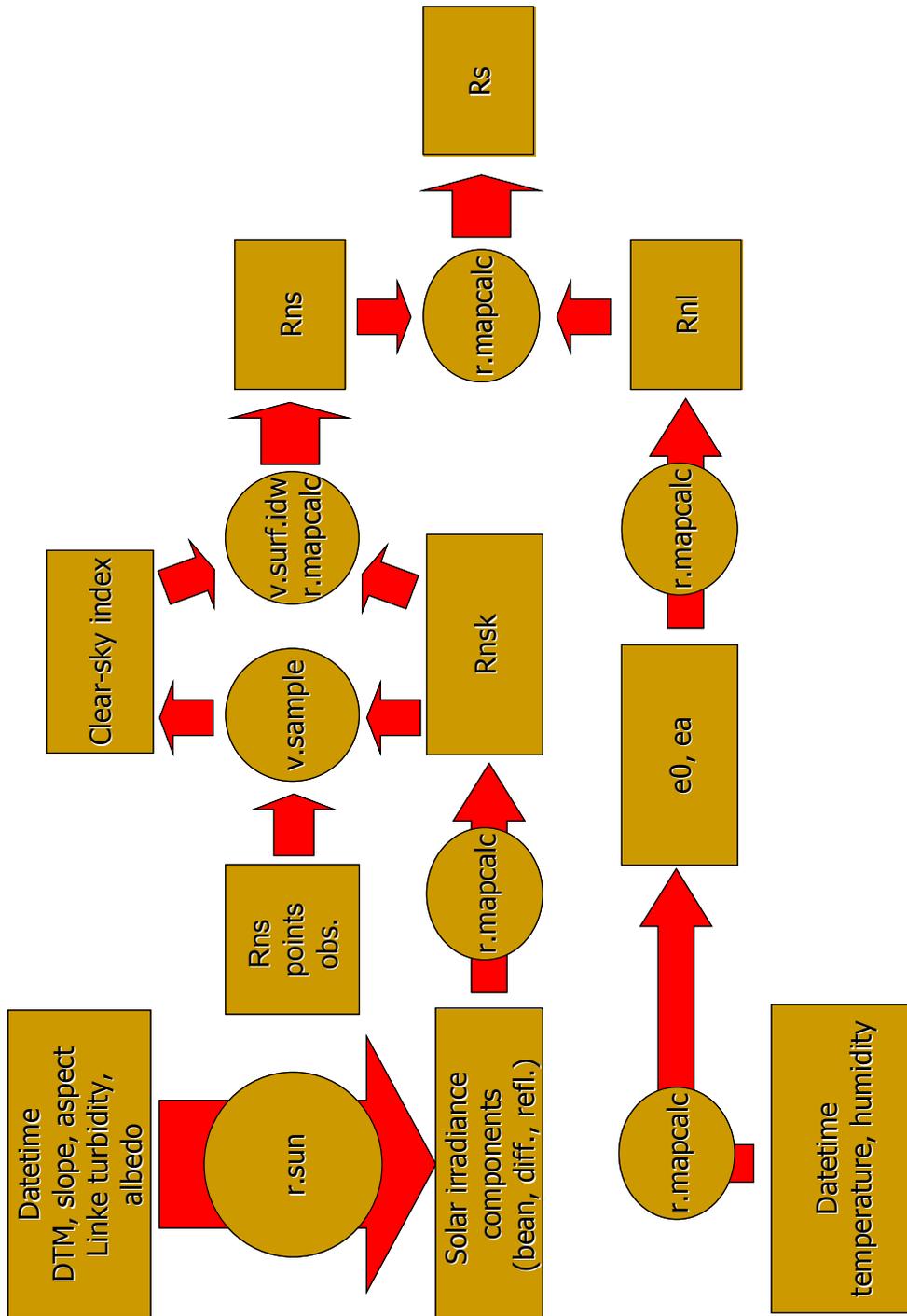


Figure 4.2 – model of the Net Radiation map calculation.

### 4.3 Snow accumulation and melt maps generation

The equations described in paragraph 3.3.3 have been implemented in a new GRASS command: *h.snow*.

At first the algorithm estimates the liquid fraction ( $Fr$ ) of the precipitation at a given air temperature ( $T$ ):

$$\begin{aligned} Fr &= 1 && \text{if } T > Tr; \\ Fr &= (T + Tb) / (Tr + 1) && \text{if } Tb < T < Tr; \\ Fr &= 0 && \text{if } T < Tb; \end{aligned} \quad [\text{eq. 4.4}]$$

and calculates the solid ( $Ps$ ) and liquid ( $Pr$ ) parts:

$$\begin{aligned} Ps &= R \cdot (1 - Fr); \\ Pr &= R \cdot Fr; \end{aligned} \quad [\text{eq. 4.5}]$$

whit  $R$ =rainfall [mm/h],  $Tr=3^\circ\text{C}$  and  $Tb=-1^\circ\text{C}$ .

Then it calculates the advected heat from precipitation ( $Qp$ ) with the equation:

$$\begin{aligned} Qp &= 0.001 \cdot \{ [Ps \cdot Cs \cdot \rho_w \cdot \min(T, 0)] + \\ &+ [Pr \cdot (hf \cdot \rho_w + (C_w \cdot \rho_w \cdot \max(T, 0)))] \}; \end{aligned} \quad [\text{eq. 4.6}]$$

where  $\rho_w$  is the water density [ $1000 \text{ Kg} \cdot \text{m}^{-3}$ ];  $Cs$  the specific heat of the ice [ $2.09 \text{ KJ} \cdot \text{Kg}^{-1} \cdot ^\circ\text{C}^{-1}$ ];  $Ps$  the solid part of the precipitation [mm/h];  $Pr$  the liquid part of the precipitation [mm/h];  $hf$  the heat of fusion [ $333.5 \text{ KJ} \cdot \text{Kg}^{-1}$ ]; and  $C_w$  the specific heat of water [ $4.18 \text{ KJ} \cdot \text{Kg}^{-1} \cdot ^\circ\text{C}^{-1}$ ].

With this estimation ( $Qp$ ) it is now possible to solve the equations 3.6 and estimates the new snowpack water equivalent  $Z_{t+\Delta t}^*$  and the needed energy to maintain the snowpack in a solid phase  $E_{t+\Delta t}^*$  in a zero snowmelt hypothesis.

## Chapter 4 - System implementation

The energy needed to maintain the snowpack in a solid phase  $E_n$  can be calculated by means of:

$$E_n = C_s \cdot Z^*(t+\Delta t) \cdot T_0; \quad [\text{eq. 4.7}]$$

where  $T_0$  is the ice temperature [273.5 K];

Now it is possible to evaluate if the available energy is enough to melt part of the snowpack by comparing  $E_n$  with  $E^*$ .

The algorithm calculates the snowmelt ( $Sm$ ), the snowpack water equivalent  $Z_{t+\Delta t}$ , and the snowpack energy content  $E_{t+\Delta t}$  for each cell by setting:

- $Sm=0$ ,  $Z_{t+\Delta t}=0$  and  $E_{t+\Delta t}=0$  where there is no snow ( $Z^*_{t+\Delta t}=0$ );
- $Sm=0$ ,  $Z_{t+\Delta t}=Z^*_{t+\Delta t}$ , and  $E_{t+\Delta t}=E^*_{t+\Delta t}$  where  $E_n \geq E^*_{t+\Delta t}$  (the available energy is not sufficient to melt part of the snowpack);
- $Sm = \min((E^*_{t+\Delta t} - E_n) / C_f, Z^*_{t+\Delta t})$ ,  $Z_{t+\Delta t} = \max(0, Z^*_{t+\Delta t} - Sm)$ , and  $E_{t+\Delta t} = E^*_{t+\Delta t} - ((C_s \cdot T_0 + hf) \cdot Sm)$  where  $E_n < E^*_{t+\Delta t}$  (the available energy is sufficient to melt part of the snowpack);

The new maps of  $Sm$ ,  $Z_{t+\Delta t}$ , and  $E_{t+\Delta t}$  are finally generated. The input-output schema of the *h.snow* command is shown in figure 4.3.

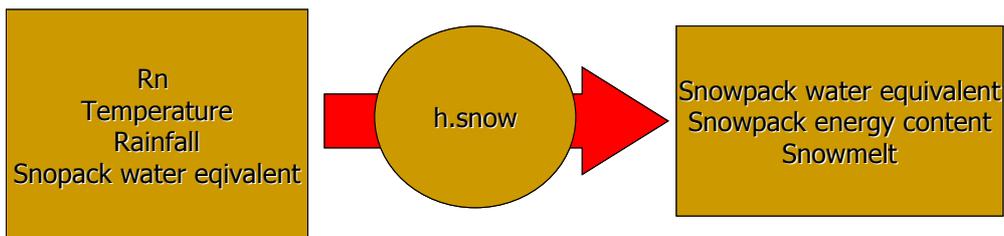


Figure 4.3 – input-output schema of the *h.snow* command.

#### 4.4 Potential evapotranspiration map generation

Following the approach described in paragraph 3.3.1, I developed the new GRASS command *h.evapo.PM* that, given the vegetation height ( $hc$ ), humidity ( $Rh$ ), wind speed at two meters height ( $u_2$ ), temperature ( $T$ ), digital terrain model ( $Z$ ), and net radiation ( $R_n$ ) raster input maps, calculates the potential evapotranspiration map. The implemented algorithm applies the open water calculation (equation 3.2) where the vegetation height is negative and the Penman-Monteith (equation 3.1) elsewhere.

The equation 3.1 and 3.2 can be solved after the computation of the single variables found in those formulas:

- mean saturation vapour pressure  $e_a$  [KPa]:  

$$e_a = 0.61078 \cdot \exp[(17.27 \cdot T)/(T+237.3)];$$
 [eq. 4.8]

- slope of vapour pressure curve  $\Delta$  [Kpa/°C]:  

$$\Delta = (4098 \cdot e_a)/(237.3+T)^2;$$
 [eq. 4.9]

- latent heat vaporisation  $\lambda$  [MJ/kg]:  

$$\lambda = 2.501 - (0.002361 \cdot T);$$
 [eq. 4.10]

- atmospheric pressure  $P$  [KPa]:  

$$P = P_0 \cdot [(T_{k0} - \eta \cdot (Z - Z_0))/T_{k0}]^{g/(\eta \cdot R)};$$
 [eq. 4.11]

- psychrometric constant  $\gamma$  [KPa/°C]:  

$$\gamma = 0.001 \cdot [(c_p \cdot P)/(\epsilon \cdot \lambda)];$$
 [eq. 4.12]

- aerodynamic resistance  $ra$  [s/m]:  
 where the vegetation height ( $hc$ ) is less than 2m:  
 $d = (2/3) \cdot hc; Z_{om} = 0.123 \cdot hc; Z_{oh} = 0.1 \cdot Z_{om};$   

$$ra = [\log((Z_w - d)/Z_{om}) \cdot \log((Z_h - d)/Z_{oh})] / (k^2 \cdot u_2);$$
 [eq. 4.13]

## Chapter 4 - System implementation

elsewhere:

$$\begin{aligned}u_{10} &= u_2 \cdot (\log((67.8 \cdot 10) - 5.42)) / 4.87; \\ r_a &= 94 / u_{10};\end{aligned}\quad [\text{eq. 4.14}]$$

- surface resistance  $r_s$  [s/m]:

$$r_s = 100 / (0.5 \cdot 24 \cdot h_c); \quad [\text{eq. 4.15}]$$

- modified psychrometric constant  $\gamma^*$  [KPa/°C]:

$$\gamma^* = \gamma \cdot (1 + (r_s / r_a)); \quad [\text{eq. 4.16}]$$

- soil heat flux  $G$  [MJ/m<sup>2</sup>\*d]:

$$\text{during the day:} \quad G = 0.1 \cdot R_n; \quad [\text{eq. 4.17}]$$

$$\text{during the night:} \quad G = 0.5 \cdot R_n; \quad [\text{eq. 4.18}]$$

- actual vapour pressure  $e_d$  [KPa]:

$$e_d = R_h \cdot e_a / 100; \quad [\text{eq. 4.19}]$$

- virtual temperature  $T_{kv}$  [°C]:

$$T_{kv} = (T + 273.15) / (1 - (0.378 \cdot e_d / P)); \quad [\text{eq. 4.20}]$$

- atmospheric density  $\rho$  [Kg/m<sup>3</sup>]:

$$\rho = P / (T_{kv} \cdot R / 100); \quad [\text{eq. 4.21}]$$

where:

$c_p$  is the specific heat of moist air [1.013 KJ·Kg<sup>-1</sup>·°C<sup>-1</sup>];

$\epsilon$  is the ratio of molecular weight of water to dry air [0.622];

$P_0$  is the atmospheric pressure at sea level [101.3 KPa];

$T_{k0}$  is the reference temperature at sea level [293.16 K];

$\eta$  is the constant lapse rate [0.0065 K/m];

$Z_0$  is the altitude at sea level [0 m];

$Z$  is the elevation above  $Z_0$ ;

## Chapter 4 - System implementation

$g$  is the gravitational acceleration [9.81 m/s];  
 $R$  is the specific gas constant [287 J·kg<sup>-1</sup>·K<sup>-1</sup>];  
 $Z_w$  is the height of wind measurements [2 m];  
 $Z_h$  is the height of humidity measurements [2 m];  
 $k$  is the Von Karman constant [0.41];  
 $d$  is the zero plane displacement of wind profile [m];  
 $U_z$  is the windspeed at height  $z$  [m/s];  
 $Z_{om}$  is the roughness parameter for momentum [m];  
 $Z_{oh}$  is the roughness parameter for heat and water vapour [m];

Optionally the user can activate a flag (-z) that allows him setting to zero all of the negative evapotranspiration cells; in fact these negative values motivated by the condensation of the air water vapour content, are sometime undesired because they can produce computational problems.

The usage of the flag -d detect that the module is run in day hours and the appropriate soil heat flux is calculated.

The input-output schema of the *h.evapo.PM* command is shown in figure 4.4.

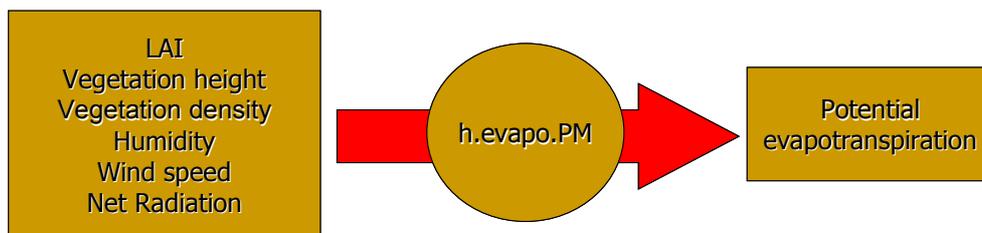


Figure 4.4 – input output schema of the *h.evapo.PM* command.

### 4.5 Canopy interception maps generation

The canopy interception mechanism is described in the paragraph 3.3.2. In order to proceed with the interception losses estimation, the potential evaporation from a liquid surface ( $ETP_{ls}$ ) has to be known. Thus the previously described command *h.evapo.PM* has to be run with a “virtual” map of the vegetation height set to negative values. Given the LAI [-], the  $ETP_{ls}$  [mm/h], the leaf storage level at the previous step  $C_0$  [mm], the rain  $R$  [mm/h], the timestep  $T_s$  [h], and the vegetation density  $V_c$  [%] the command estimates the canopy storage maximum capacity  $C_m$  [mm]:

$$C_m = 0.35 \cdot LAI \quad [\text{eq. 4.22}]$$

and the needed time to evaporate a saturated canopy  $\tau$  [h]:

$$\tau = C_m / ETP_{ls}; \quad [\text{eq. 4.23}]$$

if the  $ETP_{ls}$  is zero, in order to avoid division by zero it is simply set to a large value:

$$\tau = T_s \cdot 1000 \quad [\text{eq. 4.24}]$$

As Zengh (2000) discussed, the storage level dynamic during a rainfall event can be splitted in three different phases (see figure 4.5):

- A) filling – it is raining and the canopy storage is not saturated, thus the storage level is growing;
- B) saturation – it is raining and the canopy storage has reach the maximum capacity, thus its level remain at the maximum;
- C) drying – it is not raining and the canopy storage is not empty, thus the level is going down.

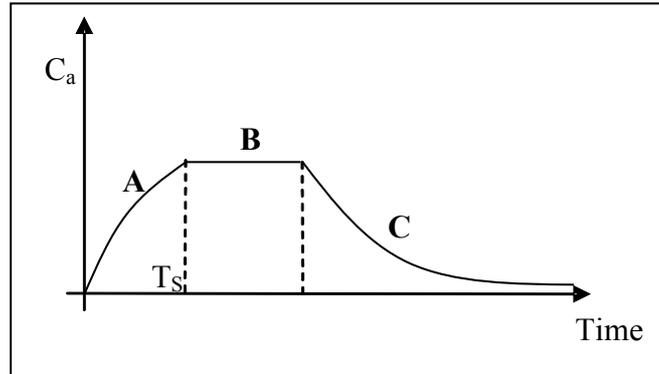


Figure 4.5 – canopy storage level dynamic during storm: filling (A), saturation (B), and drying (C).

Depending on the storage level and rain duration/intensity during a storm all the three phases (A+B+C, if saturation is reached) or only two of them (A+C, if saturation is not reached) can occur.

As a function of the phases taking place into the single time step ( $t_0-t$ ), different solutions of the equations 3.3, 3.4, and 3.5 has to be applied.

Thus the algorithm at first, being  $C_0$  the storage level at  $t_0$ , computes the rain needed to saturate the storage at the end of the interval ( $R_{sat}$ ):

$$R_{sat} = (C_0 - C_m) / [\tau \cdot (1 - \exp(-T_s / \tau))] \quad [\text{eq. 4.25}]$$

and the time when saturation is reached (dt):

$$dt = -\tau \cdot \log[1 - ((C_m - C_0) / (R \cdot \tau))] \quad [\text{eq. 4.26}]$$

Then it compares  $R$  [mm/h],  $ETP_{ls}$  [mm/h],  $C_0$  [mm],  $C_m$  [mm],  $R_{sat}$  [mm/h], and  $dt$  [h] evaluating 5 different cases.

For each occurring case the algorithm estimates the interception losses  $IL$  [mm], the drainage  $D$  [mm] and the canopy storage level at the end of the step  $C_t$  [mm] as follow:

- Case 1 – only saturation phase ( $R > ETP_{ls}$  and  $C_0 \geq C_m$ ):

$$IL = \min[(ETP_{ls} \cdot Ts), (R \cdot Ts) + (C_0 - C_m)]$$

$$D = R \cdot Ts - IL + (C_0 - C_m)$$

$$C_t = C_c \quad [eq. 4.27]$$

- Case 2 – only filling phase ( $R > ETP_{ls}$  and  $C_0 \leq C_m$  and  $R_{sat} > R$ ):

$$C_t = R \cdot \tau \cdot (1 - \exp(-Ts/\tau)) + d \cdot W_0;$$

$$IL = (R \cdot Ts) - (C_t - C_0);$$

$$D = 0; \quad [eq. 4.28]$$

- Case 3 – mixed filling and saturation phase ( $R > ETP_{ls}$  and  $C_0 \leq C_m$  and  $R_{sat} \leq R$ ):

$$IL = (R \cdot Ts) - (C_m - C_0) + (Ts - dt) \cdot ETP_{ls}$$

$$D = (R \cdot Ts) - (C_m - C_0) - IL;$$

$$C_t = C_m \quad [eq. 4.29]$$

- Case 4 – only drying phase ( $0 < R < ETP_{ls}$ ) or ( $R = 0$  and  $C_0 \leq C_m$ ):

$$D = 0;$$

$$C_t = \max(W_0 \cdot \exp(-Ts/\tau), 0);$$

$$IL = (C_0 - C_t) + R; \quad [eq. 4.30]$$

- Case 5 – empty phase ( $R = 0$  and  $C_0 = 0$ ):

$$IL = 0;$$

$$D = 0;$$

$$C_t = 0; \quad [eq. 4.31]$$

The input output schema of the implemented command *h.interception* is shown in figure 4.6.

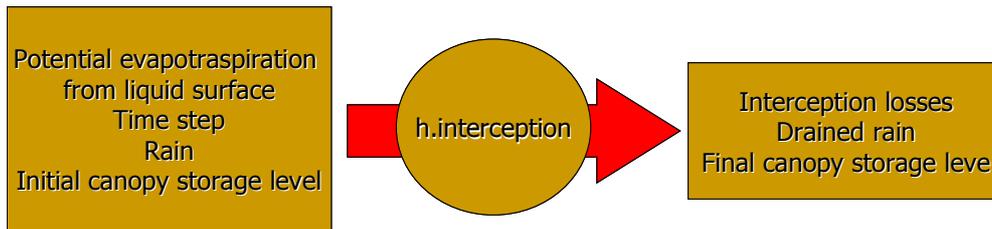


Figure 4.6 – input output schema of the *h.evapo.PM* command.

#### 4.6 Rainfall runoff maps generation

The governing equation for the runoff module has been presented in paragraph 3.3.4.

Solving those formulas implies the knowledge of the values of the following parameters at each cell:

- saturated hydraulic conductivity  $k_s$  [m/s] – soiltype dependent factor;
- saturated soil water content  $\theta_s$  [m<sup>3</sup>/m<sup>3</sup>] – soiltype dependent factor;
- residual soil water content  $\theta_r$  [m<sup>3</sup>/m<sup>3</sup>] – soiltype dependent factor;
- soil type dependent elevation parameter  $\alpha$  [-];
- soil thickness of the soil affected by subsurface flow  $L$  [m] – soiltype dependent factor;
- manning friction coefficient  $n$  [m<sup>-1/3</sup>s<sup>-1</sup>] – landuse dependent factor;
- terrain slope  $\beta$  [deg];
- terrain hydrological type  $h_t$  [-] – terrain or river;
- river width [m];

The terrain hydrological type allows to select which equations has to be solved, with this aim a map has to be derived. The river cells (where subsurface and eventually overland flow equation will be solved) will be identified by the value 1, while the terrain cells (where river, subsurface and eventually overland flow equation will be solved) will be identified by the value 0. Further development of

the runoff module, which may include lake or artificial basin modelling, could be achieved by using different hydrological type values.

Furthermore additional information are needed to solve in the correct order the flow that drains from one cell into another:

- flow direction – that shows in which of the 8 neighbour cell the water drains;
- flow accumulation – that shows the number of up-cells draining into.

Finally, with the aim to prevent any out of memory operations, each basin has to be solved separately but in a consequential order that permits to consider the outflow of a basin as one inflow for the following; thus the last information needed is:

- basin number – unique basin identifier.

Given a digital terrain model (DTM), the slopes, flow directions, flow accumulations, rivers and basins maps can be derived with the *r.solpe.aspect* and the *r.watershed* GRASS commands.

In order to correctly extract these maps with the *r.watershed* command a correct threshold value has to be chosen, in fact this value affects both the river and basin delineation. The right value can be chosen by visually comparing different *r.watershed* output river maps (obtained with different threshold values) with an existing vector map and selecting the one that gives the best fitting. Even if *r.watershed* do not ask for an elevation map without sinks, the result can be slightly different in the two cases (original DTM or filled DTM as input). The user can check if a filled DTM can improve the results by generating a filled elevation map with the GRASS *r.fill.dir* or *r.terraflow* command.

Note that for sake of simplicity the runoff are solved only in the four main directions, thus the *r.watershed* command has to be run with

the flag -4 that allows only horizontal and vertical flow of water. Future studies could implement other more elegant solution with multiple flow direction.

### 4.6.1 Initialization command: *h.hydroFOSS.init*

The flows (subsurface, surface and channel) have to be computed cell by cell following the flow accumulation order: this ensures that every incoming discharge has been already estimated, but also implies that raster cells have to be accessed in a sparse order.

These tasks are commonly performed by using temporary files or by loading the entire map into the memory. While the first method is time consuming because it needs multiple file position indicator settings (by means of the C ANSI fseek function), the second could imply limitation in solving large watersheds due to insufficient physical memory size.

In order to skip these two drawbacks the chosen approach makes extensive use of the database: this allows extracting the basins, and within each basin the cells with their own associated values, in the correct order of solution.

To benefit of the database properties a table with the previously mentioned informations has to be generated: the new developed command *h.hydroFOSS.init* has been designed for this purpose.

The input raster maps containing the cells properties and values for the cells of the region are: the filled DEM (without sinks), the slopes, the basins, the flow directions, the soiltypes, the landuses, the hydrotypes, the flow accumulations, and the riverwidths.

Furthermore the algorithm extracts informations from two tables, named *soiltype\_cat* and *landuse\_cat*, where *soiltype* and *landuse* dependent parameters values are stored.

The *soiltype\_cat* table contains six columns named *cat* (the *soiltype* map category values), *kappa\_s* (the saturated hydraulic conductivity values), *teta\_s* (saturated soil water content values), *teta\_r* (residual soil water content values), *alfa* (soil type dependent values), *thick*

(the thickness of the soil affected by subsurface flow values).

The `landuse_cat` table contains two columns, named `lu_id` (the landuse map category values) and `manning` (the Manning friction coefficient). The `h.hydroFOSS.init` uses the values found in the input raster maps and in the two tables to populate the initialization table that will be accessed for discharge maps computation.

The overall process that brings to the generation of the initialization table is represented in figure 4.7.

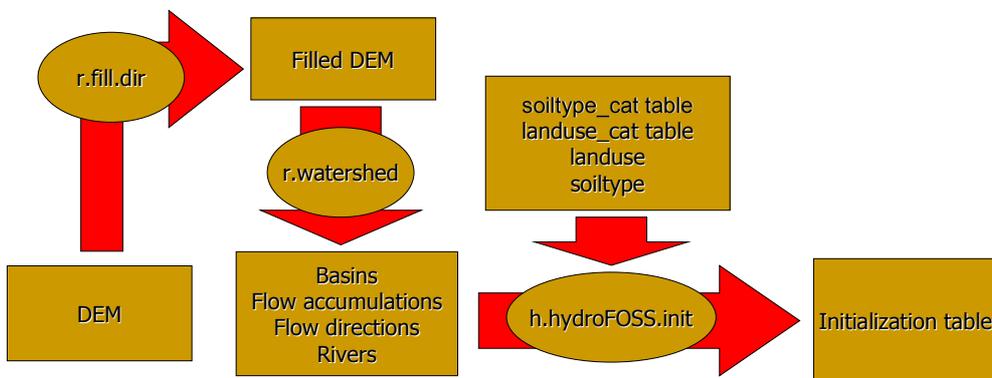


Figure 4.7 – rainfall runoff initialization procedures.

The output initialization table will store the informations of one cell for each row and will have the following attributes:

- `cat [int4]` – the cell index  $C_{idx}$ . Referring to figure 4.8, the cell index is a progressive integer number calculated with reference to the current GRASS region by the formula:

$$C_{idx} = (\text{ncolumns} * i) + j \quad [\text{eq. 4.32}]$$

where `ncolumns` are the total columns of the region, `i` is the GRASS row index and `j` is the GRASS column index;

- `ks [float8]` – the saturated hydraulic conductivity  $k_s$ ;
- `teta_s [float8]` – the saturated soil water content  $\theta_s$ ;
- `teta_r [float8]` – the residual soil water content  $\theta_r$ ;
- `alfa [float8]` – the soil type dependent  $\alpha$ ;
- `L [float8]` – the soil thickness  $L$ ;
- `n [float8]` – the manning coefficient  $n$ ;

## Chapter 4 - System implementation

- basin [float8] – the basin index  $B_{idx}$ : a unique index of the basin whose the cell belongs;
- direction [float8] – the flow direction index  $F_{idx}$ : provides the drainage direction values as defined in the *r.watershed* command.
- accumulation [float8] – the flow accumulation value  $A_{cc}$ : number of up-cells that drain into each cell as defined in the *r.watershed* command.
- slope [float8] – the terrain slope  $\beta$ : value, stated in degrees of inclination from the horizontal, as described in the *r.slope.aspect* command;
- h\_type [float8] – the hydrological type value: currently only two values are defined “0” for terrain and “1” for river;
- width [float8] – the riverwidth value;
- l\_use [int4] – the landuse category value  $lu_{cat}$ : the coded class which the cell belongs;
- s\_type [int4] – the soiltype category values  $st_{cat}$ : the coded class which the cell belongs;

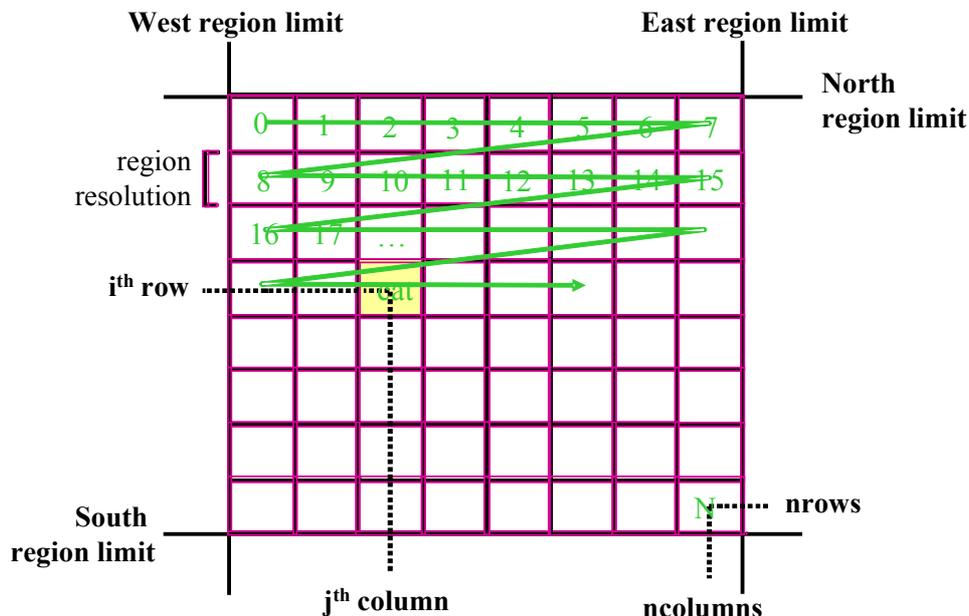


Figure 4.8 – cells indexing in cat attribute of the initialization table outputted from the *h.hydroFOSS.init* command.

4.6.2 Runoff computation command: *h.hydroFOSS.runoff*

The runoff computation algorithm has been developed into the new module *h.hydroFOSS.runoff*.

Before to proceed with the runoff estimation, the correct amount of water available for the routing (AW) and the actual evapotranspiration (ETa) rate has to be calculated.

The AW can be calculated as the water drained from canopy interception plus the rainfall directly falling over the ground plus the snowmelt.

The ETa can be estimated following a Topmodel approach. It applies a simple reduction function to the potential evapotranspiration (ETp) on the base of the soil water content as follows:

$$\begin{aligned}
 ETa &= (ETP \cdot V_s) / (\xi \cdot V_{s_{sat}}) && \text{if } V_s \leq \xi \cdot V_{s_{sat}} \\
 ETa &= ETP && \text{if } V_s > \xi \cdot V_{s_{sat}}
 \end{aligned}
 \tag{eq. 4.33}$$

where  $V_s$  is the soil water content,  $V_{s_{sat}}$  is the maximum soil water content and  $\xi$  is a threshold.

Following Menzel (1997) the value for the threshold  $\xi$  can be fixed equal to 0.6, that means to estimate evapotranspiration at potential rate for soil saturated with more then the 60% of their capacity and reduced otherwise.

Both, the AW and the ETa can be estimated by applying the map algebra as shown in frame 4.5.

## Chapter 4 - System implementation

```
L1> #runoff water estimation [mm/h]
L2> r.mapcalc 'NETrain.200503151300=Idrain.200503151300 +
snowmelt.200503151300 +
(1-veg_c_73)*rain.200503151300;
L3> #actual evapotranspiration estimation [mm/h]
L4> r.mapcalc 'EREAL.200503151300 =
if((soil_wc.200503151300
<(0.6*soil_wc_max)), (EPOT.200503151300*soil_wc.2005031513
00/0.6*soil_wc_max)), EPOT.200503151300 )';
```

Frame 4.5 – map algebra application for runoff water and actual evapotranspiration estimation.

Once calculated:

- the initialization table (generated with the *h.hydroFOSS.init* command);
- the AW and ETa map;
- the starting overland water content (Vo);
- the starting soil water content (Vs);
- the starting channel water content (Vc);

the new developed GRASS command *h.hydroFOSS.runoff* can be run.

The *h.hydroFOSS.runoff* algorithm flux can be described by the following steps:

1. selection of each basin index ordered by the maximum accumulation, the minimum and maximum row, the minimum and maximum column, and the minimum and maximum accumulation values.
2. allocation of a vector for storing the parameters of the outlet of each basin;
3. starting of a loop through the ordered basins;
4. allocation for the current basin of a matrix ranging from minimum column-1 to maximum column+1 and from minimum row-1 to maximum row+1. A frame of one cell around the exact matrix containing the basin guarantees the inclusion of every

## Chapter 4 - System implementation

- inlet (outlet of previously processed basin);
5. accordingly to the flow accumulation order, selection of all the cells, with their respective properties, that belong to the currently processed basin;
  6. population of the allocated matrix with the outlet falling into the matrix itself;
  7. starting of a loop through the selected cells in flow accumulation order;
  8. solution for the currently processed cell of the equations 3.10, 3.11, and 3.12 with a Runge-Kutta procedure;
  9. storing of the results for the currently processed cell in the allocated matrix;
  10. once solved the flow equations for all the cell of the currently processed basin, writing of the estimated values stored in the allocated matrix into temporary files and go on throw the basin loop by processing the following basin (go back to step 4).
  11. once solved all the basins, and transformed all the temporary files in permanent files the (i) soil water content [ $m^3$ ], (ii) overland water content [ $m^3$ ], (iii) channel water content maps [ $m^3$ ], and optionally (iv) soil water discharges [ $m^3/s$ ], (v) overland water discharges [ $m^3/s$ ], and (vi) channel water discharges [ $m^3/s$ ] maps, calculated at the end of the timestep, are outputted.

As shown in figure 4.9, the solution of the runoff equations for the cells where the river does not cover all the cell area, implies the solution of the overland and the subsurface equation over an area equal to the cell resolution minus the riverwidth, and the solution of the channel equation for an area equal to the riverwidth.

A further problem in these cells is represented by the Manning coefficient definition: in fact the initialization table assigns one value for each cell while two values are needed (one for the channel and one for the soil) in order to solve the flow equations.

In my implementation the archived Manning values for those cells

crossed by river are referred to the channel, and the values referred to the soil part of the cell is estimated as the mean values of the not river neighbours cells.

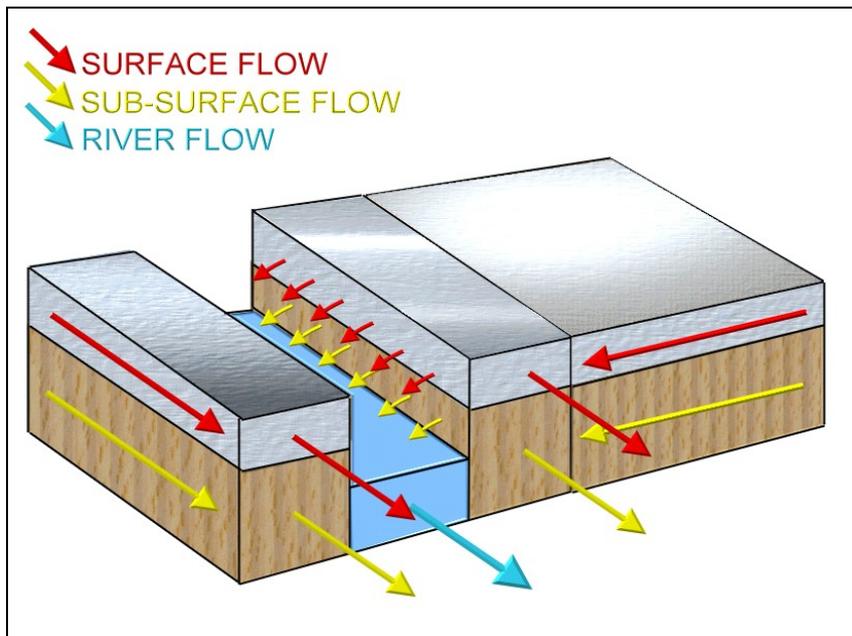


Figure 4.9 – *h.hydroFOSS.runoff* solved water flux over two cells.

The kinematic differential equations of flow have to be solved in the following order:

1. subsurface: in fact all the AW instantaneously goes into this layer until it is saturated; solving this equation it is possible to estimate the soil water discharges and the eventual ex-filtration to the overland layer.

Being:

$V_{Si}$  = the soil water content of the  $i^{\text{th}}$  cell;

$Q_{Si}$  = the soil water discharge of the  $i^{\text{th}}$  cell;

$p_i$  = the runoff water of the  $i^{\text{th}}$  cell;

$X$  = the grid resolution;

$Q_{oi}^U$  = the overland discharges of the cells drained by the  $i^{\text{th}}$  cell;

## Chapter 4 - System implementation

$Q_{si}^U$  = the subsurface discharges of the cells drained by the  $i^{\text{th}}$  cell;

$ETa_i$  = the actual evapotranspiration for the  $i^{\text{th}}$  cell;

$C_{si}$  = the subsurface local conductivity coefficient for the  $i^{\text{th}}$  cell as defined in equation 3.9;

$r_{oi}$  = soil water excess for the  $i^{\text{th}}$  cell;

the subsurface equation results:

$$\frac{\partial V_{Si}}{\partial t} = (p_i X^2 + Q_{O_i}^U + Q_{S_i}^U) - \frac{C_{S_i} \cdot X}{X^{2\alpha_s}} \cdot V_{S_i}^{\alpha_s} \quad [\text{eq. 4.34}]$$

Solving the equation 4.33 and subtracting the  $ETa_i$  the  $V_{si}(t+\Delta t)$ , the  $Q_{si}(t+\Delta t)$  and the  $r_{oi}(t+\Delta t)$  can be calculated.

2. Overland: only if the ex-filtration process is active the overland flow occurs and the overland equation can be solved in order to estimate the overland discharges;

Being:

$V_{oi}$  = the overland water content for the  $i^{\text{th}}$  cell;

$Q_{oi}$  = the overland water discharges for the  $i^{\text{th}}$  cell;

$C_{oi}$  = the coefficient relevant to the Manning formula for the  $i^{\text{th}}$  cell;

the overland equation results:

$$\frac{\partial V_{oi}}{\partial t} = (r_{oi} X^2) - \frac{C_{oi} \cdot X}{X^{2\alpha_o}} \cdot V_{oi}^{\alpha_o} \quad [\text{eq. 4.35}]$$

Solving equation 4.35, the  $V_{oi}(t+\Delta t)$  and the  $Q_{oi}(t+\Delta t)$  can be estimated.

3. Channel: if the cell contain the river, the two subsurface and overland flows are used as input for channel discharge estimation.

Being:

$V_{Ci}$  = the channel water content for the  $i^{\text{th}}$  cell;  
 $Q_{Ci}$  = the channel water discharges for the  $i^{\text{th}}$  cell;  
 $C_{Ci}$  = the coefficient relevant to the Manning formula for the  $i^{\text{th}}$  cell;  
 $W_i$  = the river width for the  $i^{\text{th}}$  cell;  
 $Q_{Ci}^U$  = the channel discharges of the cells drained by the  $i^{\text{th}}$  cell;  
 can be obtained the equation:

$$\frac{\partial V_{Ci}}{\partial t} = (r_{Ci} \cdot X \cdot W_i + Q_{Ci}^U + Q_{Si} + Q_{Oi}) - \frac{C_{Ci} \cdot W_i}{(X \cdot W_i)^{\alpha_c}} \cdot V_{Ci}^{\alpha_c} \quad [\text{eq. 4.36}]$$

Solving equation 4.36, the  $V_{Ci}(t+\Delta t)$  and the  $Q_{Ci}(t+\Delta t)$  can be estimated.

Figure 4.10 shows the input-output schema for the *h.hydroFOSS.runoff* command.

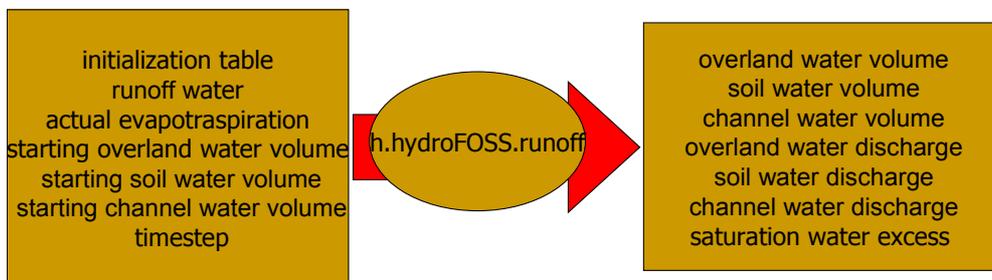


Figure 4.10 – input output schema of the *h.hydroFOSS.runoff* command.

It has to be noted that in order to select the optimal basin size to be processed at one time, the obtained basins map can eventually be aggregated in larger basins that join up-flow basins. This operation can be done by checking the maximum flow accumulation of neighbours basins and aggregating those with lower values.

### 4.7 Automatic model running

The procedures and the commands described in the previous paragraphs have to be run iteratively for each time step, that for the current model has been chosen of one hour.

This task is achieved by implementing the Shell script *h.hydroFOSS.model* that loops, from a starting date-time to an ending date-time, through the time and performs all the needed operations (catchment status maps generation, net radiation map generation, snow accumulation and melt maps generation, potential evapotranspiration map generation, canopy interception maps generation, and rainfall runoff maps generation).

The awk print function allows to parse the input dates (defining the starting and ending simulation period) and to set up the date suffix used to name the raster series maps.

```
>DATEstrFROM=`echo $GIS_OPT_from | awk -F"-" '{print($1" "$2"
"$3" "$4" "$5" 00");}'`
>DATEstrTO=`echo $GIS_OPT_to | awk -F"-" '{print($1" "$2" "$3"
"$4" "$5" 00");}'`
```

Then the while loop ending when through hours is achieved by using the shell function mktime, strftime as follow:

```
>while [ "$DATEstrFROM" != "$DATEstrTO" ]; then
>DATE_NOW=`gawk "BEGIN{print
strftime(\"%Y%m%d%H%M\",mktime(\"$DATEstrFROM\")) }"`
>DATE_LAST=`gawk "BEGIN{print
strftime(\"%Y%m%d%H%M\",mktime(\"$DATEstrFROM\")-3600) }"`
>.....
>DATEstrFROM=`gawk "BEGIN{print strftime(\"%Y %m %d %H
%M 00\",mktime(\"$DATEstrFROM\"))+3600) }"`
>fi
```

## Chapter 4 - System implementation

The number of generated maps can highly increase during the model running through time (i.e. 16376 a month), causing two problems:

- for Linux distributions using a kernel version younger than the 2.4, large numbers of files and folders in a single directory can lead to system errors due to the existence of a maximum node limit (*inode-max*: describes the maximum number of files or folders allowed for each directory);
- large numbers of nodes cause loss of performance in accessing files.

To skip these problems the directory structure schema shown in figure 4.11 has been adopted.

The active location has a “root mapset” containing all the basic maps like DTM, landuse, soiltype, riverwidth, etc. and a saved region that defines the model domain (root\_mapset of figure 4.11).

Then the model creates for each month a new mapset (identified by the root mapset name plus a month-year suffix: i.e. mapsetX.aug2005) that will store all the maps of the raster series belonging to that month.

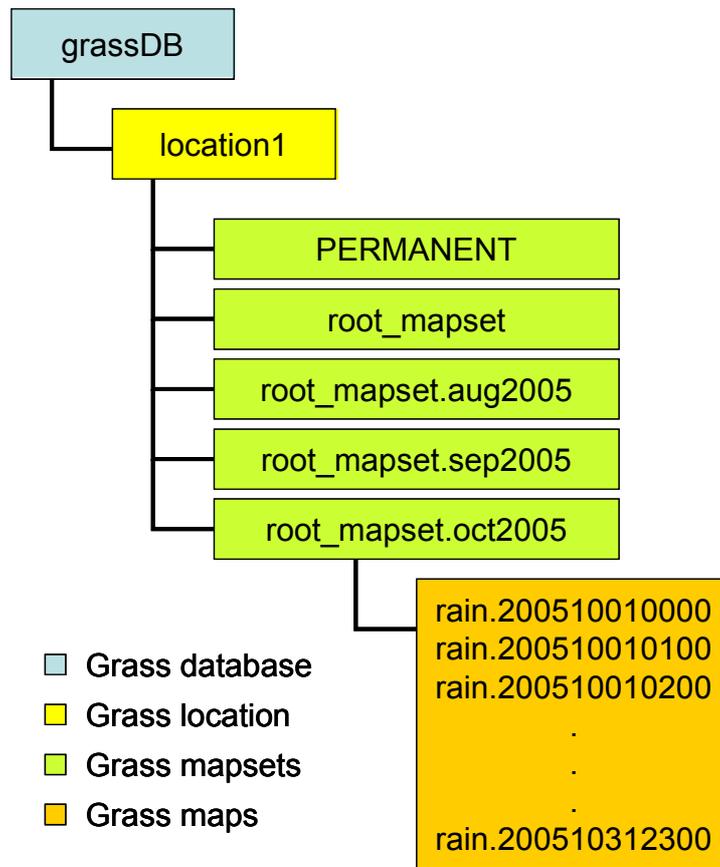


Figure 4.11 – hydrofoss directory structure.

For each modelled time step, *h.hydroFOSS.model* calculates the current date-time (DTc) and the previous step date-time (DTp), the current mapset (Mc) and the previous step mapset (Mp).

It is now possible to:

- access the last calculated maps by adding as a suffix the DTp to raster series name, and the Mp as the mapset where they reside (i.e. with raster series name = “rain”, Dtp=“200510171300”, and Mp=“oct2005”: map to read is “rain. 200510171300@ oct2005”);
- select the database informations by using the DTc (i.e. with Dtc=200510171400:  

```

“SELECT east, north, elev, value
FROM Measurement_points as p, Measurement as v
WHERE p.feature_id=v.feature_id

```

```
AND to_char(v.time,'YYYYMMDDHH24MI')='200510171400'  
AND v.param_id=3");
```

- generate new rasters maps and add to the raster series, in the Mc mapset, by using the DTc suffix  
(i.e. with raster series name = "rain", Dtc="200510171400", and Mc="oct2005": mapset to be set as active (*g.mapset*) is "oct2005" and map to be created is "rain. 200510171300");

The inputs required to run *h.hydroFOSS.model* are:

- general:  
root mapset name, and model region name;
- rasters:  
filled DEM map name, landuse map name, soiltype map name, aspect map name, and slope map name;
- database:  
climatic station database name, database driver name, table name for seasonal parameters, table name for monthly parameters, a string comma separated values of the four reference days, and initialization table name.

The generated time series maps are:

- rain; wind; humidity; temperature; directed, diffuse, reflected, and global solar radiation; clear sky index; net solar radiation; saturation vapour pressure; actual vapour pressure; longwave radiation; global radiation; snowpack water equivalent; snowpack energy content; snowmelt; potential evapotranspiration; potential evapotranspiration from a liquid surface; canopy storage level; interception loss; canopy drainage; water available for runoff; real evapotranspiration; overland discharge; subsurface discharge; channel discharge; overland water content; subsurface water content; and channel water content.

## **4.8 Sensitivity analysis and model calibration**

Even if HydroFOSS is a physical model, the values used to characterize the watershed are mean values calculated over the cells and are not directly measurable.

Therefore a sensitivity analysis and a calibration of the runoff component is needed.

### ***4.8.1 Advanced calibration method***

UCODE-2005 (Poeter et al., 2005) was used to compute sensitivities, parameter correlation coefficients, and to calibrate the hydrological model. The sensitivities are calculated by UCODE-2005 using parameter perturbation methods. UCODE-2005 is developed using the JUPITER API capabilities and conventions, which facilitates inter program communications.

In recent years, automatic calibration is becoming more diffused also for the calibration of hydrological model (Madsen, 2000). In manual calibration, a trial-and error parameter adjustment is made. In this case, the goodness-of-fit of the calibrated model is basically based on a visual judgement by comparing the simulated and the observed hydrographs. Accordingly to Madsen (2000), "For an experienced hydrologist it is possible to obtain a very good and hydrologically sound model using manual calibration. However, since there is no generally accepted objective measure of comparison, and because of the subjective judgement involved, it is difficult to assess explicitly the confidence of the model simulations. Furthermore, manual calibration may be a very time consuming task, especially for an inexperienced hydrologist."

Using the automatic calibration, parameters are adjusted automatically, according to specified rules and numerical measures of the goodness of the fit.

## Chapter 4 - System implementation

UCODE-2005 can be applied with any application model or set of models; the only requirement is that they have numerical (ASCII or text only) input and output files and that the numbers in these files have sufficient significant digits. The estimated parameters can be defined with user-specified functions: for example prior, or direct, information on estimated parameters also can be included in the regression. Observations to be matched in the regression can be any quantity for which a simulated equivalent value can be produced, thus simulated equivalent values are calculated using values that appear in the application model output files and can be manipulated with additive and multiplicative functions, if necessary. The nonlinear regression problem is solved by minimizing a weighted least-squares objective function with respect to the parameter values using a modified Gauss-Newton method. Sensitivities needed for the method are calculated approximately by forward or central differences and problems and solutions related to this approximation are discussed. Statistics are calculated and printed for use in (1) diagnosing inadequate data or identifying parameters that probably cannot be estimated with the available data, (2) evaluating estimated parameter values, (3) evaluating the model representation of the actual processes and (4) quantifying the uncertainty of model simulated values.

4.8.2 Calibration procedure implementation

The runoff model uses parameters stored in a Database Table.

For each cell, the table (see table 4.1) stores the values of the needed parameters.

<b>Column</b>	<b>Type</b>	<b>Description</b>
cat	integer	cell_ID
ks	double precision	conductivity
teta_s	double precision	saturated soil moisture content
teta_r	double precision	residual soil moisture content
alfa	double precision	soil parameter dependent
l	double precision	soil thickness
n	double precision	manning coefficient
basin	integer	basin_ID
direction	integer	flow direction
accumulation	integer	flow accumulation
slope	double precision	slope
h_type	integer	hydrological cell type (river/soil)
width	double precision	river width
landuse	integer	landuse map category
soiltype	integer	soiltype map category

Table 4.1 – initialization table fields

The model generates for each hour a new map of discharges.

The UCODE program is run with the configuration file (.in) as an argument.

The .in file specifies:

- the command to run the model;
- the parameters to be evaluated and their information (starting value, lower reasonable value, largest reasonable value, scale parameter values, adjustability (yes/no), fractional amount of perturbation for sensitivity, applied log-transformation (yes/no), maximum fractional change between iterations, etc.);

- the observation data file (.flo);
- the model input file (.if);
- the model input file template (.tpl);
- the model output file (.est);
- the model output file structure (.ins).

In order to link the model with the calibration program the following task are required:

1. create a command that samples the channel discharge raster in the given points and creates the estimation file (.est).

The command *h.sample.pt* has been generated by modifying the already existing GRASS command *r.sample*: the new command given a raster and a point vector map, prints to standard output, for each point: the point cat value, the raster name, and the value of the cell where the point falls into.

i.e. "h.sample.pt vector=st raster=channel\_wd.200408022200" prints out:

```
(1)channel_wd.200408022200 | 1.306722
```

```
(2)channel_wd.200408022200 | 5.317407
```

2. Create the output file structure files (.ins), to allow UCODE to extract estimations. This is done following the UCODE requirements for each observation by setting the line number, the observation name, and the character of the line containing the value (starting character: ending character)

i.e.

```
jif @
```

```
1 [(1)channel_wd.200408022000]18:22
```

```
2 [(1)channel_wd.200408022100]18:22
```

```
3 [(1)channel_wd.200408102200]18:22
```

```
.....
```

3. create a template file for update the parameters.

This is the file that the model runs to change the parameter values at each new call; in practice being the parameters in a

## Chapter 4 - System implementation

database table, this is an SQL file with the parameters name instead of the values.

i.e.

```
UPDATE IniTab SET ks = @ ks_soiltype_1 @ WHERE soiltype=1;
UPDATE IniTab SET ks = @ ks_soiltype_2 @ WHERE soiltype=2;
UPDATE IniTab SET l = @ l_soiltype_4 @ WHERE soiltype=4;
UPDATE IniTab SET n = @ n_landuse_11 @ WHERE landuse=11;
```

.....

4. create the observation files to be used to compare the model estimations and perform the calibration (.flo). This is the files where the river gauges observations are written.

i.e.

```
BEGIN OBSERVATION_DATA TABLE
NROW=1440 NCOL=6 COLUMNLABELS
(1)channel_wd.200408102100 0.78537
(1)channel_wd.200408102200 0.81421
(1)channel_wd.200408102300 1.13655
```

....

5. create a command that updates the parameters and run the model. This is a simple shell script, calling the db.execute GRASS command, to executes the model input file (.if) generated by UCODE and then the h.hydroFOSS.model.

i.e.

```
> db.execute input=filename.if;
> h.hydroFOSS.model date_from=200408011200
date_to=200408311200 db=verzasca drv=pg tab=ucode_test
dem=dem_fill slope=SLOPE soiltype=soiltype landuse=landuse
mapset=verz_cal region=verzasca sample=verz_lim@verz_cal
seas_tab=season_par seas_days="120,180,230,290"
mon_tab=monthly_par;
```

It is now possible, in a Shell windows, to run within GRASS the calibration process by calling the command:

## Chapter 4 - System implementation

> ucode config.in calibration1;

The described procedure realizes the linkage between GRASS and UCODE-2005, its application will be shown in the chapter 6, where a first calibration has been run.

# **5. Investigation of the radar rainfall estimates and of the temperature interpolation method**

As previously discussed in the section 4.1.1 in this chapter a detailed investigation on the estimations of rainfall intensity derived from meteorological radar observations (paragraph 5.1) and the temperature interpolation technique (paragraph 5.2) are presented.

## ***5.1 Radar rainfall estimates investigation***

Remote sensed data are attractive because they can be sampled in a distributed way but often, unlike the gauges observations, are not direct measurements of the phenomena.

In order to validate the derived information, a comprehensive analysis of rainfall estimations, by means of weather data in Southern Switzerland and Northern Italy, is carried out. Validation of the MeteoSwiss weather radar data vs. rainfall gauges data is conducted by using a data set composed of 365 days (Year 2002) per 24 hourly observations at 35 different rain gauges: these have been homogenized by classifying gauge measurements on the basis of the radar estimates classes and further subdivided in two different seasonal datasets.

### ***5.1.1 State of the art in rainfall estimate***

For operational real-time flood forecasting systems ground-based rain-gauges are still commonly used for several reasons: (i) they provide a direct measurement of rain, although in a single point; (ii) the measurement and the data transfer techniques are nowadays well established; (iii) radar and satellite systems still require high investment and maintenance cost. In the last 20 years increasing attention has been focused on weather radar system for spatial

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

measurement of precipitation. In the 1980s the EU sponsored the COST72 (1985) and COST73 programs, with the aim of establishing a European radar network.

The COST Action 717 involves radar experts, hydrologists, and meteorologists from several countries dealing with “Use of Radar Observations in Hydrological and Numerical Weather Prediction (NWP) models”. A common basis of all activities within the project is the availability of reliable radar-derived precipitation estimates, and, therefore, the study about the conversion of radar echo to precipitation intensity. According to the experience collected during the project, as summarized in Gjersten et al. (2003), no generally-accepted methodology for radar-derived precipitation estimates is available and many different approaches for improving the quantitative precipitation estimates of weather radar are in use or presently tested in different EU-countries.

The current and most applied methods of radar (R) adjustment usually rely on rain gauge data (G), although different hypothesis and aggregation procedures have been tested within this general framework, as illustrated in Krajewski (1987), Borga et al. (2002), Todini (2001) Seo (1998), and several others. Based on the definition of Michelson (2003), we concentrated on the gauge-radar adjustment problem, which involves any procedure whereby the characteristics of radar data are modified such that they correspond as much as possible with the measured quantity collected by rain gauges. According to these recent studies, a comprehensive statistical analysis and comparison of weather radar estimates and rain gauges data were carried out.

### *5.1.2 Study area*

The study area is situated upstream to the Lake Maggiore, on the border between Switzerland and Italy. The area (6599 km<sup>2</sup>) is mountainous, with elevation ranging from 200 m to about 3500 m a.s.l., particularly prone to heavy rainfall and flood events.

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

### *5.1.3 Radar data*

The study area is covered by the Mount Lema weather radar operated by MeteoSwiss (see Figure 5.1). The C-band Doppler weather radar was conceived to work continuously in an unattended operational environment and to supply high quality data even in presence of severe ground clutter. With a spatial resolution of the volume element of 1 km<sup>3</sup>, 20 different elevations are scanned every 5 minutes and finally converted in 16 reflectivity levels. A best estimate of the precipitation rate, called RAIN, containing the first 30 minutes precipitation intensity was used as data set for all the analyses. Variability in time and space of the vertical profile of reflectivity (VPR), issues related to beam geometry, ground clutter, anomalous propagation, and visibility, stability of the electronic calibration of the radar system are all already considered during the processing chain up to the RAIN product. The values of reflectivity are finally transformed by MeteoSwiss in rainfall intensity by using the standard Z-R relationship:

$$Z = A \cdot R^b \quad [\text{eq. 5.1}]$$

where Z is the reflectivity [mm<sup>6</sup>/m<sup>3</sup>], R is the corresponding rainfall intensity [mm/h]; A=316 and b= 1.5 are the parameters estimated from Joss et al. (1997) for whole Switzerland.

### *5.1.4 Rainfall Data*

Hourly rainfall measurements from 35 stations and 3 different meteorological networks (see figure 5.1) are collected in real-time: the network of the Institute of Earth Sciences (IST), the network of Piemonte Region (RP) in Italy, and the network of MeteoSwiss (MS). A data quality control is directly performed by MeteoSwiss and a plausibility check is carried out, before the data are transferred into the database. Rainfall data and radar estimates were available for the whole Julian year 2002.

Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

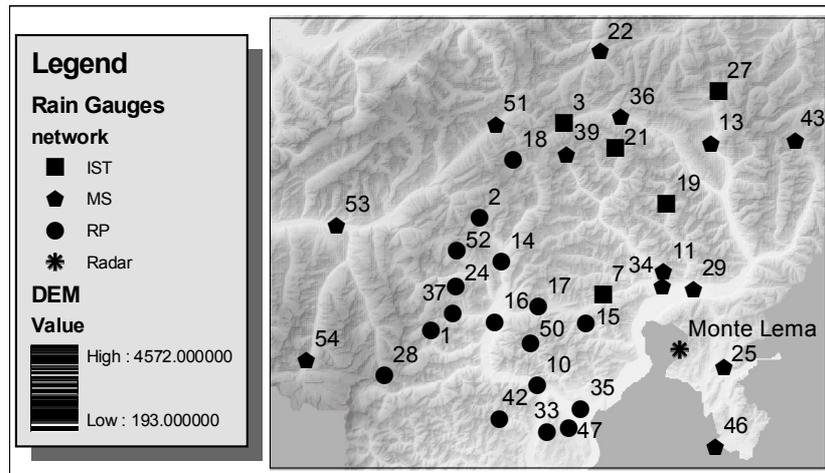


Figure 5.1 – Radar and rain gauge networks

In table 5.1 some basic statistics of the two data sets are illustrated.

	Radar	Rain Gauges
Observations No.	251068	251068
Mean	0.1	0.26
Variance	0.33	1.4
Min	0	0
Max	50	58.2

Table 5.1 – Basic statistics of the data sets in mm/h.

*5.1.5 Preliminary analysis*

A preliminary analysis of the data samples has been carried out, in order to (i) verify absence of anomalies, (ii) detect presence of outliers in the radar images, and (iii) homogenize the different rainfall data sets.

Persistent anomalies have been excluded by verifying that none of the rain gauges falls into a cell where the reflectivity in the RAIN product exceeds 0.16 mm/h in at least 20% of the time in clear sky conditions. Gross errors are identifiable by considering a 3x3 pixel grid centered on the i-th pluviometer and by calculating, for each time step, the mean value  $\mu$  and the median  $m$  of the 9 observations. By comparing the two values, the difference  $d=\mu-m$  confirms the

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

absence of many outliers in the radar estimates corresponding to the rain gauge positions (average=0.00, median 0.01, min.=-7.67, max.=9.00 in mm/h).

The homogenization of the rainfall data were performed throughout (i) a temporal synchronization of the rainfall data collected from different networks, and (ii) a reclassification procedure. In order to compare radar estimates with rain-gauge measurements, the rainfall data were classified according to the boundaries of the radar classes (3dBZ amplitude), as provided by MeteoSwiss.

To each class, for both radar and rain-gauge, the central value of the class was finally associated. Any further analysis will refer to this data set, whereas error sources not considered in RAIN were at the moment neglected.

Finally, I assumed that gauge measurements are appropriate for each gauge location and they are therefore referred to as the "ground truth". This is only partially true, since precipitation measurements are affected by different errors.

Nevertheless Lombardo and Staggi (1998) showed that the systematic error is finally a bias, which could be corrected using a simple calibration curve. Additionally to the gauge measurement errors, the question of using point measurements in representing areal rainfall of a radar pixel has been emphasized by many authors (see, e.g., Chumchean et al., 2003).

### *5.1.6 Rainfall detection analysis*

As first test of the radar reliability, the simple detection of a rainfall event, without any consideration of its magnitude was carried out, by considering the binomial process rain no-rain, both for gauges and radar. The relative concordance and discordance frequency of the four possible cases are summarized in table 5.2.

A high agreement level (over 90%) is achieved in the case of both rain and no-rain cases. Furthermore, referring the gauges as ground-truth, the detection of a no-rain event by the radar is very reliable, not only in the frequency (97%), but also in terms of

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

intensity, since in the empirical cumulative distribution function of no-concordance cases only 0.66% correspond to error higher than 2 mm/h. The detection of a rainfall event is on the other side more uncertain, with an error by the radar in 47% of the cases.

		RAIN GAUGE		
		rain	no-rain	tot
RADAR	rain	16218	6482	22700
	no-rain	18050	210318	228368
	tot	34268	216800	251068

Table 5.2 – Four cell contingency table, gauges and radar

### 5.1.7 Rainfall intensity analysis

By visually comparing rain-gauge measurements (named RG) and radar estimates (named RR), a tendency of the radar to underestimate was evident. A seasonal analysis on the conditional sample with  $RR > 0$  and  $RG > 0$  (herein referred to as “active-set”), aiming at distinguish different characteristics of the rainfall events (advective events in the cold season, convective events in the hot season) confirms this tendency. The standard deviation of RR is smaller than the one observed by RG and seasonally variable. Moreover the values of  $D = RG - RR$  on the “active-set” are bias affected since their mean and median are strongly different (average=1.69, median=0.96, min.=-48.00, max.=56.00 in mm/h).

### 5.1.8 Spatial correlation analysis

A spatial correlation analysis of rainfall measurements was undertaken, aiming at identifying spatial properties of the rainfall field. This would allow radar adjustment with gauge data, e.g. by Kriging or Cokriging procedures as in Krajewski (1987) or Todini (2001).

The empirical semivariogram on the “active-set” was calculated as follow:

$$\gamma(h_k) = \frac{1}{2N(h_k)} \sum_t \sum_{s_i, s_j=1}^{35} \{y_t(s_i) - y_t(s_j)\}^2 \text{ with: } 5k \leq |s_i - s_j| \leq 5(k+1) \quad [\text{eq. 5.2}]$$

where  $h_k = [5k, 5(k+1)]$  km,  $\gamma(h_k)$  is the mean semivariance in the range  $h_k$ ,  $N(h_k)$  is the corresponding number of observed pairs in the range  $h_k$ ,  $y_t(s_i)$  is the rainfall observation at gauge  $s_i$  at time  $t$ . Both radar and gauges display a phenomenon with similar correlation length (sill), but their  $\gamma(h_k)$  values, as illustrated in figure 5.2, present a non constant ratio among the empirical variograms calculated at different months.

The different spatial correlation of the rainfall field, as observed by gauges and radar, prevent us from directly and trivially integrating the two datasets into a unique geostatistical model.

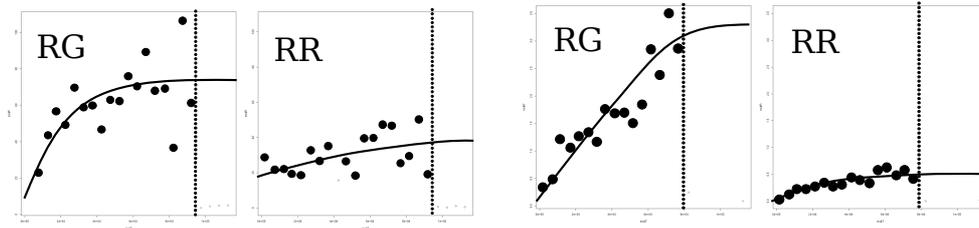


Figure 5.2 – Monthly semivariograms for RG and RR  
(June on the left and December on the right).

### 5.1.9 Linear and functional correlation analysis

By using linear correlation analysis a functional relationship between RR and RG was investigated. The global correlation index is extremely low ( $\rho = 0.324$ ), and the distribution of the single  $\rho$  values at each gauge does not reveal any spatial pattern, therefore indicating that a linear radar vs. rain gauges adjustment procedure can not be implemented. A more general functional dependence was thus investigated.

A general dependence of two variables  $Y$  and  $X$  can be estimated by means of a functional index  $\eta^2_Y$ , defined according to Sansò (1996):

$$\eta_Y^2(x, y) = \frac{\sigma_{S_Y}^2}{\sigma_Y^2} \quad [\text{eq. 5.3}]$$

where:  $\sigma_{S_Y}^2 = E\{(E\{Y|X=x\} - \mu_Y)^2\}$ ,  $\sigma_Y^2 = E\{(Y - \mu_Y)^2\}$ .

The index ranges from 0 to 1;  $\eta_Y^2 = 1$  if and only if Y is a function of X, and  $\eta_Y^2 = 0$  if X and Y are stochastically independent. By comparing this value with  $\rho$  a certain, but not satisfactory, functional dependency can be identified. Taking into account the characteristics of the data, particularly: the radar data availability (rainfall estimates of the first 30 minutes of each hour), and the physical behaviour of the rainfall phenomenon in narrow and deep alpine valleys, a 3 hours moving average operator was applied to RR.

The correlation between radar and rain gauges significantly increased suggesting to look for a new parameters estimate of the Z-R relation using  $Z_{Rm}$  (3 hours moving average reflectivity) and RG. The results are summarized in figure 5.3.

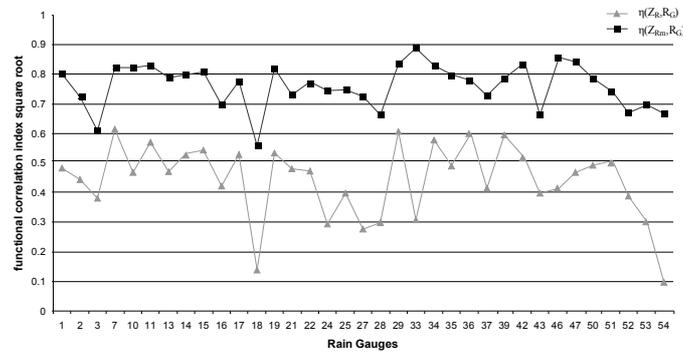


Figure 5.3 – Square root of the functional correlation index for different gauges and radar.

### 5.1.10 Z-R relationship adjustment

According to the previous analysis,  $\alpha$  and  $\beta$  of the inverse of the radar equation

$$R_G = \alpha \cdot Z_{Rm}^\beta \quad [\text{eq. 5.4}]$$

were fitted for each station, by non-linear least-square estimation and further converted into the commonly used A and b (see figure 5.4 and equation 5.1). The result of the fitting shows that A values mostly range from 0 to 50 and b values range from 2 to 4 and therefore they greatly differ from those operationally used by MeteoSwiss. A high spatial variability can be also observed.

By using an outliers identification procedure (Grubbs' Test) 3 stations (No. 18, 36, 54) were excluded from further analysis. The result of the Grubbs' Test fairly agrees with consideration related to the local meteorological and topographical conditions of the identified critical sites. By omitting these stations, a general Z-R relation for the whole data set was estimated; nevertheless this approach is still unsatisfactory, since the residual variance curve is higher than the one computed at each gauge (see as example figures 5.5 and 5.6).

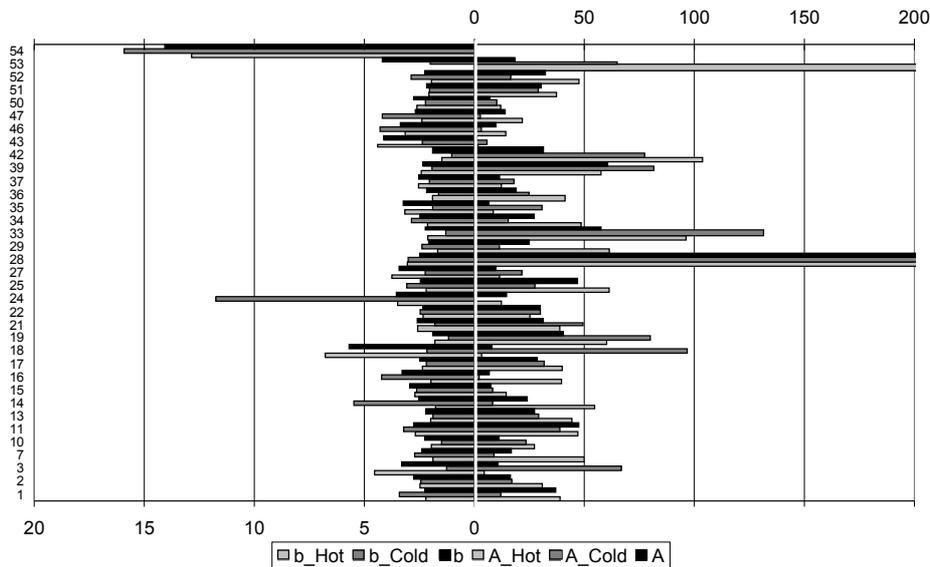


Figure 5.4 – Estimated Z-R parameters in different seasons (Hot and Cold) and all over 2002.

Chapter 5 - Investigation of the radar rainfall estimates  
and of the temperature interpolation method

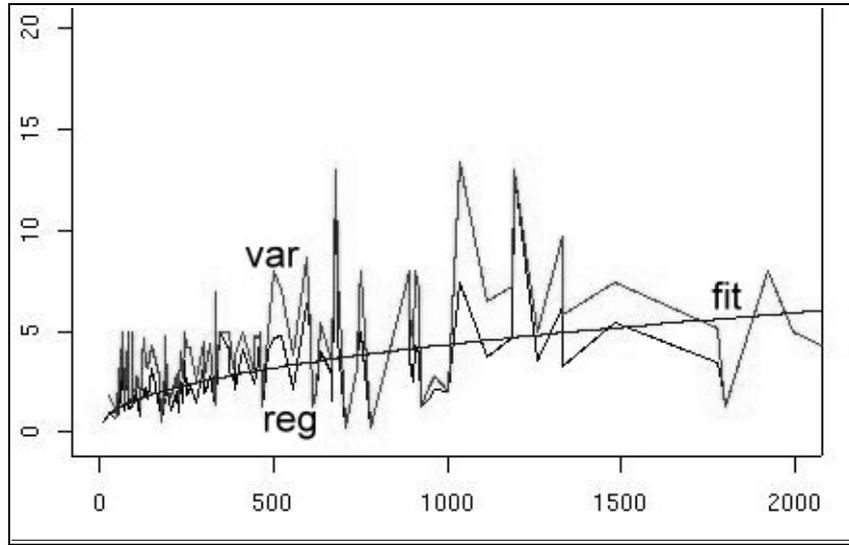


Figure 5.5 – Regression curve (reg), residual variance curve (var), and fitted model (fit) on gauge No. 1 data

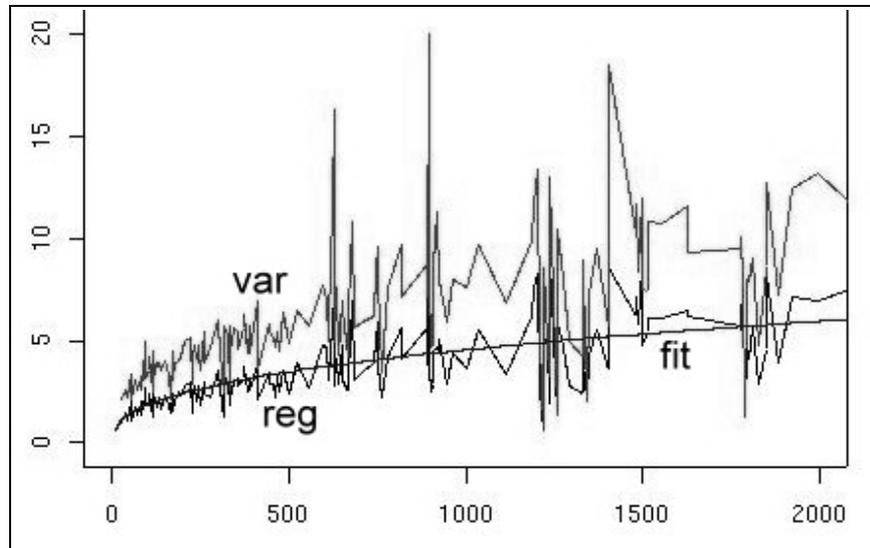


Figure 5.6 – Regression curve (reg), residual variance curve (var), and fitted model (fit) on the whole data set (right).

*5.1.11 Conclusions of the validation of radar rainfall estimates*

It has been shown that, in the study area, the MeteoSwiss RAIN product:

- can be used as reliable tool for no-rain detection;
- underestimates rain average and variance values in comparison with gauge observations;
- shows different spatial correlations compared to those usually observed from gauge data.

To achieve a reliable rainfall intensity map, a recalibration of radar measures is needed.

The use of a moving average operator on radar reflectivity observations generates a high non linear functional correlation between the consequent radar data and the rain gauges measurements. This functional dependence could be modelled by means of locally recalibrated Z-R relations.

The strong space-wise and time-wise variability for the computed parameters of the Z-R relation suggests that a local fit and a further interpolation, instead of using constant  $A$  and  $b$  values for the whole area, could be a methodology to achieve a reliable rainfall intensity map. Anyway the studied product (RAIN) has been shown to be not yet mature for quantitative precipitation estimations and therefore a classical approach has been chosen.

One question still remains open: is it higher the radar error in magnitude estimates or the interpolation error in spatialize point-wise observation? No answer can be found at the moment because of the lack of any reliable measures.

It could be argument for further research to compare hydrological models simulations (estimated river flow discharges) obtained with gauges inputs and with radar estimates inputs to the observed discharges, in order to verify the reliability of radar rainfall estimates in hydrological modelling.

## **5.2 Investigation of interpolation methods for temperatures maps derivation**

Surface air temperature is an important meteorological input in most formulas for the evapotranspiration and snowmelt calculation. Therefore, an accurate interpolation of temperature is necessary.

Due to the lack of numerous observations, the topography dependence of this variable has been often taken into account in order to improve the spazialization of the sampled points.

Different interpolation schemes were evaluated in literature using the leave-one-out cross-validation method (see e.g. Collins and Bolstad, 1996).

The interpolation schemes often applied are:

1. constant lapse rate;
2. inverse distance weighting (IDW) without consideration of elevation effects;
3. ordinary Kriging without consideration of elevation effects;
4. hybrid scheme that combined a constant lapse rate with ordinary Kriging;
5. elevationally detrended kriging approach, which is a combination of regression against elevation and ordinary Kriging (Li et al., 2003).

After an investigation of the different techniques and the influence on the temperature of various parameters, the elevation (Z) was found as the most significant one. As balance between accuracy and simplicity three models was tested:

1. MODz – a simple linear relationship between the temperature (T) and the height (Z):

$$T = aZ + b \quad [\text{eq. 5.5}]$$

2. MODz2 – a 2<sup>nd</sup> order polynomial relationship between temperature and height trying to take into account the

Chapter 5 - Investigation of the radar rainfall estimates  
and of the temperature interpolation method

phenomenon of the inversion temperature profile:

$$T = aZ^2 + bZ + c \quad [\text{eq. 5.6}]$$

3. MODxyz – a 3<sup>rd</sup> order polynomial in X and Y with an additional term in Z to consider temperature attenuation with elevation:

$$T = b + b_1X + b_2Y + b_3X^2 + b_4Y^2 + b_5XY + b_6X^3 + b_7Y^3 + b_8X^2Y + b_9XY^2 + b_{10}Z \quad [\text{eq. 5.7}]$$

This three methods were validated by using the leave-one-out cross-validation procedure: one observation at time is leaved out from the interpolation; the omitted observation is then used to compute the residual between its value and the one generated, by means of interpolation, without its contribution.

For each map (for each hour) it is possible to evaluate the MBE (mean bias error) and the MAE (mean absolute error) following the equation 5.8 and 5.9:

$$MBE = \frac{\sum_{i=1}^n Es_i - Eo_i}{n} \quad [\text{eq. 5.8}]$$

$$MAE = \frac{\sum_{i=1}^n |Es_i - Eo_i|}{n} \quad [\text{eq. 5.9}]$$

where:

$n$  = number of observation,  $Es$  = estimated value,  $Eo$  = omitted observation,  $i$  = observation index.

Applying this procedure for multiple time series it is possible to evaluate a series of these indexes that indicates the goodness of the interpolations.

The data set for the cross-validation procedure was composed by a total of 251'959 observations coming from 32 stations (see table

Chapter 5 - Investigation of the radar rainfall estimates  
and of the temperature interpolation method

5.3) with hourly sampling lap and referring to the whole year 2002. This approach has been used to evaluate the performances of the MODz, MODz2 and MODxyz model for temperature interpolation.

Nid	Station Name	Descr	Est	Nord	Quota
1	Alpe di Cheggio	ant	652171	104165	1510
2	Alpe Devero	bac	663117	129561	1634
10	Candoglia	can	676102	91818	201
11	Cimetta	cim	704370	117515	1672
13	Comprovasco	com	714998	146440	575
14	Crodo	cro	668170	119610	560
15	Cursolo Orasso	cur	687181	105678	940
16	Domodossola	dom	666627	105974	252
17	Druogno	dru	676289	109454	831
18	Piano dei Camosci	for	670747	142739	2450
22	Guetsch	gue	690140	167590	2287
24	Lago Paione Sup.	lpa	657981	114162	2269
25	Lugano	lug	717880	95870	273
28	Passo del Moro	mac	641739	93894	2820
29	Locarno-Magadino	mag	711160	113540	197
30	Monte Matro	mat	714260	140940	2171
33	Mottarone	mot	678326	81222	1302
34	Locarno-Monti	olm	704160	114350	366
35	Pallanza	pal	685952	86297	211
36	Piotta	pio	694930	152500	1007
37	Pizzanico	piz	657116	108101	1142
39	Robiei	roe	682600	144075	1898
42	Sambughetto	smb	667557	84089	800
43	S. Bernardino	sbe	734120	147270	1639
46	Stabio	sta	716040	77970	353
47	Someraro	str	683252	82060	430
50	Alpe Mottac	tro	674650	101342	1690
51	Ulrichen	ulr	666740	150760	1345
52	S. Domenico	var	658120	122068	1300
53	Visp	vis	631150	128020	640
54	Zermatt	zer	624350	97550	1638
55	Lodrino	ldr	719650	128500	261
56	Generoso	gen	722250	128500	1608
57	Cevio	cev	689700	130550	416

Table 5.3 – stations used in temperature interpolation methods analysis

Based on this data set the MBA and MBE where calculated at each hour: in table 5.4 a statistic summary of these indexes is presented. The model MODxyz results in being too inaccurate with a mean

Chapter 5 - Investigation of the radar rainfall estimates  
and of the temperature interpolation method

MAE of 4.3°C while MODz and MODz2 results to be fairly good with a mean MEA values of 0.42°C.

		MODxyz	MODz	MODz2
MAE	max	195.589	5.8153	5.8321
	min	0.418	0.4275	0.4405
	mean	4.368	1.5298	1.5219
	median	1.778	1.3532	1.3451
MBE	max	56.09081	0.243848	0.70255
	min	-56.38430	-0.284020	-0.18990
	mean	-0.09006	0.001829	0.01669
	median	0.09309	0.001834	0.01203

Table 5.4 – MAE and MBE statistics for the 2002 temperature (°C) time series

Figure 5.7 shows the MAE behaviour along the whole year 2002 with the three evaluated methods.

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

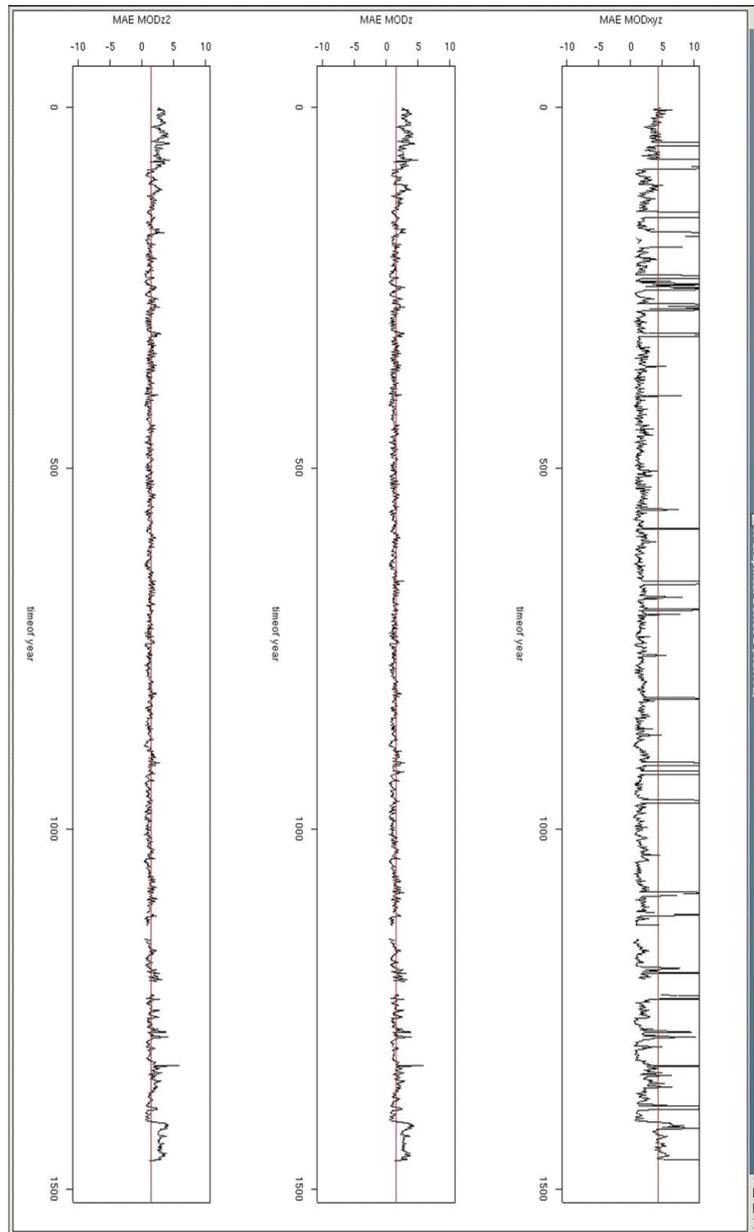


Figure 5.7 – temporal distribution of MAE with three interpolation methods; the horizontal lines represent the mean value.

Also figure 5.8 shows at a given time (2002-01-14 00:00) that MODz has better fitting than MODxyz.

Chapter 5 - Investigation of the radar rainfall estimates  
and of the temperature interpolation method

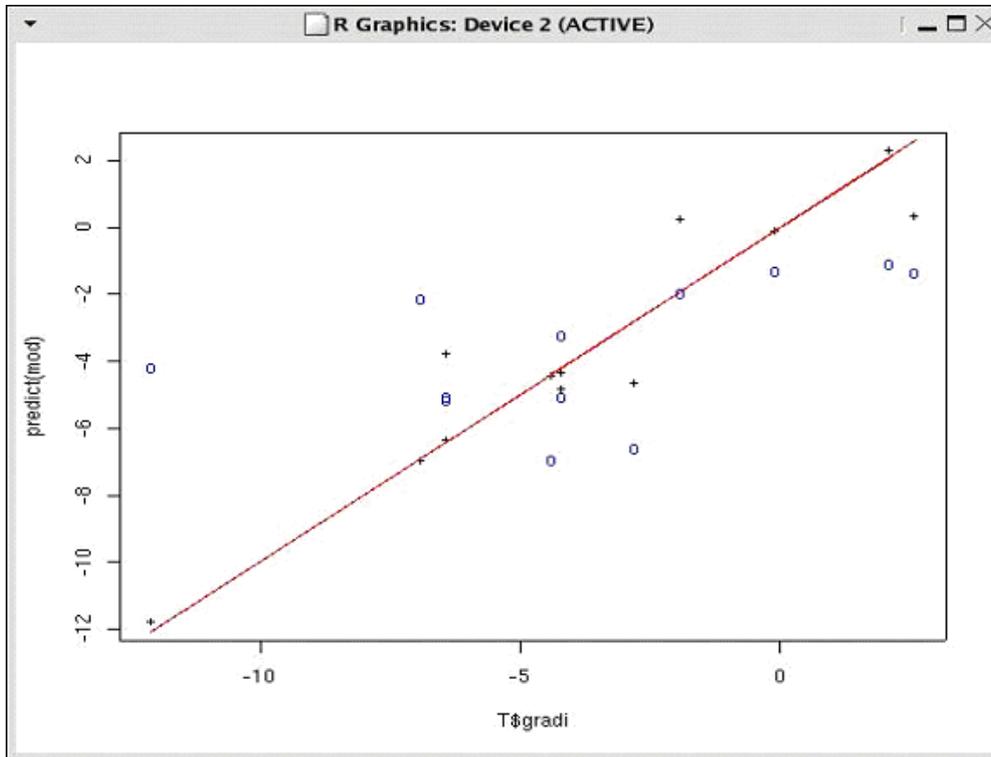


Figure 5.8 – observations vs. estimations: with the “o” symbol the MODxyz prediction and with the “+” symbol the MODz prediction.

A further confirm of that is given by the visualization of these maps, as shown in figure 5.9.

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

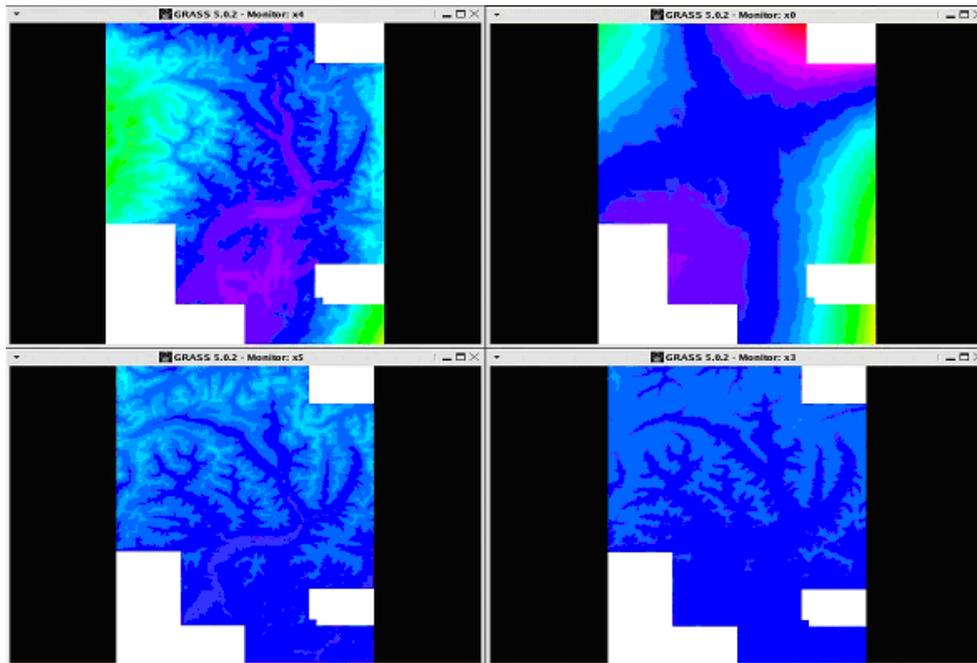


Figure 5.9 – maps generated with different interpolation methods. In clock wise order from the top left map: weighed inverse distance, MODxyz, MODz and MODz2.

The cross-validation shows how the two interpolation methods MODz2 and MODz are very close in generating interpolation errors; therefore the simplest one has been selected.

The implemented procedure is composed of three steps:

- querying of the MNSD database and extraction of the temperatures and relative elevation of observation;
- regression analysis and calculation of the regression parameters;
- application of the estimated relationship to the elevation map by means of map algebra;

As shown in frame 5.1 a pearl script executes the query (based on a system variable defining the correct timestamp) and the regression analysis returning the parameters A and B used for mapalgebra by the *r.mapcalc* GRASS command.

## Chapter 5 - Investigation of the radar rainfall estimates and of the temperature interpolation method

```
L1> # query and calculation of reg. parameters
L2> COEFF=`perl temp_regression.pl`
L3> A=`echo $COEFF | awk -F" " '{print($1);}'`
L4> B=`echo $COEFF | awk -F" " '{print($2);}'`
L5> r.mapcalc temp.200503151300 =$A+$B*dem
```

Frame 5.1 - temperature map derivation procedure.

## 6. Case study: the Verzasca basin

In order to test the developed procedures and modules, they have been applied to the Verzasca basin as a case study.

The analysed period is the August 2004, where some short showers and one heavy rainfall event occurs.

After a general introduction to the geographic region of the Verzasca basin, the most significant produced maps have been analysed and validated by means of: verified reliable spatial representation, and comparisons with other available data and “in situ” observations.

### 6.1 General characteristics

The Verzasca valley, is located in southern Switzerland, Canton Ticino (see figure 6.1). The Verzasca river flows from the Alps to the Lake Maggiore and has a length of 34 Km with a mean discharges of  $1 \text{ m}^3/\text{s}$  that, during floods events, can easily rise up to  $250 \text{ m}^3/\text{s}$ . The drained area has an extension of  $218 \text{ Km}^2$  and registers a mean annual precipitation of 1976 mm with a mean of 129 rainy days a year. The basin shows the typical Alpine characteristics with its high elevation ranges (from 193 m to 3596 m a.s.l.), narrow valley (2 Km wide), and high mean slope ( $33.93^\circ$ ).



Figure 6.1 – Verzasca valley localization in Switzerland.

## Chapter 6 – Case study: the Verzasca basin

Figure 6.2 and 6.3 show a picture and a satellite image of the valley.



Figure 6.2 – picture of the Verzasca valley at Ponte Brolla



Figure 6.3 - Satellite image of the Verzasca valley  
(source <http://maps.google.com>)

## 6.2 Projection and region properties

The projection used in the case study, following the proj structure([www.remotesensing.org/proj/gen\\_parms.html](http://www.remotesensing.org/proj/gen_parms.html), April 2006) is:

- name : Swiss. Obl. Mercator
- datum : ch1903
- towgs84 : 674.253,15.053,405.324
- proj : somerc
- ellps : bessel
- a : 6377397.1550000003
- es : 0.0066743722
- f : 299.1528128000
- lat\_0 : 46.9524055560
- lon\_0 : 7.4395833333
- k\_0 : 1.0000000000
- x\_0 : 600000
- y\_0 : 200000
- unit : meters
- units : meters
- meters : 1

The region of application, that will be maintained during all the processing phases and saved with the name of Verzasca\_reg (*g.region save=Verzasca\_reg;*) is identified by the following coordinate of the limits:

north=140636.61458333; south=115129.97395833;  
west=695548.32142829; east=715255.35714256.

The resolution in north-south and east-west has been fixed at 100 m and the rows number and columns number are respectively 255 and 197, for a total of 50235 cells.

### 6.3 Topography

The topography is described by the Swiss elevation model DHM25 at 25 meter resolution (derived from topographic maps at scale 1:25'000), with a vertical accuracy of 2 m for the Ticino area. A 3D view of the basin, generated with the GRASS *nviz* command, is shown in figure 6.4.

The DTM over the basin area has minimum value of 193 m, maximum value of 3596 m, and mean value of 1608 m a.s.l.

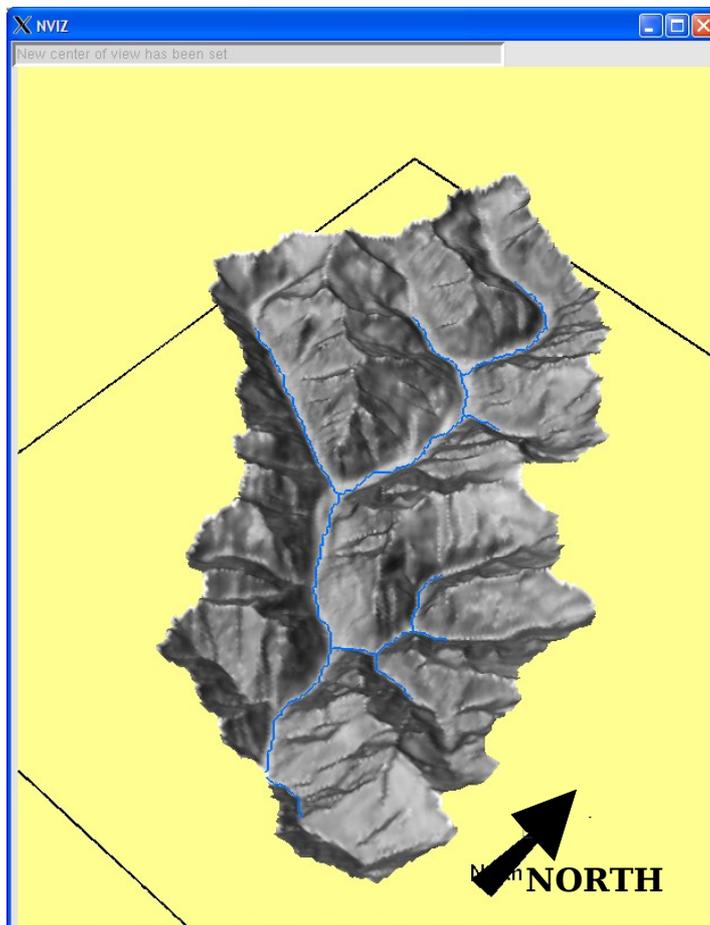


Figure 6.4 – 3D view of the Verzasca basin.

## 6.4 DTM derived maps

As discussed in paragraph 4.6, in order to calculate the runoff discharges some maps has to be derived from the elevation model. The DTM has been resampled at 100 m. cell size by mean of the `r.resample` GRASS command, this because the other informations (e.g. Landuse and soil map) are available with this resolution. In this section the application of the procedures described in chapter 4 and the generated map are shown and discussed.

### 6.4.1 Filled DEM

The digital elevation model has been modified in order to remove sinks, that generally would not permit the continuous flow over the basin. This map can be either generated with the `r.fill.dir` or the `r.terraflow` commands (see figure 6.5). The second one has a more sophisticate algorithm that better solves problems of flow in lowland areas and thus it has been preferred (`r.terraflow elev=dem filled=dem_fill dir=direction watershed=watershed accumulation=accu tci=tci;`).

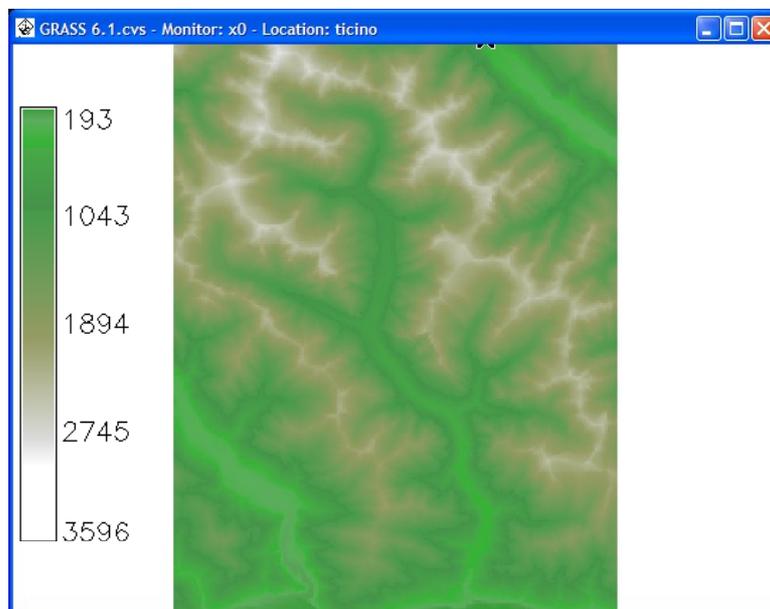


Figure 6.5 - Filled DEM [m] generated with the `r.terraflow` algorithm.

Unfortunately the maps of flow direction and flow accumulation that can be generated by *r.terraflow* cannot be used by the hydroFOSS model.

In fact the *r.terraflow* command solve the flow directions with a so called “D8-MFD” algorithm (Tarboton, 1997) that spread the flow direction among all the neighbour cells with lower elevation.

While, as discussed in chapter 4.6, for sake of simplicity the *h.hydrofoss.runoff* model solves the runoff only in one of the four cardinal direction, and therefore the flow direction map has to be extracted with the so called “D4-SFD” algorithm.

For this reason the *r.watershed* command has been chosen (as described in the following section) and a filled DEM is no more necessary because the used algorithm directly override sink problems.

In the next pages the command applied for this purpose has been described.

#### 6.4.2 Stream, basin, flow accumulation, and flow direction maps

Following the procedure reported in chapter 4.6 all these maps have been created with the *r.watershed* command:

```
r.watershed -4 elevation=dem_fill threshold=100  
accumulation=ACCUMUL drainage=DIRECTION basin=BASIN  
stream=STREAM;
```

The threshold value of 100 has been chosen among some test values because it is the one that, in the Verzasca basin case, permits to extract a stream map in good agreement with the “observed” river layer from the Swiss “vector25” maps of the Canton Ticino: the fitness of the two maps is represented in figure 6.6 and 6.7.

In figure 6.8, 6.9, and 6.10 the basin, the flow accumulation, and the flow direction maps are shown.

Note that the basins that are not completely defined within the region limits cannot be solved in the runoff process, because the incoming flow could be unknown.

Chapter 6 – Case study: the Verzasca basin

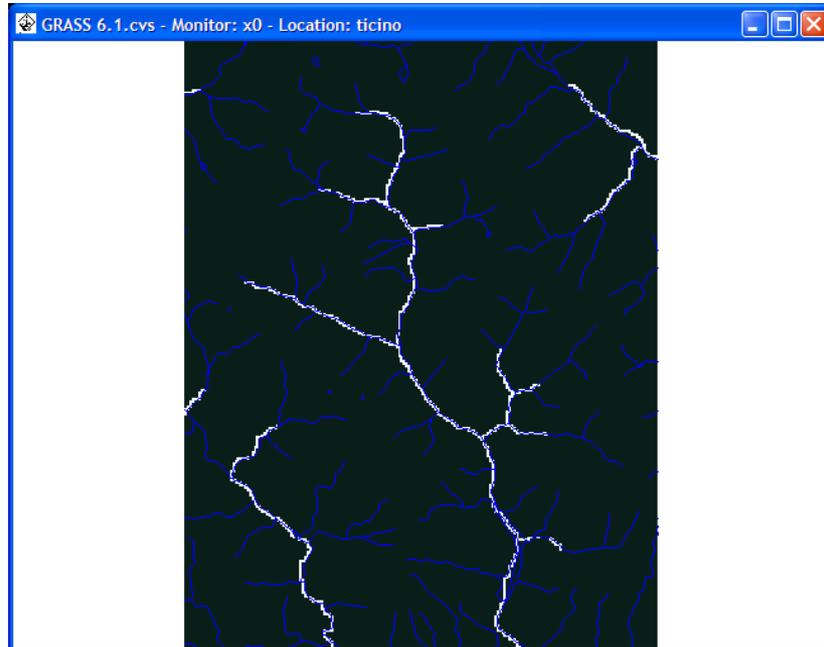


Figure 6.6 – extracted river (white) and “observed” river (blue) (threshold=100).



Figure 6.7 – zoom of extracted river (white) and “observed” river (blue) (threshold=100).

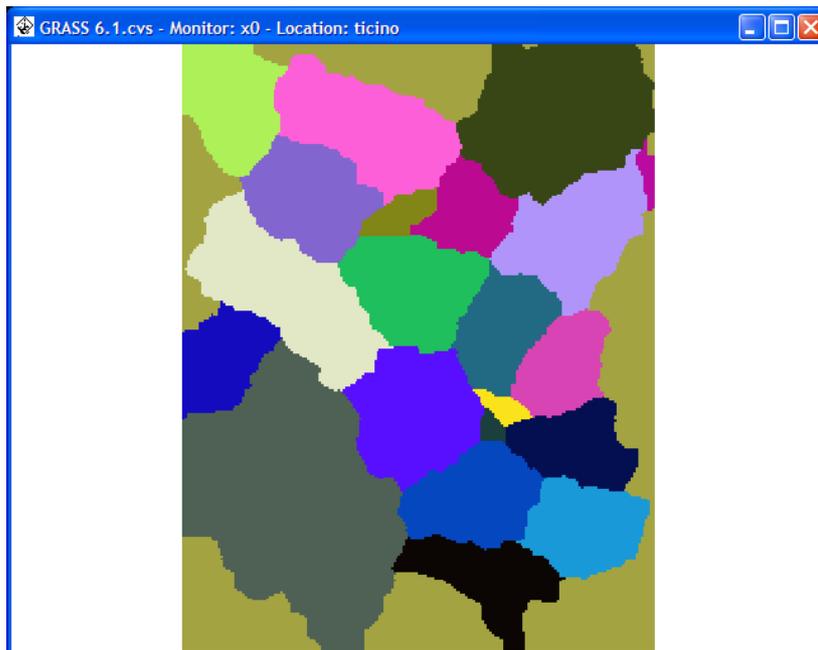


Figure 6.8 – Basins map calculated with the *r.watershed* command (threshold=100).

The flow accumulation map defines the number of cells that drain through each cell.

The *r.watershed* command extracts a map whose absolute cell values represent the amount of upland cells plus one. Negative numbers indicate that those cells possibly have surface runoff from outside of the current geographic region.

Thus the not completely defined basins will have in the flow accumulation map (figure 6.9) some cells with negative values and will not be solved in the *h.hydroFOSS.runoff* module.

It is therefore really important to define the correct region limits containing all the basins that have to be processed.

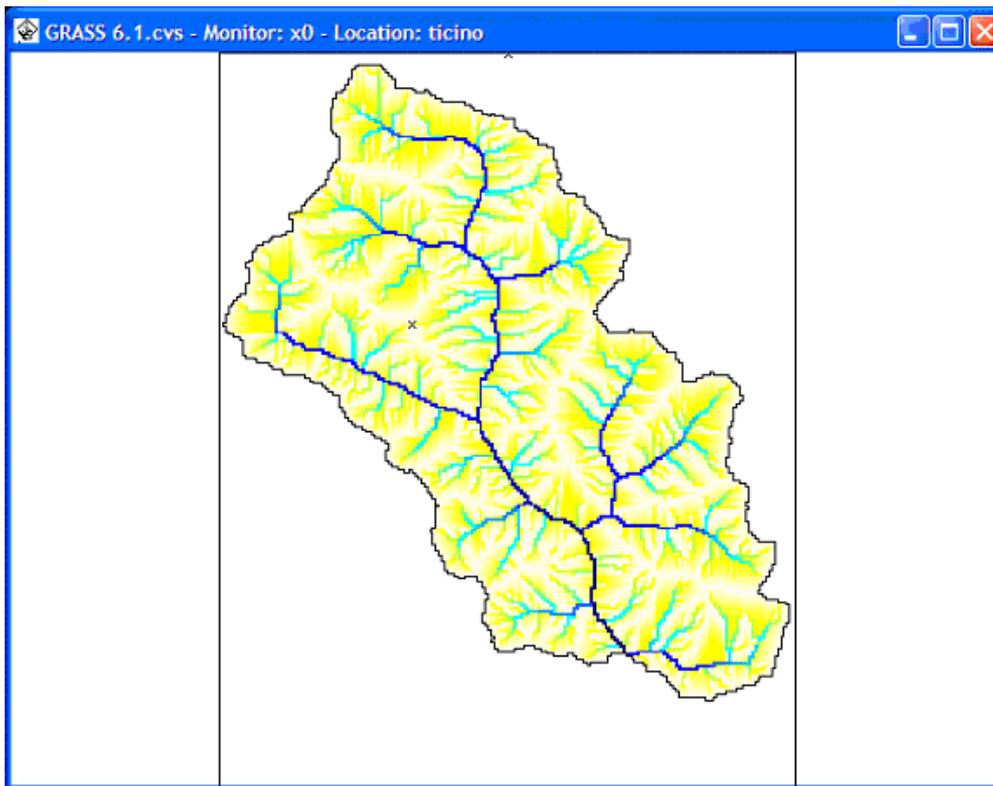


Figure 6.9 – Flow accumulation map of the Verzasca region.

The flow direction map represents the drainage direction.

The `r.watershed` command extracts a map whose absolute cell values, multiplied by 45 will give the direction in degrees that the surface runoff will travel from that cell. Negative values indicate that surface runoff is leaving the boundaries of the current geographic region.

Because the `r.watershed` has been run with the `-4` flags, that indicate the usage of the “D4” algorithm, only the absolute values 2 (90°, North), 4 (180°, West), 6 (270°, South), and 8 (360°, East) are present (see figure 6.10) while some negative values can be noted in correspondence of the region borders.

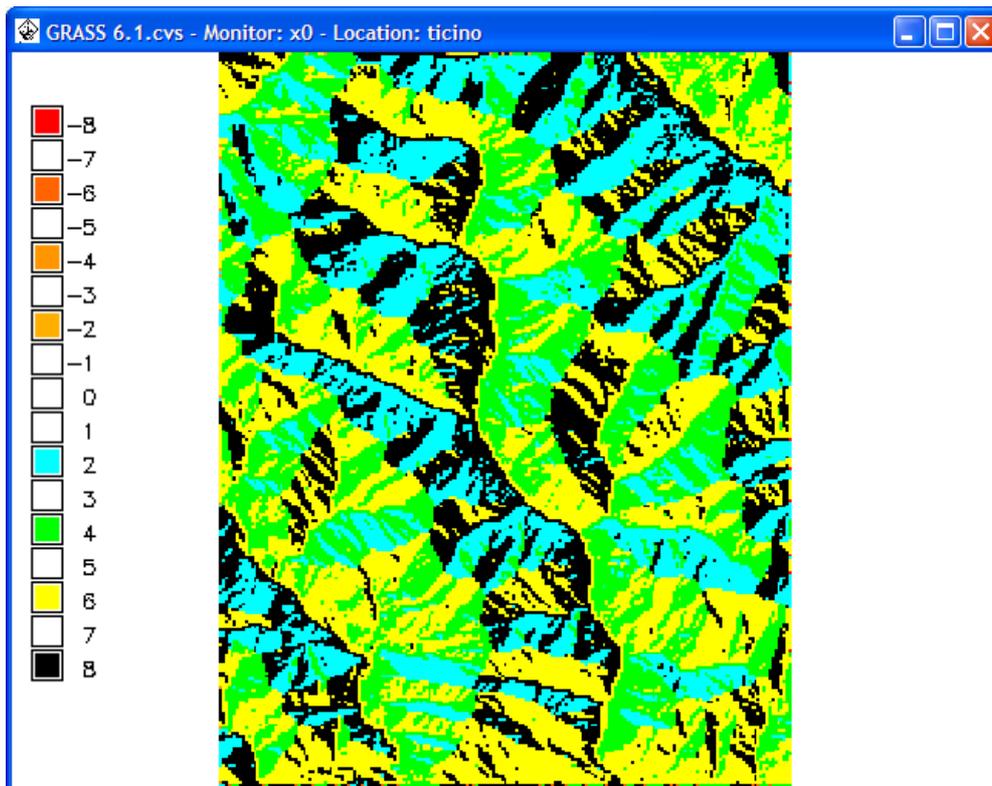


Figure 6.10 – Flow direction map for the Verzasca region.

As the basin is not too large and our computer power is elevated (one cell requires 132 byte of memory) we decided to aggregate the basins; this operation can be done with two commands:

- `r.statistics` in order to extract a map with the minimum accumulation value for each basin: `"r.statistics base=BASIN cover=ACCUMUL method=min output=min_accumul;"`
- `r.mapcalc` in order to generate two main basins, one completely defined in the region and one with all the other basins: `"r.mapcalc 'VerzBasin=if(@min_accumul>0,1,2)'"`

The new generated basins map (see figure 6.11) will be used in the `h.hydroFOSS.runoff` command.

For the runoff processing it will require 2.75 MB (21'837 cells \* 132 byte = 2'882'484 byte).

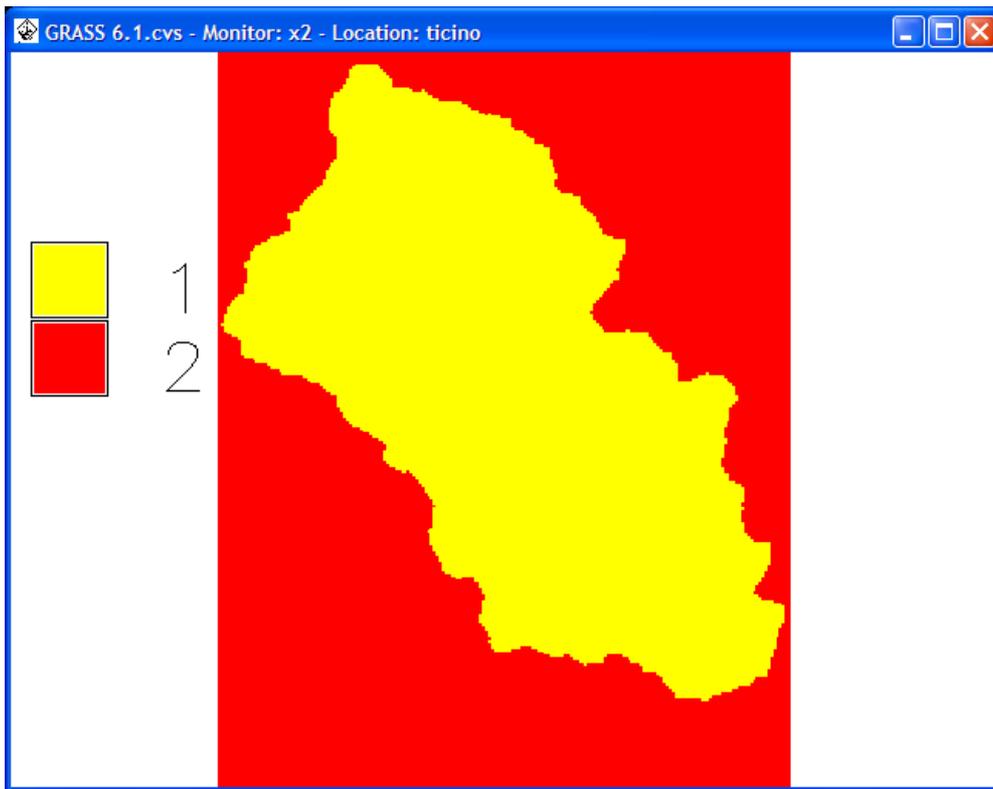


Figure 6.11 – Accumulated basins for runoff calculation.

6.4.3 Slope map

The slope map can be derived by running the `r.slope.aspect` command with the filled DEM as input map:

```
"r.slope.aspect elevation=dem_fill slope=slope;"
```

Figure 6.12 shows the generated map (masked with the Verzasca basin map), while the table 6.1 shows same basic statistics.

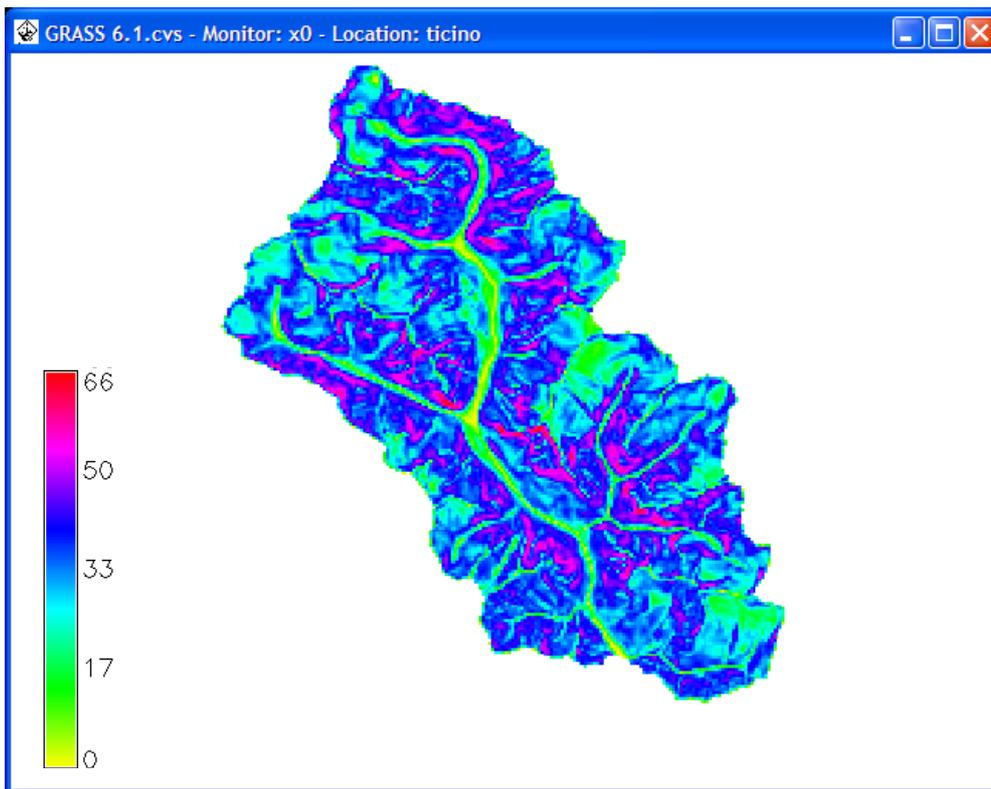


Figure 6.12 – Slope map [deg] generated with *r.slope.aspect* algorithm.

Average	36.72
Stdv	10.44
Min	0.1
Max	66.34

Table 6.1 – statistical values for slope map in degrees;

## 6.5 Landuse

The used landcover map is the “Arealstatistik 1972”, derived from topographic maps. It was elaborated from the Swiss Federal Office of Statistic which classified the landuse in 12 categories over a grid with resolution of 100 m according to the dominance principle that assigns to the cell the most extended category.

Ten of these categories are present in the Verzasca basin, while one additional class (river) has been added, for hydrological modelling purposes, superimposing the extracted stream map.

In fact, due to the dominance principle, the water class do not represent all the cell where the river is present while the characterization of the river properties is required from the r.hydrofoss.init module (see section 4.6.1).

Figure 6.13 shows the modified “Arealstatistik 1972” map for the test case area, while table 6.2 shows class statistics and figure 6.14 shows classes distribution over the entire basin where class with small extension are not represented.

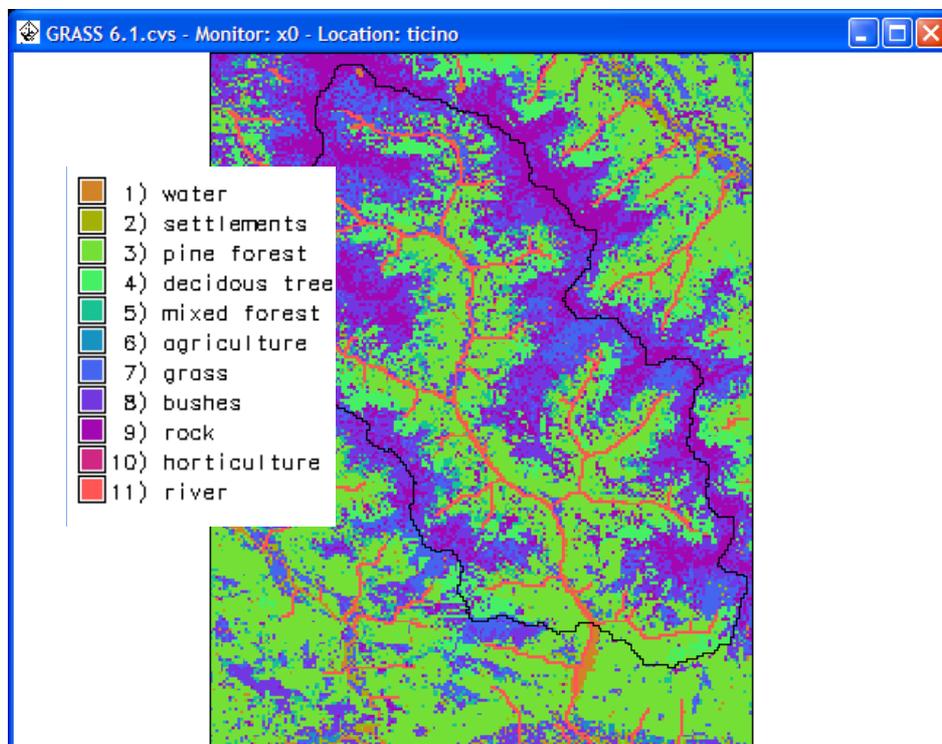


Figure 6.13 - modified “Arealstatistik 1972” map over the Verzasca basin.

## Chapter 6 – Case study: the Verzasca basin

class	classification	area (Km <sup>2</sup> )	cells	Area %
1	water	2.93	293	1.34%
2	settlements	0.64	64	0.29%
3	pine forest	53.43	5340	24.47%
4	deciduous tree	25.59	2557	11.72%
5	mixed forest	11.47	1146	5.25%
6	agriculture	0.11	11	0.05%
7	grass	56.08	2001	9.17%
8	bushes	20.02	5605	25.68%
9	rock	37.40	3738	17.13%
10	horticulture	0.01	1	0.01%
11	river	10.83	1082	4.95%

Tbale 6.2 – landuse classes statistics for the Verzasca basin.

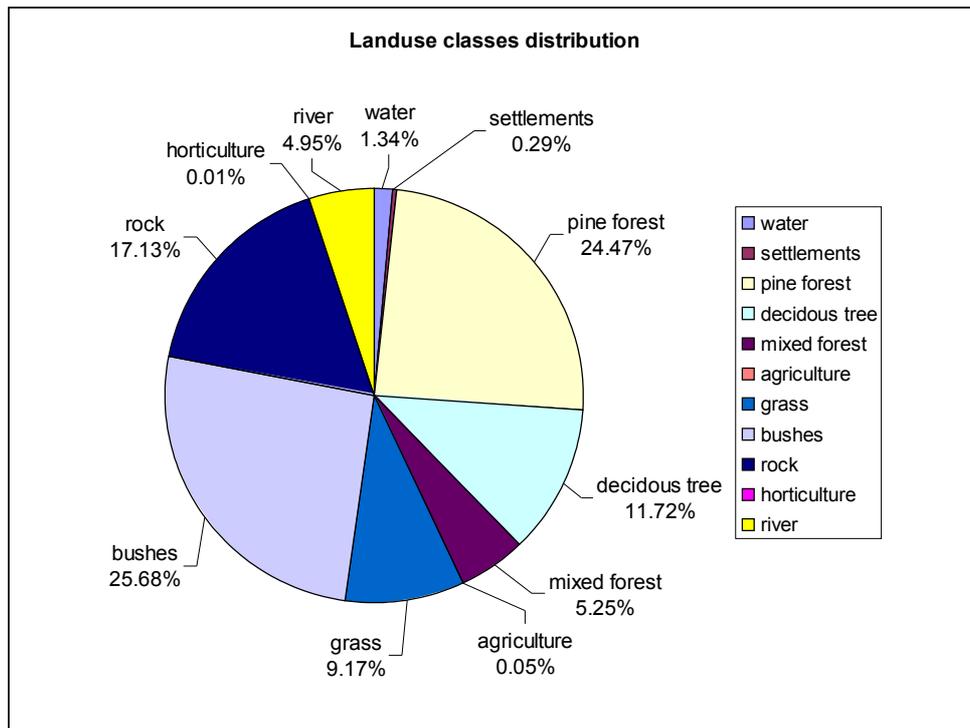


Figure 6.14 – landuse classes distribution in the Verzasca basin.

It is clear how the major landuse classes for the Verzasca basin are the pine forests (class 3), the bushes (class 8), and the rocks (class 9).

## 6.6 Soiltype

The soil map of the basin has been based on the Swiss soil data set (known as “Bodeneignungskarte”), which includes information about the suitability of soils for agricultural use at a resolution of 100 m. General information about soil dependent parameters was retrieved from the so-called “Bodeneignungskarte” of Switzerland, which describes soil attributes according to a proprietary classification that aims at evaluating the agricultural attitude of soils. The depths are therefore referred only to those soil layers that are influenced by biological activities.

In figure 6.15 the soiltype maps for the Verzasca basin is shown, while in table 6.3 and figure 6.16 information on the soiltype classes distribution are represented.

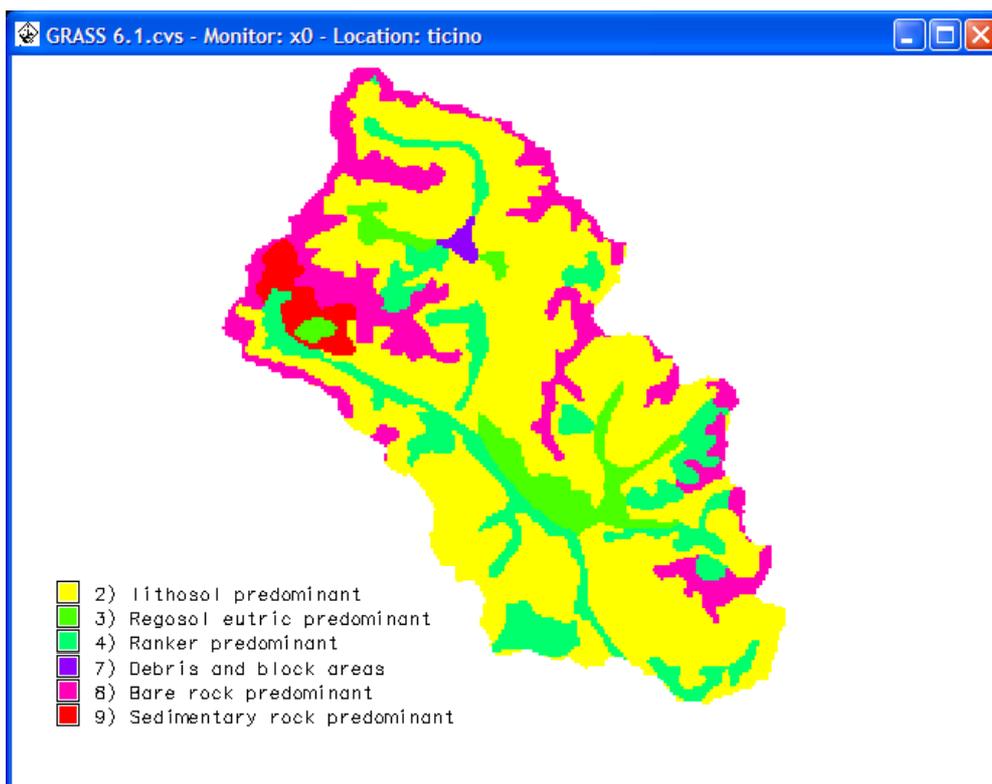


Figure 6.15 – map of soiltypes found in the Verzasca basin (derived from the “Bodeneignungskarte”).

## Chapter 6 – Case study: the Verzasca basin

class	classification	area (m <sup>2</sup> )	cells	Area %
2	Lithosol predominant	130470536	13039	59.71%
3	Regosol eutric predominant	12887955	1288	5.90%
4	Ranker predominant	33080420	3306	15.14%
7	Debris and block areas	1030636	103	0.47%
8	Bare rock predominant	35762075	3574	16.37%
9	Sedimentary rock predominant	5273255	527	2.41%

Table 6.3 – soiltype classes statistics for the Verzasca basin.

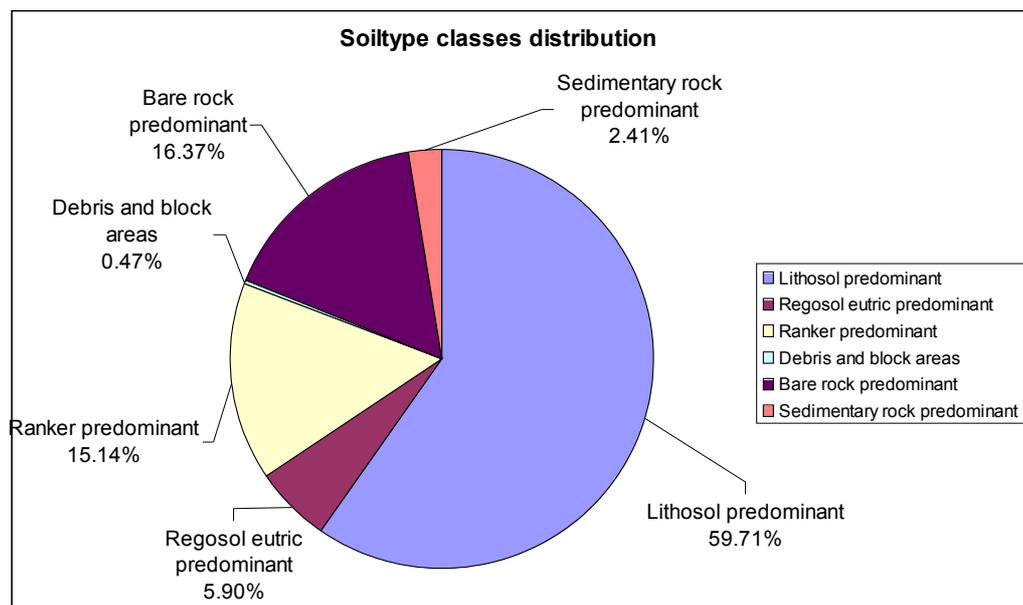


Figure 6. 16 – soiltype classes distribution in the Verzasca basin.

Lithosol predominant (class 2) is largely the most diffuse soil type in the Verzasca area, and therefore it is expected that their derived parameters will greatly influence the model simulations.

## 6.7 Riverwidth

Not commonly adopted methodology to estimates the riverwidth (rw) can be found in literature, thus a simple but logic procedure has been used. With the hypothesis that the river width grows following the drained area dimension, a linear factor growing with the flow accumulation values has been estimated by means of:

$$F_{rw} = [\max(rw)-\min(rw)]/[\max(acc)-\min(acc)] \quad [\text{eq. 6.1}]$$

where:

$\max(rw)$  is the maximum riverwidth [m],

$\min(rw)$  is the minimum riverwidth [m],

$\max(acc)$  is the maximum flow accumulation at  $\max(rw)$ ,

$\min(acc)$  is the minimum flow accumulation at  $\min(rw)$ ,

$F_{rw}$  is the riverwidth growing factor [m/accumulation];

After some “in situ” observations the river has been detected to reach a maximum width of 10 m.

Observing that the maximum flow accumulation of the river's cell is 28500 and the minimum is 100 (this is the threshold applied to derive the stream map) for the case study the used values are:

$\max(rw)=10$ ,  $\min(rw)=0$ ,  $\max(acc)=28500$ , and  $\min(acc)=100$

The resulting linear factor is therefore:  $F_{rw} = 0.00035$ .

The  $F_{rw}$  was then applied by means of map algebra to the flow accumulation map and the riverwidth map was extracted (figure 6.17) by means of the following command:

```
r.mapcalc 'riverwidth=if(accumulation>10,accumulation*0.00035,0)';
```

Furthermore the hydrological type map (figure 6.18) can be easily extracted from the riverwidth map by means of map algebra with the command:

```
r.mapcalc 'hydrotype=if(riwerwidth=0, 0, 1)';
```

## Chapter 6 – Case study: the Verzasca basin

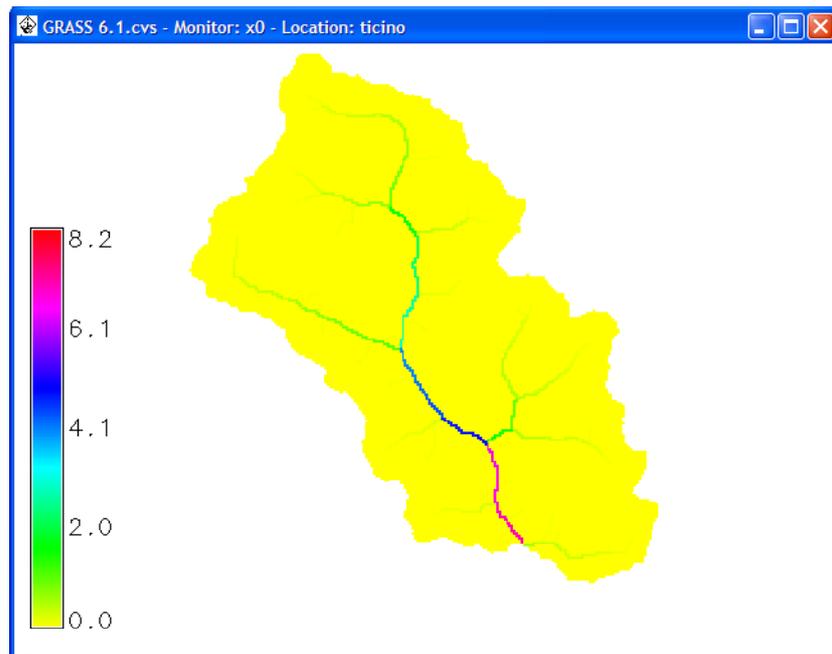


Figure 6.17 – riverwidth map for the Verzasca basin.

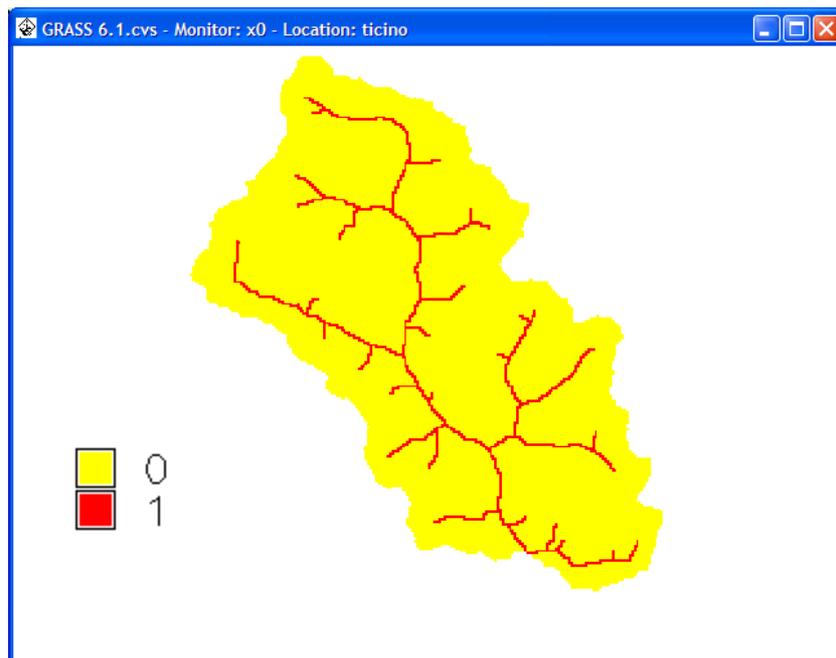


Figure 6.18 – hydrotype for the verzasca basin: 0=terrain, 1=river

The substantial correctness of the derived maps has been verified by means of visual comparing with aerial photos of the Verzasca basin.

### **6.8 Model verification and running strategy**

Known the basic maps (filled DTM, slope, soiltype, landuse, riverwidth, hydrotype, flow accumulation, flow direction, and basin) and defined the parameters tables (landuse\_cat and soiltype\_cat tables, as defined in section 4.6.1 and shown in tables 6.4 and 6.5) it is possible to run the simulation of the rainfall-runoff processes during a given time interval.

The starting values has been selected as follow:

- Manning (surface roughness coefficient), estimated from the literature (Chow, 1973) for each land cover class, and subsequently adjusted to match the timing concentration towards the drainage network;
- kappa\_s (horizontal permeability), estimated for each soil class by using values from the US Geological Survey (USGS) “South Florida Information Access” (SOFIA) database (<http://sofia.usgs.gov/publications/wri/02-4050/hydcondsa.html>, 2006);
- theta\_s and theta\_r (saturated and residual soil water content), estimated by assigning for each soil type values taken from the digital thematic soil map of Switzerland (Bodeneignungskarte, see section 6.6);
- alpha (transmissivity law exponent), was fixed at 2.5, being the value suitable for most of the soil as suggested in literature on the basis of the application of the Topkapi model to other basins; a lower value would imply a decrease in soil water drainage, therefore an increase in the surface runoff;
- thik (soil thickness), derived from the “Bodeneignungskarte” as an average value of the corresponding soil class in the soil map of switzerland adjusted within 30 cm.;

In order to verify and calibrate the model, the full script was divided in two parts:

- the first (h.hydroFOSS.pre) generates all the maps requested by the flows computation (climatic maps, solar radiation maps, clear

## Chapter 6 – Case study: the Verzasca basin

sky index map, global radiation maps, snow accumulation & melt maps, evapotranspiration maps, canopy interception maps, and available water for flows map);

- the second (h.hydroFOSS.cal) run and calibrates the runoff model.

<b>lu_id</b>	<b>cat_name</b>	<b>manning</b>
1	lakes	0.8
2	settlements	0.01
3	pine forest	0.2
4	deciduous tree	0.1
5	mixed_forest	0.15
6	agriculture	0.15
7	grass	0.18
8	bushes	0.09
9	rock	0.1
10	horticulture	0.15
11	rivers	0.04

Table 6.4 – landuse\_cat table.

<b>cat</b>	<b>smb</b>	<b>descr</b>	<b>kappa_s</b>	<b>teta_s</b>	<b>teta_r</b>	<b>alfa</b>	<b>thik</b>
2	l	Lithosol	5.56E-005	0.524	0.141	2.5	1.1
3	re	Regosol	1.10E-004	0.582	0.123	2.5	1.05
4	ra	Ranker	1.78E-005	0.533	0.181	2.5	1.3
7	de	Debris	5.56E-005	0.524	0.141	2.5	1.1
8	rk	Bare rock	1.66E-004	0.506	0.123	2.5	1.1
9	se	Sediment	1.31E-004	0.451	0.123	2.5	0.9

Table 6.5 – soiltype\_cat table of the categories found in the Verzasca Valley.

The division of the model in two different scripts, allows the user to compute the raster series not affected by calibration only once and to use these maps, already extracted, as inputs for the further runoff running required by the calibration process. This solution

allows to save time and computational power.

The raster series was hourly calculated all along the 2004 with *h.hydroFOSS.pre* and verified:

- by displaying raster map series and visually verify the correctness spatial component;
- by plotting sampled series at samples points and visually verify the correct behaviour of the variables;
- by quantitatively compare annual estimated values with annual observed values.

The calibration was run for the whole August 2004: the used methods, described in chapter 4.8, was applied by comparing the simulations with the observations at two river gauges stations (Pincascia and Lavertezzo) that are shown in figure 6.19.

The parameters to be calibrated are those represented in the tables *landuse\_cat* and *soiltype\_cat*, and their values can change during the calibration process.

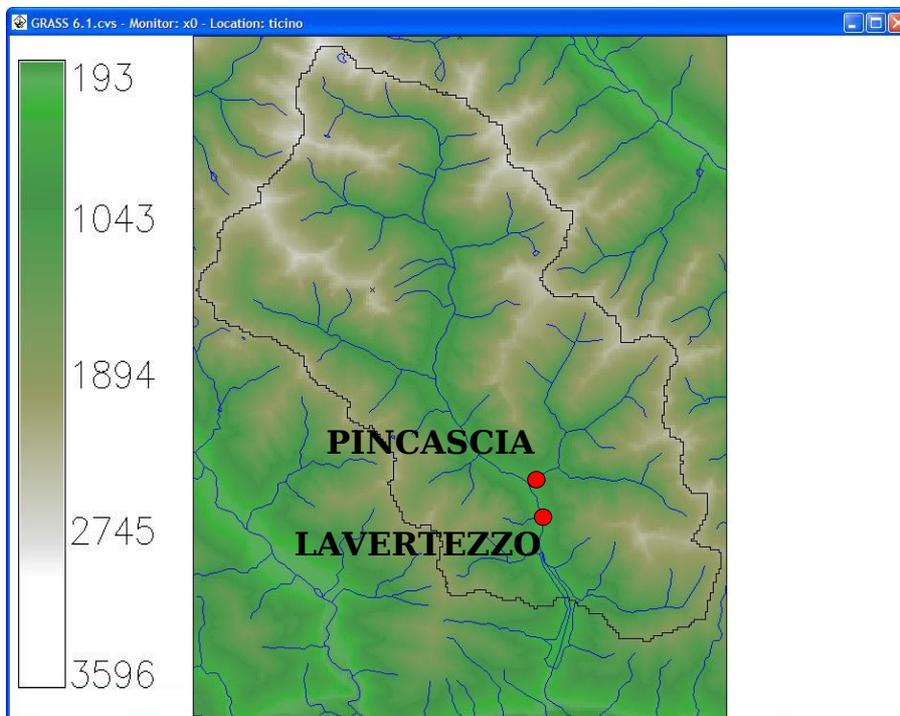


Figure 6.19 – Raiver gauges used in model calibration.

## **6.9 Climatological maps**

The processes described in section 4.1.2 was successfully applied, and the hourly climatic raster series was generated.

As already discussed, the application of Thiessen tassellation do not alters the observed values but just identify for each measure an area of influence where the observation is uniformly applied.

This kind of interpolation, used for the rainfall, wind, and humidity maps generation, produces maps with constant values that change with abrupt artificial break lines. This representation unlikely reproduce the reality and therefore a validation with other information would not be meaningful.

For this reason a full validation of these maps with other data type (e.g. remote sensing) cannot be performed. A simple check, that verified the absence of null values and missing maps, validated the generated series.

Furthermore, considering the general small number of climatic measurements points (in this case study a maximum of 5 station influencing the basin can be counted), a leave-one-out cross-validation procedure will result in being inadequate due to the great influence of every observation in the map generation: the omission of one of the already limited interpolation points would greatly affect the result.

From this point of view, the remotely sensed data, once demonstrated their correctness in quantities estimation, could greatly improve the representation of these fields.

Other improvements in the generation of these maps could be achieved by taking advantage of dependencies from other variables (like it has been done for the temperature maps interpolation in relation with the quota, see chapter 5.2).

Further investigation on each of these climatic variable (wind velocity, humidity, and rainfall) should be carried on in future studies.

Validation and investigation on temperature map series has been discussed in chapter 5.2.

## Chapter 6 – Case study: the Verzasca basin

In figures 6.20-27 some examples of rainfall maps (figure 6.20-21), wind velocity maps (figure 6.22-23), temperature maps (figure 6.24-25), and humidity maps (figure 6.26-27) calculated at different instants are shown.

With exception for the temperature maps, where a topographic variable is considered, these figures reveal the already mentioned presence of artificial break lines. This behaviour is going to influence all the maps derived by the usage of these climatological maps, as can be easily noted looking at the map derived with the snow module (Figures 6.34, 6.35, 6.36), the interception module (Figures 6.40, 6.41, 6.42) and the runoff module (Figures 6.53, 6.54, 6.55).

This kind of approach (Thiessen tassellation) for climatological maps extraction has been widely used mainly for two reason:

1. hydrologists are mostly interested in computing variables (mainly discharges) for an outlet, and therefore, in this way, they consider a sort of cumulative values (in space, zime and processes) that hardly suffer from those artificial boundary;
2. proven methodology for distributed observations (quantitatively valuables) or more dense samples are missing, therefore no sensible advantage can be achieved without exhaustive researches.

Anyway, still remaining known errors in the spatial description of same variables, as shown in the further sections, the general behaviour of the basin can still be simulated.

Chapter 6 – Case study: the Verzasca basin

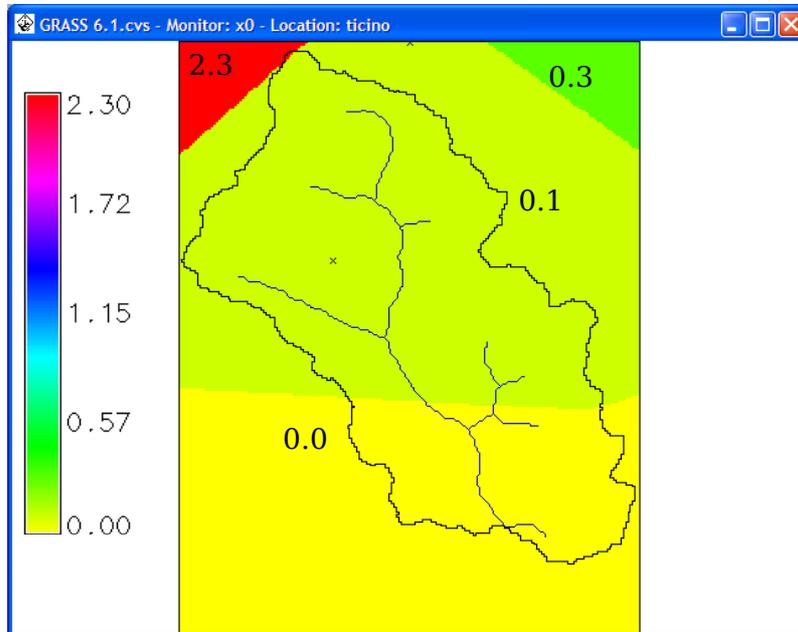


Figure 6.20 – Rainfall map [mm/h] for the 19/08/2004 at 18:00

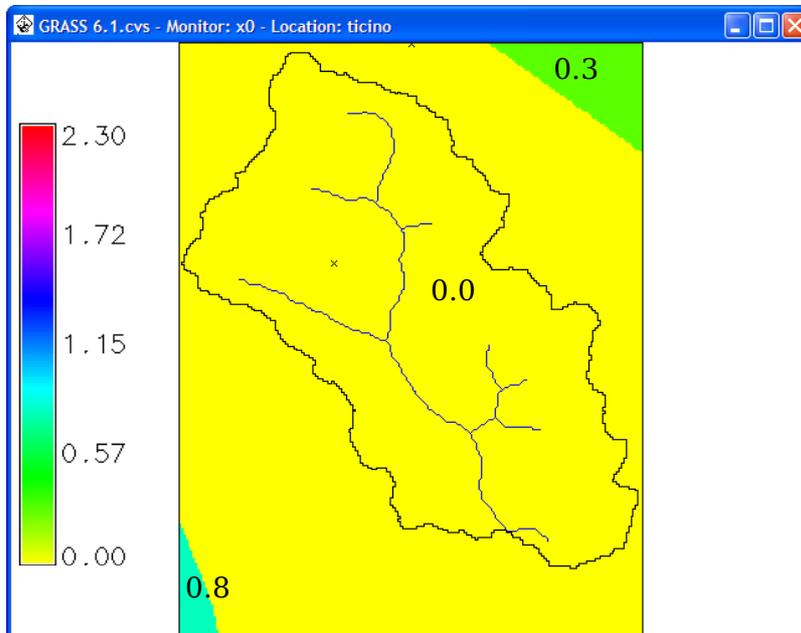


Figure 6.21 – Rainfall map [mm/h] for the 20/08/2004 at 18:00

Chapter 6 – Case study: the Verzasca basin

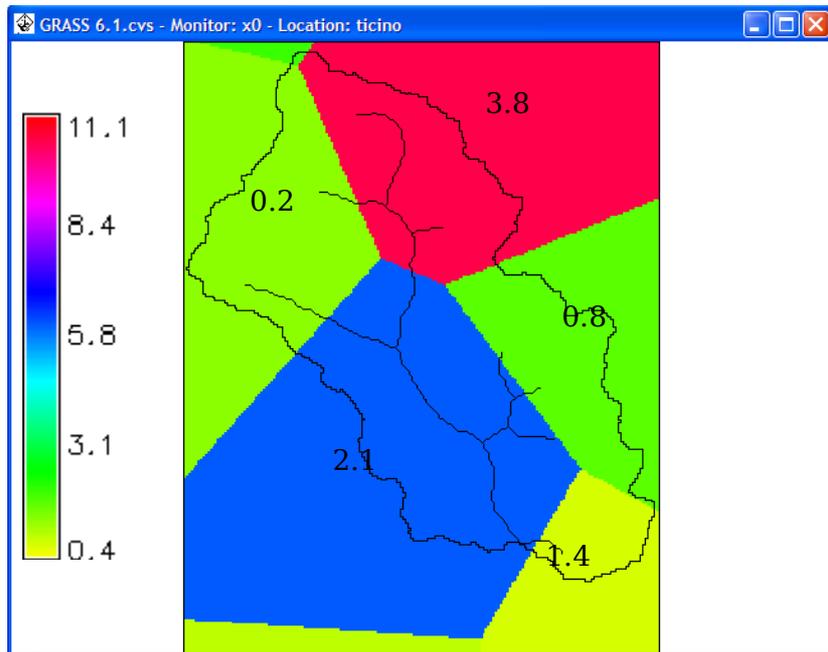


Figure 6.22 – Wind velocity [m/s] map for the 19/08/2004 at 18:00

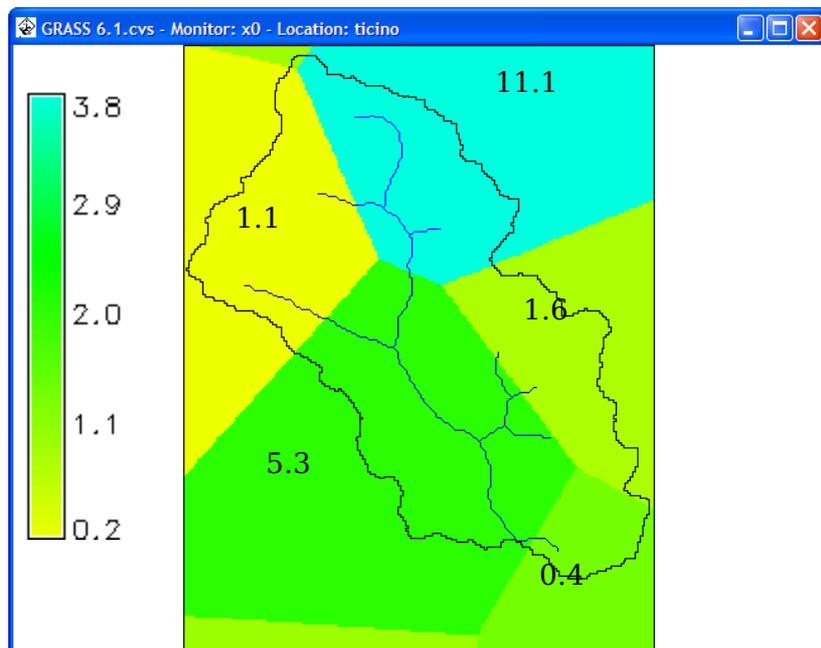


Figure 6.23 – Wind velocity [m/s] map for the 20/08/2004 at 18:00

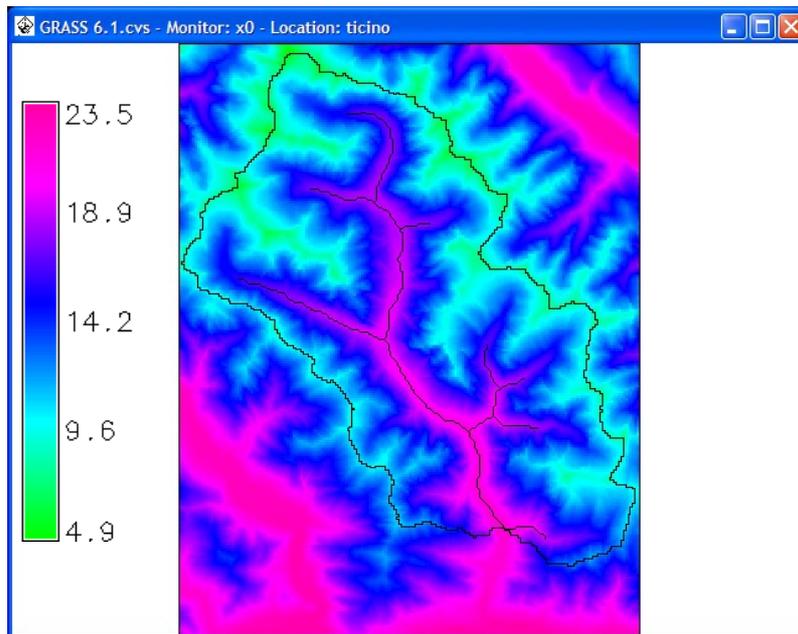


Figure 6.24 – Temperature [°C] map for the 20/08/2004 at 18:00 (day)

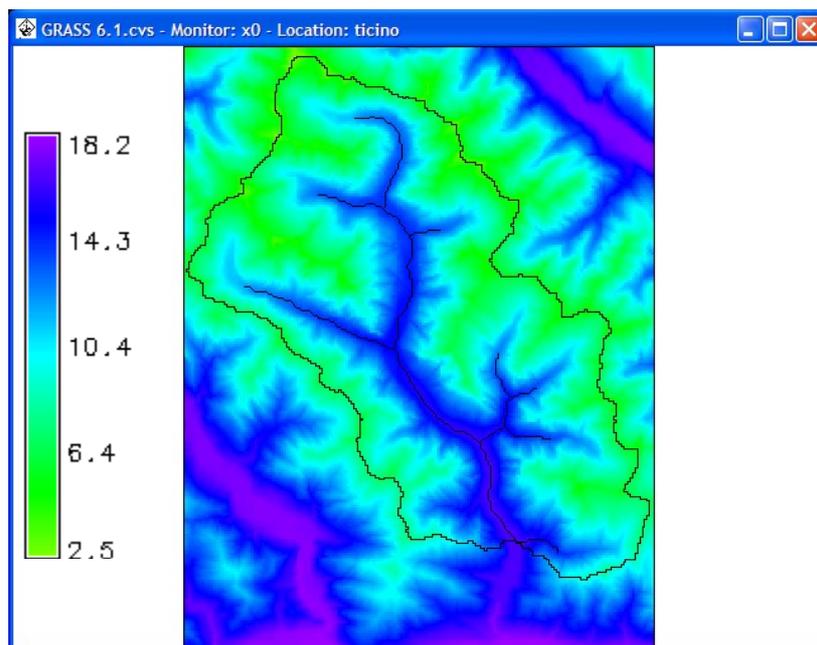


Figure 6.25 – Temperature [°C] map for the 21/08/2004 at 01:00 (nigh)

## Chapter 6 – Case study: the Verzasca basin

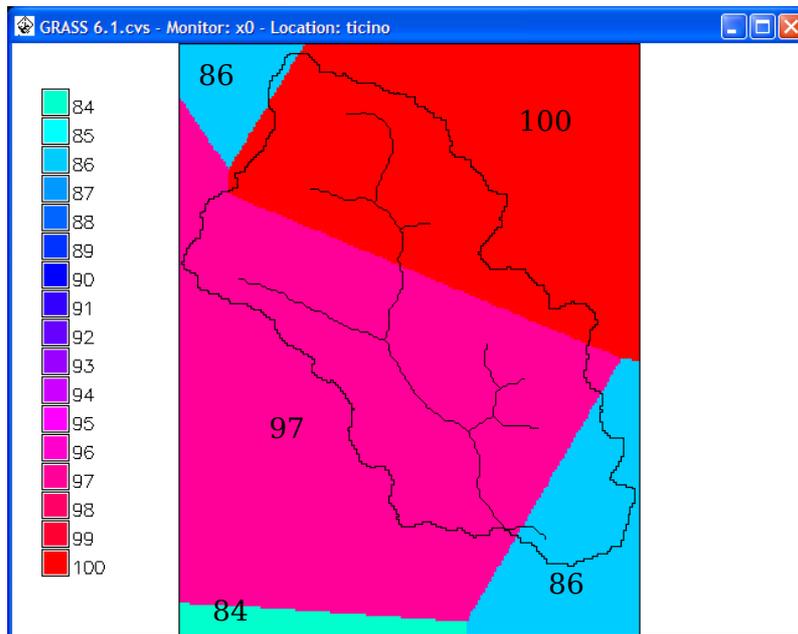


Figure 6.26 – Humidity [%] map for the 19/08/2004 at 18:00

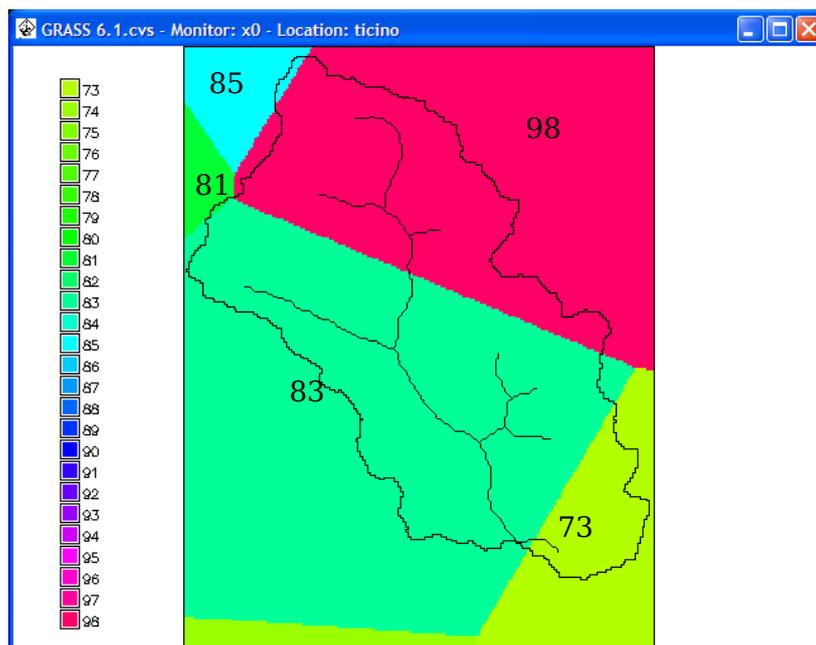


Figure 6.27 – Humidity [%] map for the 20/08/2004 at 18:00

### 6.10 Watershed characterization maps

These maps has been extracted from literature values with the commands *h.seasonal.par* and *h.monthly.par* described in section 4.1.3. The used monthly parameters table (table 6.6) and seasonal parameters table (table 6.7) were derived from a WaSiM model application in the Alpine region (Schulla and Jasper, 1999).

Since the albedo is time independent, in table 6.6, the same values are repeated through columns.

PAR	RASTER	CAT	JEN	FEB	..	..	..	..	..	..	SET	OCT	NOV	DEC
alb	landuse cat	1	0.05	0.05	..	..	..	..	..	..	0.05	0.05	0.05	0.05
alb	landuse cat	2	0.1	0.1	..	..	..	..	..	..	0.1	0.1	0.1	0.1
alb	landuse cat	3	0.12	0.12	..	..	..	..	..	..	0.12	0.12	0.12	0.12
alb	landuse cat	4	0.17	0.17	..	..	..	..	..	..	0.17	0.17	0.17	0.17
alb	landuse cat	5	0.15	0.15	..	..	..	..	..	..	0.15	0.15	0.15	0.15
alb	landuse cat	6	0.25	0.25	..	..	..	..	..	..	0.25	0.25	0.25	0.25
alb	landuse cat	7	0.25	0.25	..	..	..	..	..	..	0.25	0.25	0.25	0.25
alb	landuse cat	8	0.2	0.2	..	..	..	..	..	..	0.2	0.2	0.2	0.2
alb	landuse cat	9	0.12	0.12	..	..	..	..	..	..	0.12	0.12	0.12	0.12
alb	landuse cat	10	0.25	0.25	..	..	..	..	..	..	0.25	0.25	0.25	0.25
alb	landuse cat	11	0.05	0.05	..	..	..	..	..	..	0.05	0.05	0.05	0.05
rsc	landuse cat	1	20	20	..	..	..	..	..	..	20	20	20	20
rsc	landuse cat	2	100	100	..	..	..	..	..	..	100	100	100	100
rsc	landuse cat	3	80	80	..	..	..	..	..	..	55	75	80	80
rsc	landuse cat	4	100	100	..	..	..	..	..	..	65	85	100	100
rsc	landuse cat	5	90	90	..	..	..	..	..	..	60	80	90	90
rsc	landuse cat	6	80	80	..	..	..	..	..	..	65	75	90	90
rsc	landuse cat	7	90	90	..	..	..	..	..	..	60	70	90	90
rsc	landuse cat	8	80	80	..	..	..	..	..	..	55	70	80	80
rsc	landuse cat	9	250	250	..	..	..	..	..	..	250	250	250	250
rsc	landuse cat	10	100	100	..	..	..	..	..	..	60	80	100	100
rsc	landuse cat	11	20	20	..	..	..	..	..	..	20	20	20	20

Table 6.6 – monthly parameters table.

The used seasonal parameters, of table 6.7, were referred to the four reference days of the year: DAY1=110, DAY2=150, DAY3=250, and DAY4=280. Also here some values are repeated trough columns: they are clearly no time dependent.

Chapter 6 – Case study: the Verzasca basin

PAR	RASTER	CAT	DAY1	DAY2	DAY3	DAY4
LAI	landuse cat	1	1	1	1	1
LAI	landuse cat	2	1	1	1	1
LAI	landuse cat	3	8	12	12	8
LAI	landuse cat	4	0.5	8	8	0.5
LAI	landuse cat	5	2	10	10	2
LAI	landuse cat	6	1	5	3	1
LAI	landuse cat	7	2	4	4	2
LAI	landuse cat	8	3	5	5	3
LAI	landuse cat	9	1	1	1	1
LAI	landuse cat	10	0.5	5	5	0.5
LAI	landuse cat	11	1	1	1	1
vegh	landuse cat	1	0.01	0.01	0.01	0.01
vegh	landuse cat	2	10	10	10	10
vegh	landuse cat	3	10	10	10	10
vegh	landuse cat	4	0.3	10	10	0.3
vegh	landuse cat	5	3	10	10	3
vegh	landuse cat	6	0.05	0.5	0.2	0.05
vegh	landuse cat	7	0.15	0.4	0.3	0.15
vegh	landuse cat	8	1.5	2.5	2.5	1.5
vegh	landuse cat	9	0.05	0.05	0.05	0.05
vegh	landuse cat	10	0.4	3	3	0.4
vegh	landuse cat	11	0.01	0.01	0.01	0.01
veg	landuse cat	1	0.1	0.1	0.1	0.1
veg	landuse cat	2	0.5	0.5	0.5	0.5
veg	landuse cat	3	0.9	0.9	0.9	0.9
veg	landuse cat	4	0.7	0.95	0.95	0.7
veg	landuse cat	5	0.8	0.92	0.92	0.8
veg	landuse cat	6	0.3	0.8	0.7	0.3
veg	landuse cat	7	0.95	0.95	0.95	0.95
veg	landuse cat	8	0.9	0.95	0.95	0.9
veg	landuse cat	9	0.8	0.8	0.8	0.8
veg	landuse cat	10	0.75	0.75	0.75	0.75
veg	landuse cat	11	0.1	0.1	0.1	0.1

Table 6.7 - seasonal parameters table.

The maps generated with the *h.monthly.par* commands have been derived with a time resolution of one month, while the ones generated with the *h.seasonal.par* command have been extracted with a time resolution of one week.

Higher time resolution are not justified, because the approximation of the parameters themselves does not permit more precise estimates.

These maps generated by means of literature values spatialization, take in account the mean behaviour of these quantities (seasonal or monthly general trend) but do not consider any particular

conditions.

For instance, the LAI (Leaf Areal Index) can be greatly affected by the presence of dry periods that would limit the vegetation growth.

In this sense remote sensed maps of LAI, vegetation height, vegetation cover, and albedo, being an updated observation of the reality, could be a really valuable resource for a better watershed characterization.

The investigation of remotely sensed data is not the argument of this thesis that indeed focus its attention to the realization of a open source GIS embedded hydrological model, therefore further research in this area will be remanded to future investigations.

The generated map series has been verified by visually comparing their maps and checking their plausibility both in spatial distribution (e.g. reflecting the landuse distribution) and in time distribution (e.g. growing LAI values from winter to summer).

Like for climatic maps a simple check to verify the absence of missing maps or nulls values in the generated series was successfully executed.

In the top of figure 6.28 the albedo map for the August month is shown. Due to its time independence, the albedo raster series shows constant values through months; the accordance with the landuse map (see figure 6.13) can be observed.

In order to take into account the presence of the snow, before the running of the snow accumulation and melt module, the albedo map will updated by means of map algebra. Where, at the previous step, the snowpack water content was not zero the current albedo map assume a LAI value of 3.8, value estimated to be a mean value of the snow reflectance (Tarboton, 1996). The bottom of figure 6.28 show an example of updated albedo map for the 15<sup>th</sup> August in full agreement with the estimated snow coverage (see figure 6.34).

In figure 6.29, 6.30, 6.31 two maps of the vegetation height, vegetation cover and LAI, one for winter time and one for summer time, are represented as an example.

In figure 6.29 the correspondence with the modified landuse map can be noted, in fact the river cells are well represented with low

## Chapter 6 – Case study: the Verzasca basin

vegetation height ( $\sim 0.1$  m) and the pine forest cells have larger vegetation height values ( $\sim 10$  m). Moreover the extension of red cells (indicating higher values) became larger in summer due to the growing of the deciduous tree and mixed forest (respectively from  $\sim 3$  and  $\sim 0.3$  to  $\sim 10$  m).

Also in figure 6.30 the good correspondence with the landuse map can be noted. Moreover the full blooming of the deciduous tree and mixed forest in the summer season shows large values of coverage ( $\sim 90\%$ ) if compared to the winter season values ( $\sim 0.7\%$ ), all in full agreement with the vegetation growing cycle.

Comparing figure 6.31 with figure 6.13 the correspondence with the landuse can be identified. The comparison between winter map and summer map reveal, in full accordance with the evolution of the vegetation, a general increasing LAI values from January to August. This behaviour, motivated by the increasing leaf area of the vegetation, is not present in the water, settlement, rock, and river landuse classes where vegetation is mostly absent.

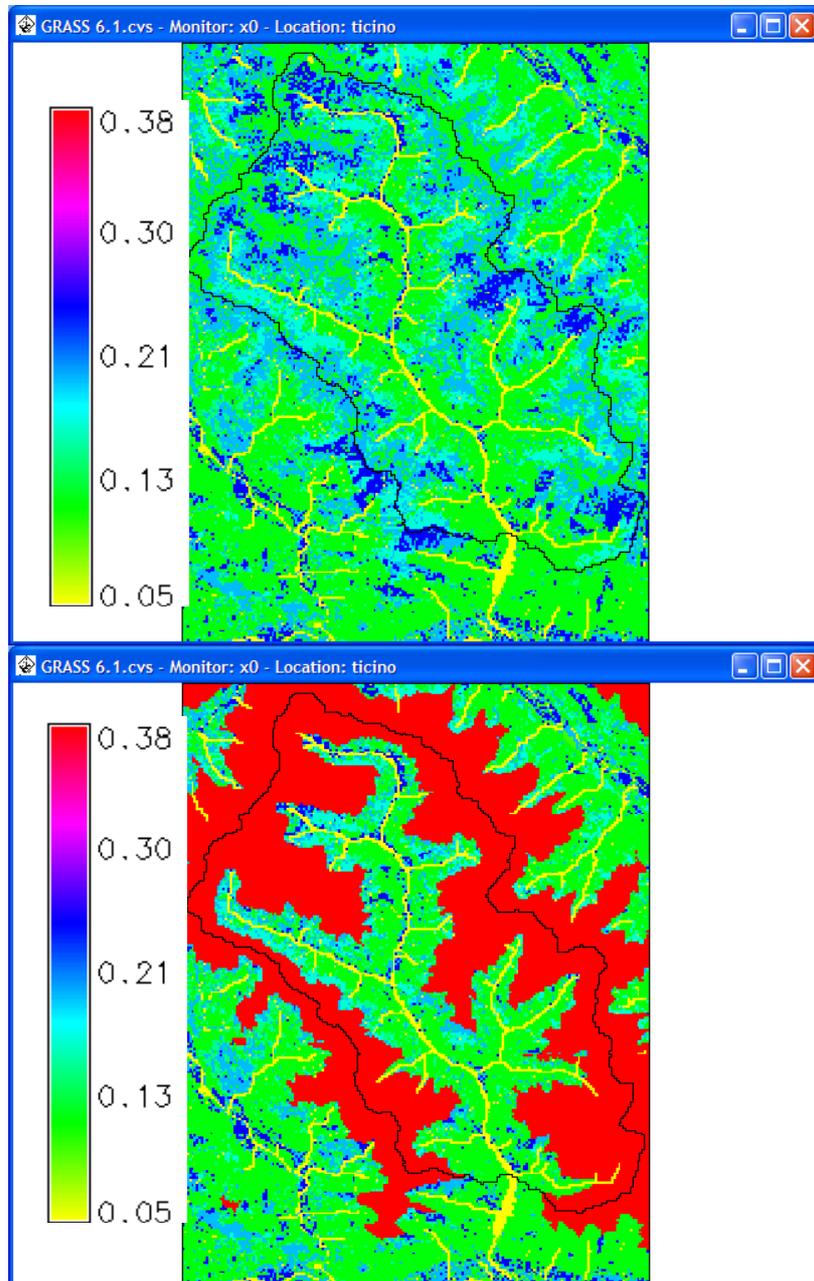


Figure 6.28 – albedo map for August 2004 (top) and updated albedo map for the 15<sup>th</sup> August 2004.

## Chapter 6 – Case study: the Verzasca basin

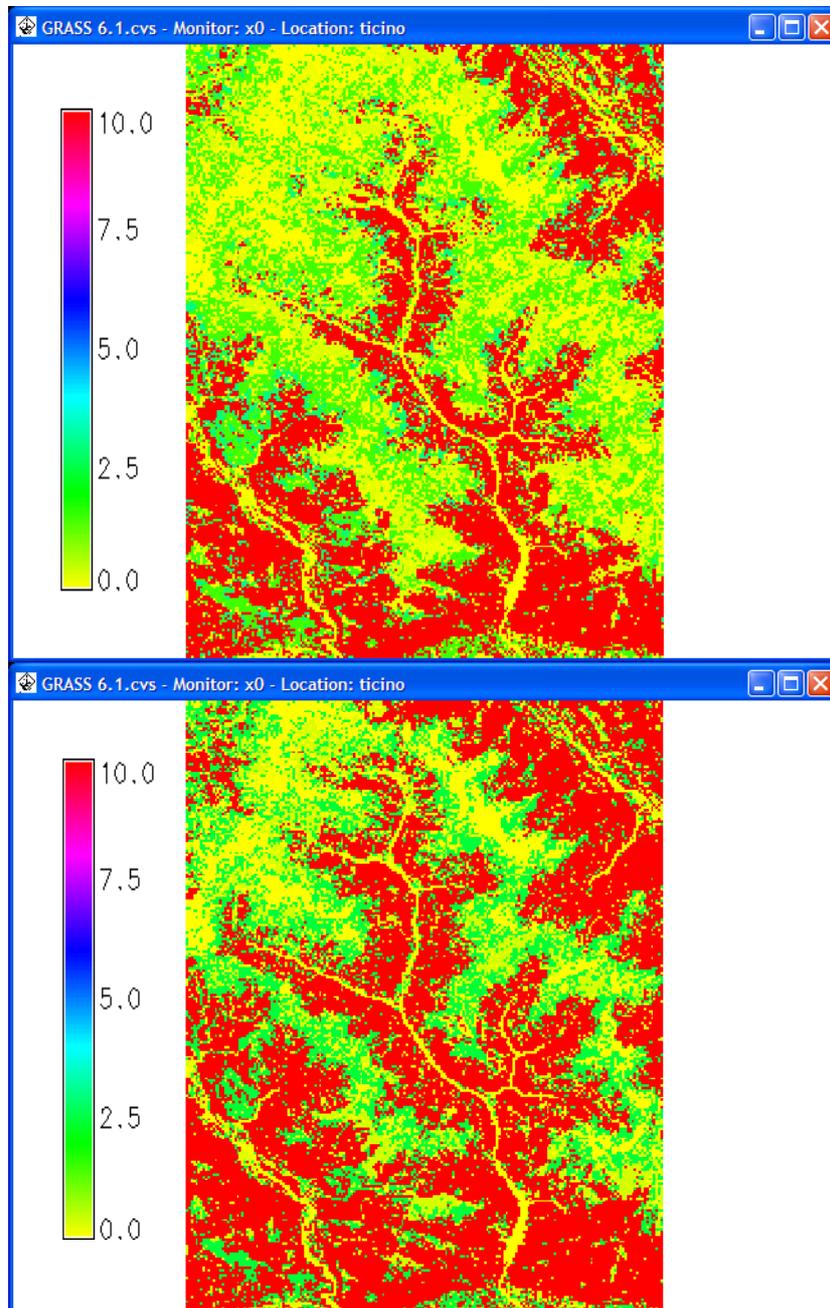


Figure 6.29 – vegetation height [m] on the 15<sup>th</sup> January (top) and 15<sup>th</sup> August (bottom).

Chapter 6 – Case study: the Verzasca basin

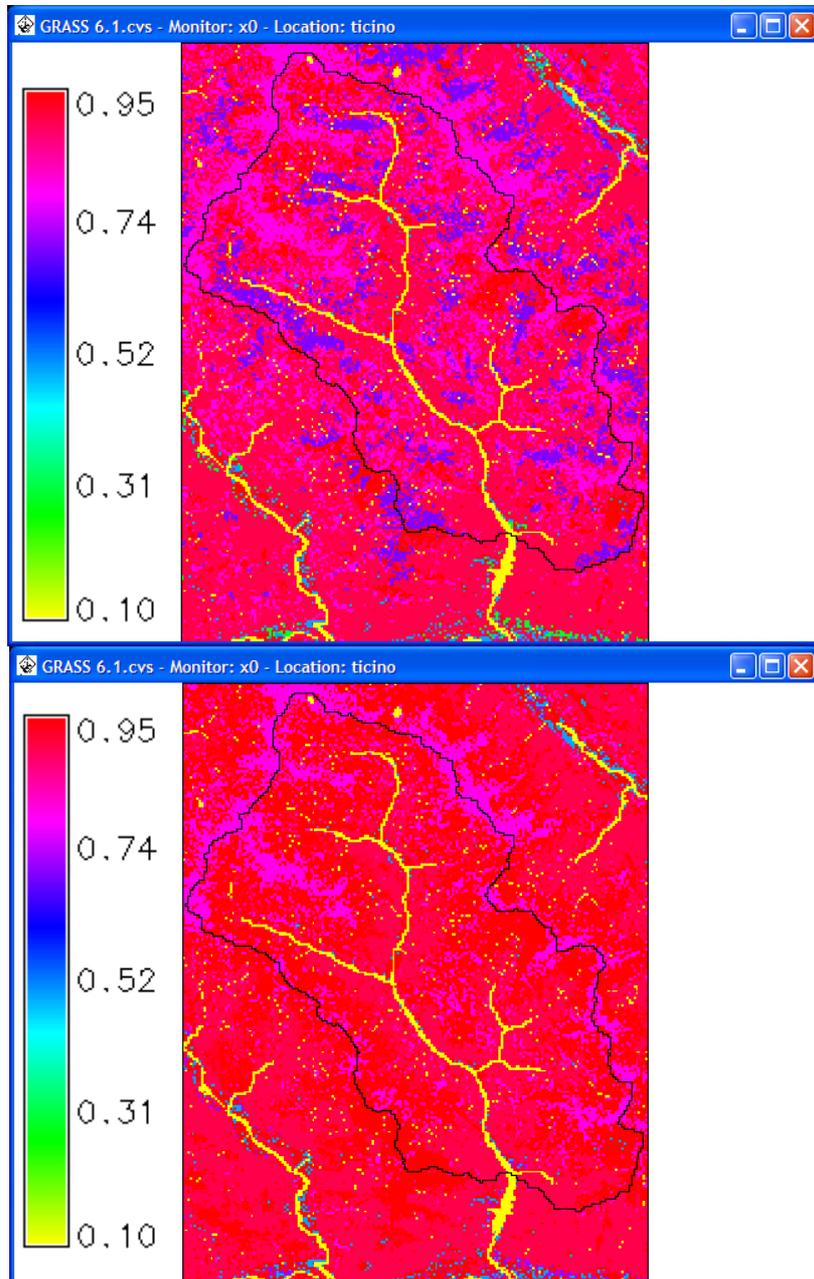


Figure 6.30 – vegetation cover [%] on the 15<sup>th</sup> January (top) and 15<sup>th</sup> August (bottom).

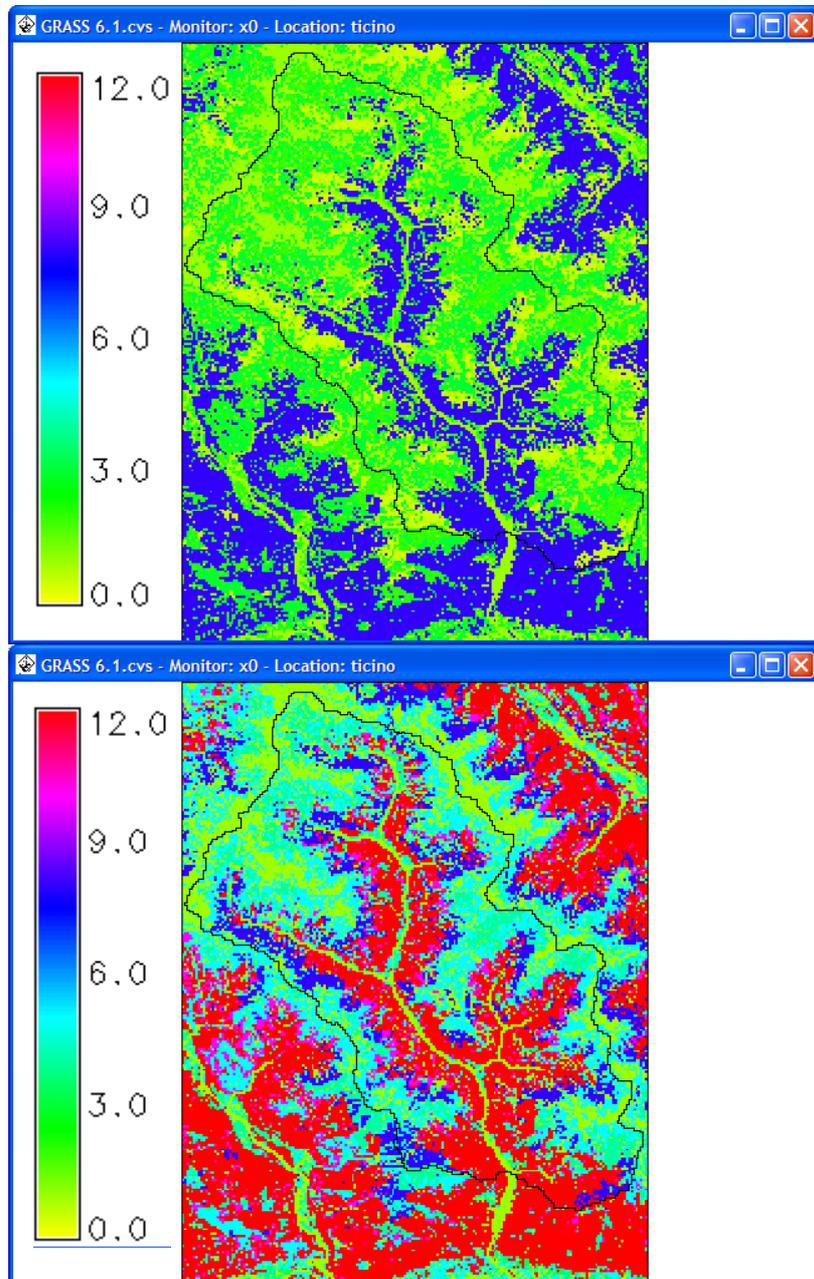


Figure 6.31 – LAI index [-] on the 15<sup>th</sup> January (top) and 15<sup>th</sup> August (bottom).

### **6.11 Net radiation**

The net energy entering in the catchment system has been hourly calculated following the procedures illustrated in figure 4.2.

The correct spatial distribution of the net radiation maps can be detected through a comparison with the aspect map. In fact, the side exposure of the mountains, with respect to solar position, is the driving factor in the net radiation values establishment.

As an example, figure 6.32 shows the aspect map of the Verzasca region and the net radiation map calculated on the 18th August 2004 at 10:00 am. At that hour of the morning the sun is positioned at east and its radiations directly hit the side of the mountain exposed to east. From the comparison of the two maps of Figure 6.32 the correct reproduction of this situation can be detected.

The temporal evolution of the radiation is shown in the figure 6.33, where each frame represent the estimated net radiation ad different hours. The presence of clouds on the weast mountains of the basin can be observed from 11 a.m. to 4 p.m..

Validity of the estimations has been evaluated by comparing extracted mean values with other study previously carried on in this region (Menzel, 1997).

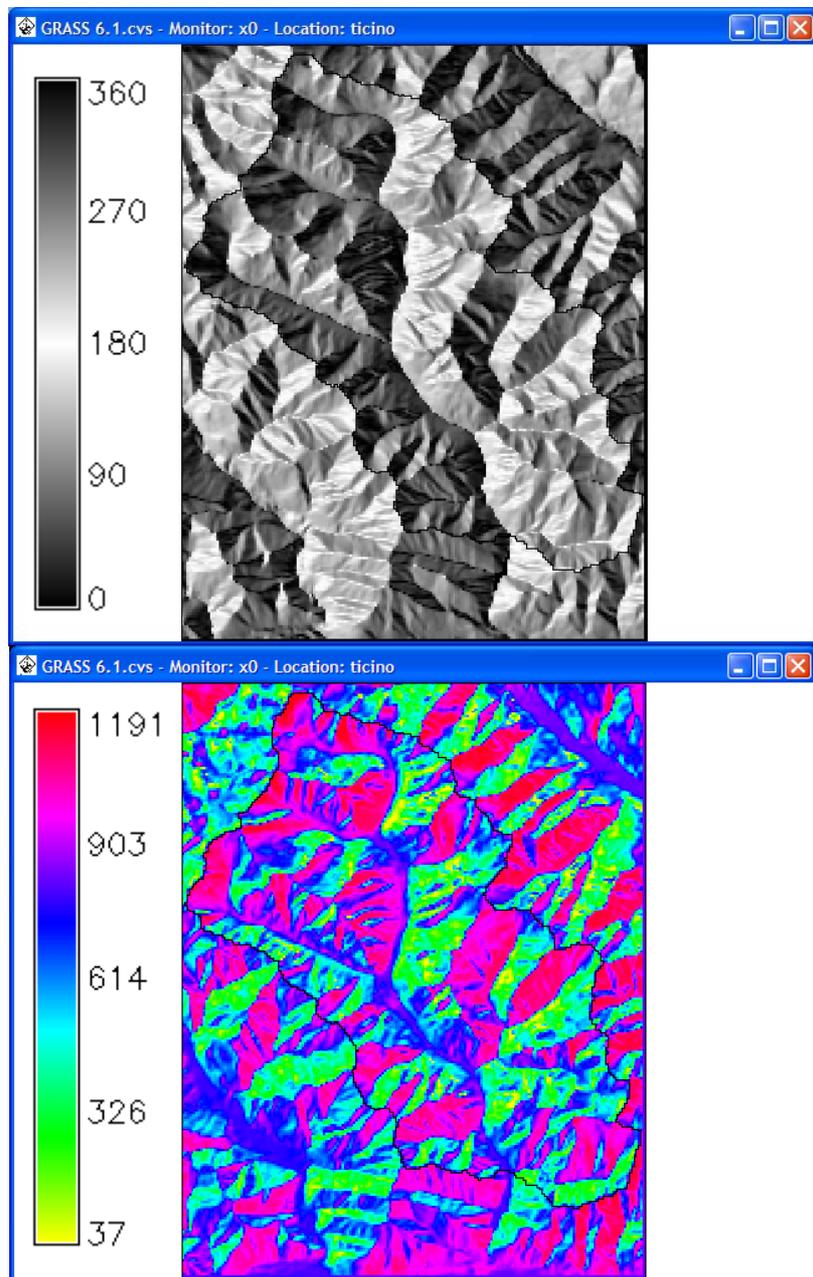


Figure 6.32 – map of the aspects [degree] (top) and net radiation [MJ]/(m<sup>2</sup>·h) for the 18<sup>th</sup> August 2004 at 10:00 am (bottom).

Chapter 6 – Case study: the Verzasca basin

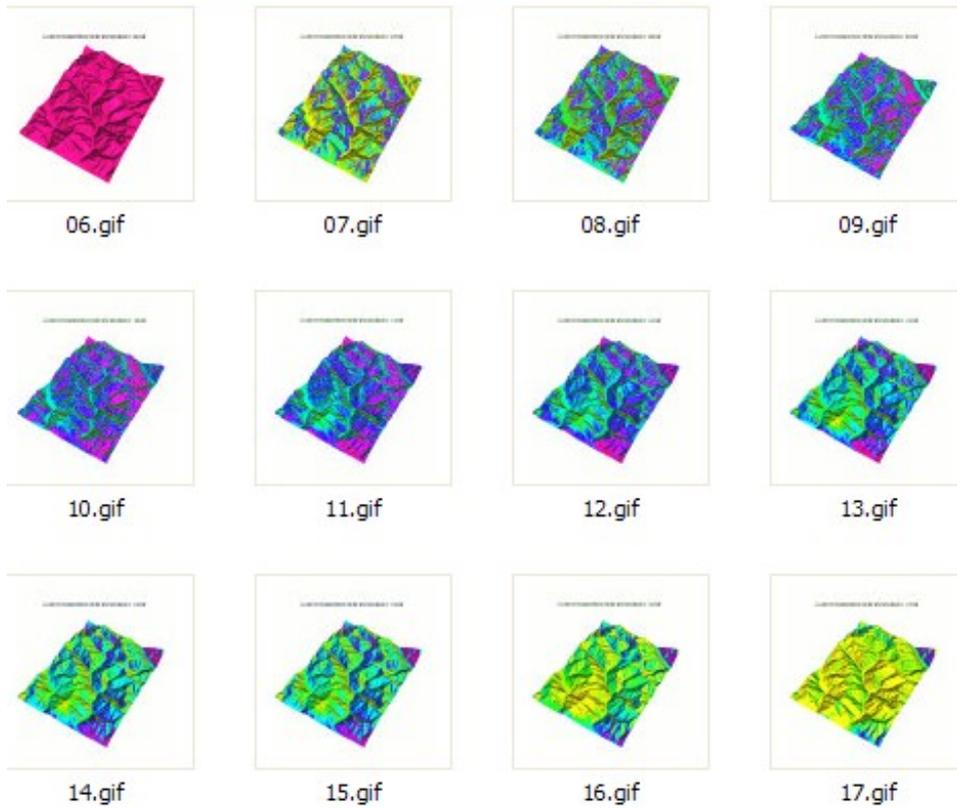


Figure 6.33 – evolution of the net radiation series from 6 a.m. to 5 p.m. On the 5<sup>th</sup> October 2004, (legend like in figure 6.32).

### **6.12 Snowmelt**

According to the procedures described in section 4.3, the snow raster series (snowpack water equivalent, energy content, and snowmelt) has been generated.

The figures 6.34-37 show some generated maps of the snow accumulation and melting process (figure 6.34 shows the snowpack water equivalent, figure 6.35 the snowpack energy, and figure 6.36 the snowmelt components).

From a spatial point of view the generated maps are correct, in fact the snow is accumulated on the top of the mountains, with a snowpack water equivalent decreasing with elevation, and the snowmelt occurs on the limits of the snow covered areas, where temperatures are higher.

Also the quantitative simulated values show a correct behaviour.

As an example, in figure 6.37 a growing snowpack water equivalent in correspondence of rain events, and a snowpack energy content growing following the day-night cycle, can be detected. Moreover, when the energy content reaches the needed value, the snowmelt occurs following day-night cycle, and reducing the snowpack water equivalent and energy content.

Comparing the simulated values with the estimates carried out in the Alpine region by Menzel (1997) a good agreement can be found.

The simulations for August 2004 show the following mean values:

- snowpack water equivalent: 281.34 mm
- snowmelt: 0.26 mm
- snowpack energy content: 31514.79 Kcal\*m<sup>-2</sup>

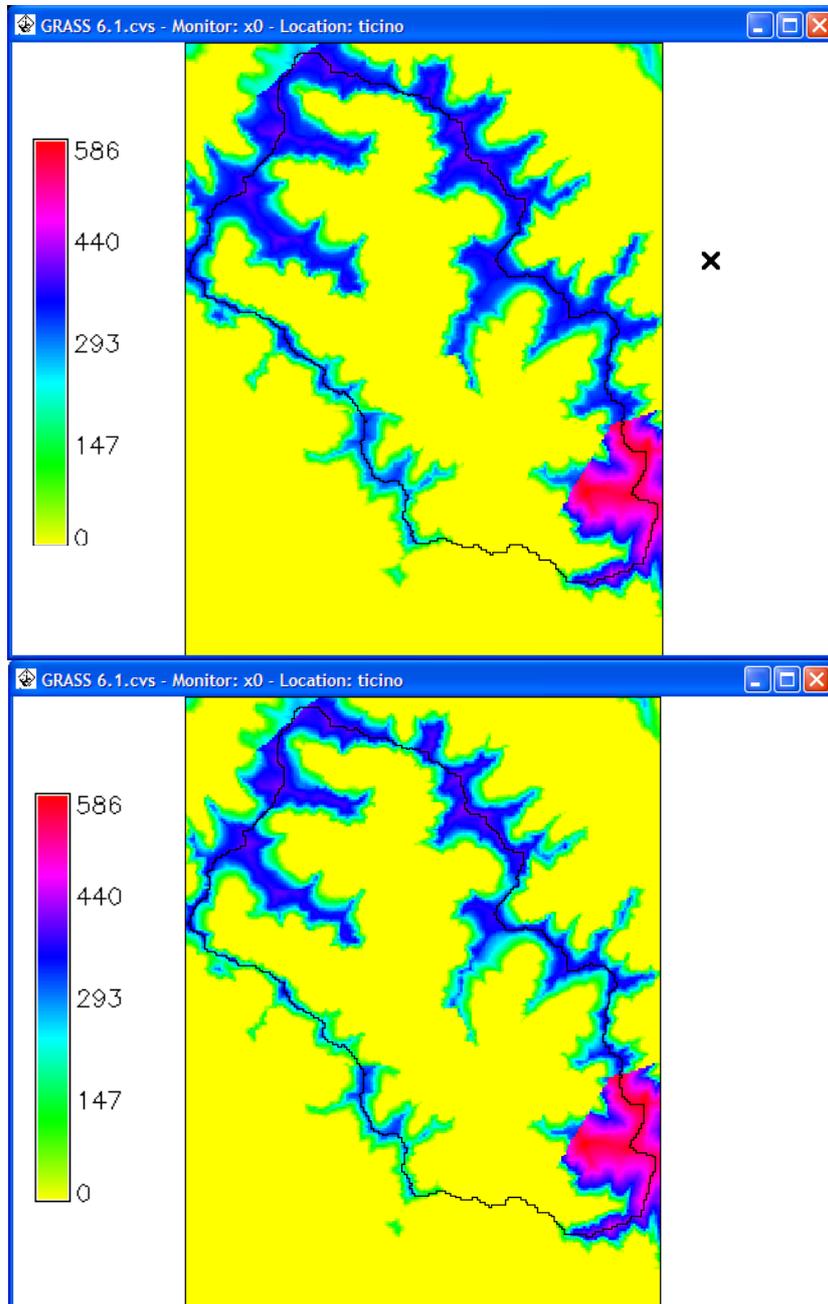


Figure 6.34 – snowpack water equivalent [mm] evolution (above: 17<sup>th</sup> August 2004 at 00:00 am.; bottom: 18<sup>th</sup> August 2004 at 10:00 am.)

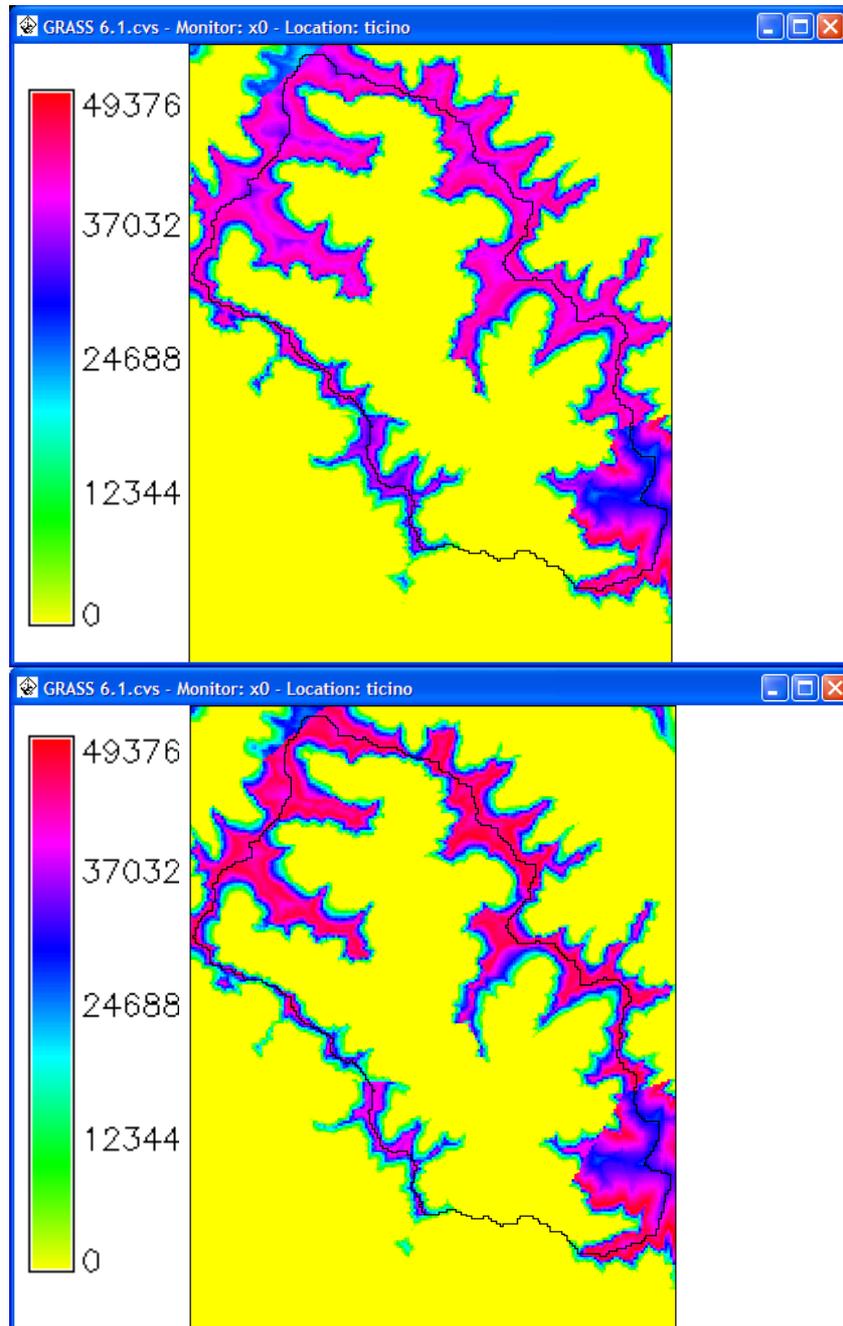


Figure 6.35 – snowpack energy [ $\text{Kcal}\cdot\text{m}^{-2}$ ] evolution. Above: 17<sup>th</sup> August 2004 at 00:00 am. (night); Bottom: 18<sup>th</sup> August 2004 at 10:00 am. (day).

Chapter 6 – Case study: the Verzasca basin

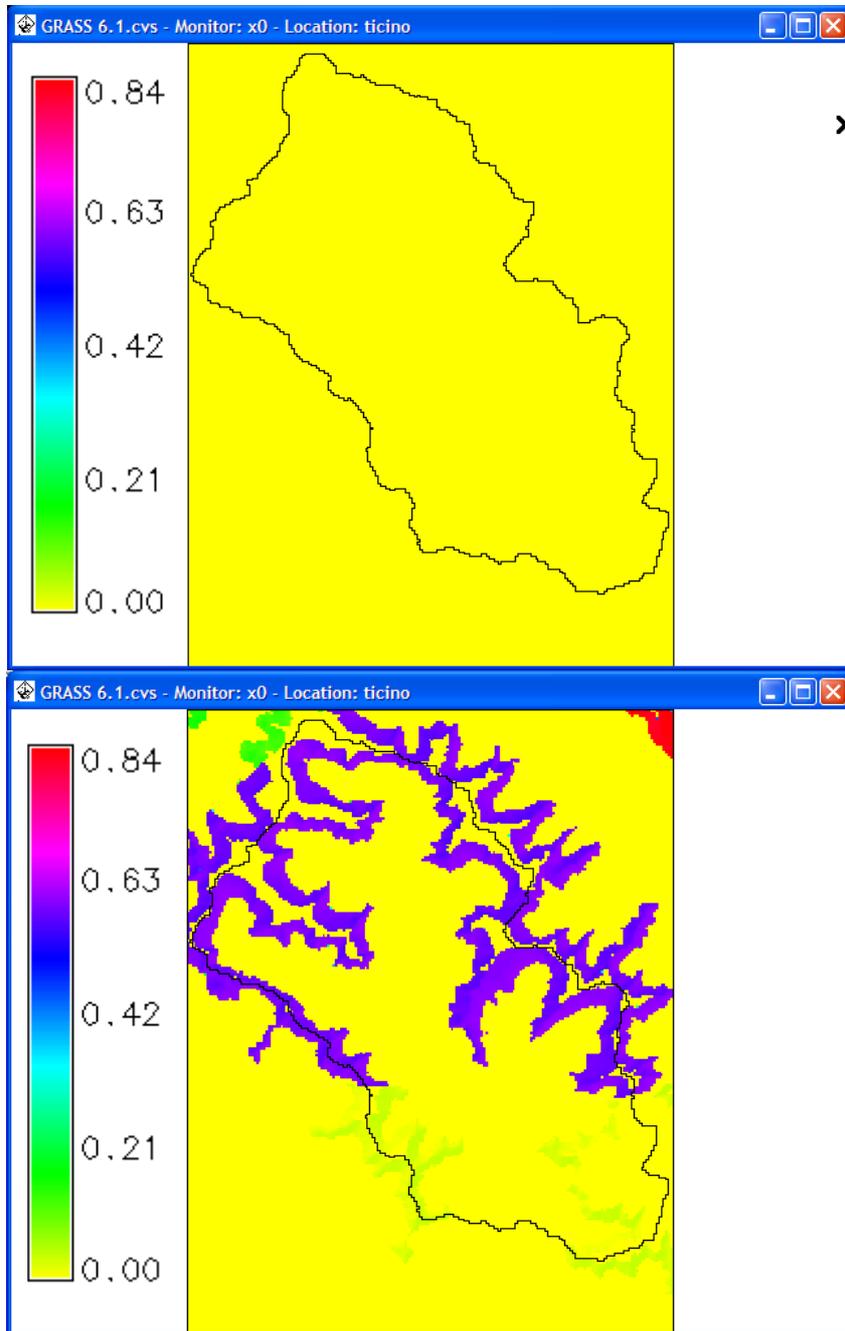


Figure 6.36 – snowmelt [mm] evolution (above: 17<sup>th</sup> August 2004 at 00:00 am.; bottom: 18<sup>th</sup> August 2004 at 10:00 am.)

## Chapter 6 – Case study: the Verzasca basin

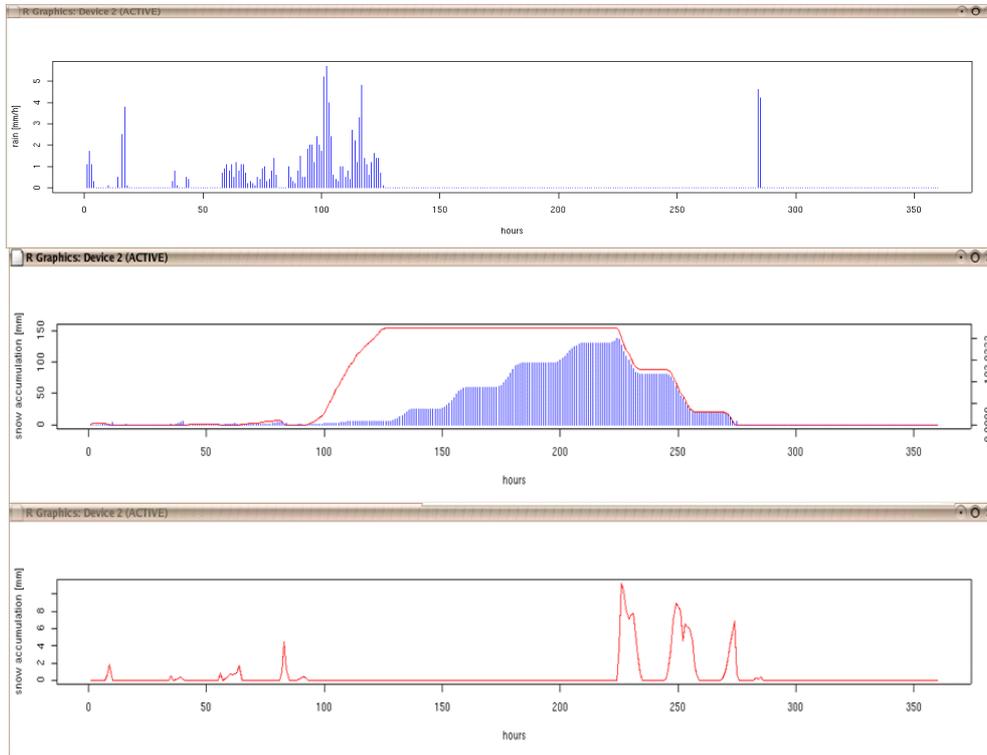


Figure 6.37 – from top to bottom: (1) rainfall (mm), (2) snowpack water equivalent (mm) in red and snowpack energy content in blue, and (3) snowmelt (mm). All plots extracted for May 2004 from a sampling point.

### 6.13 Evapotranspiration

The potential evapotranspiration maps, computed as described in section 4.4, have been extracted and analysed in order to verify their correctness.

No missing maps have been detected and the spatial behaviour agree with consideration of landuse, exposition, and quota of the terrain as shown in figure 6.38. The potential evaporation from a liquid surface is in fact higher then the one from the terrain, moreover the areas where rock is present the evapotranspiration result lower then where the vegetation is present (landuse map in figure 6.13).

Another important consideration, derived from the comparison of the two maps of figure 6.38, is that the clear sky index greatly influence the evapotranspiration values. In fact, when clouds are present, the energy reaching the ground is lower and therefore less energy is available for the water evaporation and for the vegetation transpiration.

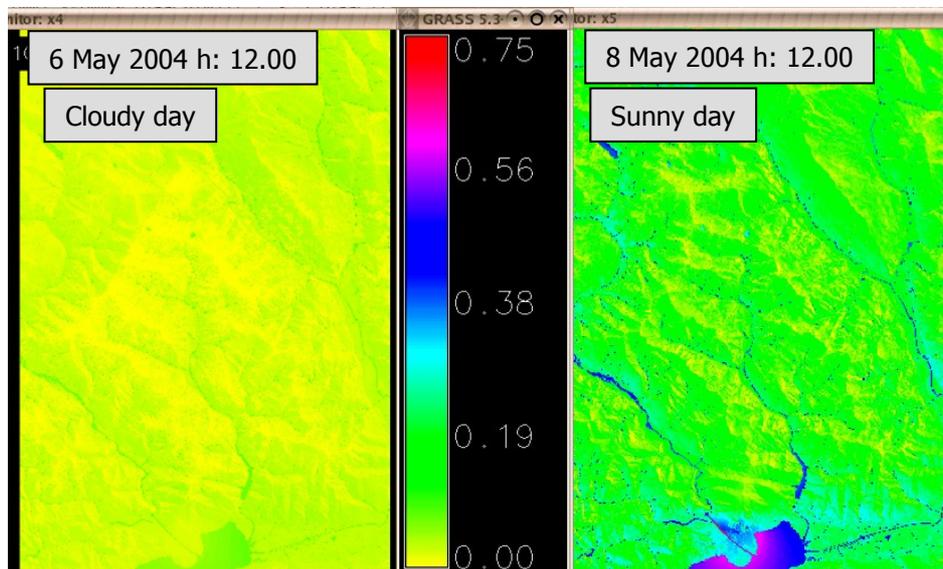


Figure 6.38 – potential evapotranspiration maps for two different days of May (6<sup>th</sup> and 8<sup>th</sup>) at the same hour (midday) in two different sky condition (cloudy and sunny).

## Chapter 6 – Case study: the Verzasca basin

The simulations sampled in same check points, homogeneously positioned in the region and different each other for quota, exposure and landuse, reveal the correctness of the values from a qualitative and quantitative point of view.

Looking at figure 6.39, where as an example the temporal evolution of the evapotranspiration at one of the check points is plotted, can be observed that the evaporation from a liquid surface is always higher than from terrain and that the day-night cycle is well reproduced.

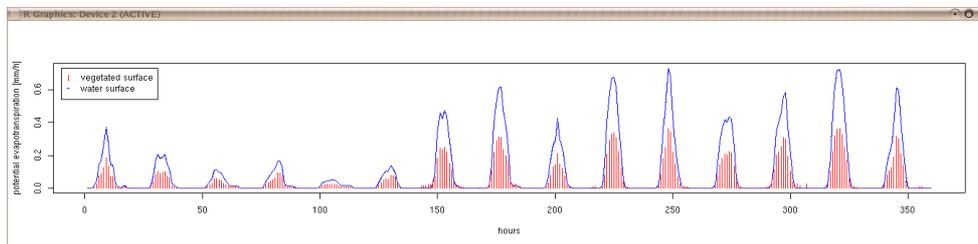


Figure 6.39 – Potential evapotranspiration (red vertical lines) and potential evaporation (blue lines) during May 2004.

From a quantitative point of view:

- the potential evapotranspiration in May 2004 shows a mean value of 0.08 mm/h, a maximum of 0.92 mm/h and a minimum of 0.00 mm/h
- the potential evaporation from a liquid surface in the same month shows a mean value of 0.14 mm/h, a maximum of 1.38 mm/h, and a minimum of 0.00 mm/h.

All these values are in good agreement with other previous estimates carried on in the study region (Menzel, 1997).

### **6.14 Canopy interception**

The canopy module running produces three raster series: the canopy storage level, the interception drainage and the interception loss.

Figures 6.40, 6.41 and 6.42 show the three components at 24 hour distance (17th Aug 2004 at 12:00 and 18th Aug 2004 at 12:00).

The comparison of these figures with the figure 6.43, showing the rain distribution generated for those times, allow us to detect the correct spatial distribution of the phenomena. In fact the canopy storage level is distributed where vegetation is present and the higher values are localized where the rain is more intense (on the 17<sup>th</sup> August it is the upper part of the map).

Also the drainage from the canopy is coherent with the rain distribution and shows non zero values where and when the canopy is saturated (red values on figure 4.41).

To analyse the the correct behaviour of the simulated process the generated maps has been samplet at different check points distributed over the region. The goodness of the simulation has been detected by plotting the temporal evolution of the sampled values. As an example, in figure 6.44 the plot derived at one of the sample point is represented. The correct behaviour can be observed by the fact that:

- the canopy storage level grows with rainfall until the maximum capacity is reached, and then decrease proportionally to the evapotranspiration;
- the canopy drainage occurs only where rain falls and the canopy is saturated;
- the interception losses are function of the evapotranspiration and of the canopy water storage content;
- the interception losses, caused from evaporation of the intercepted water, occurs wherever the canopy storage level is not empty.

The quantitative correctness of the simulations has been verified by means of the estimations statistical values reported in table 6.8.

## Chapter 6 – Case study: the Verzasca basin

<b>Map series</b>	<b>mean</b>	<b>maximum</b>	<b>minimum</b>
interception loss	0.02	0.4	0
interception drainage	0.38	8.99	0
canopy storage level	0.67	3.34	0

Table 6.8 – statistics in mm/h of the series generated with the h.interception command.

These values are in agreement with previous estimated values in the Ticino region (see Menzel, 1997).

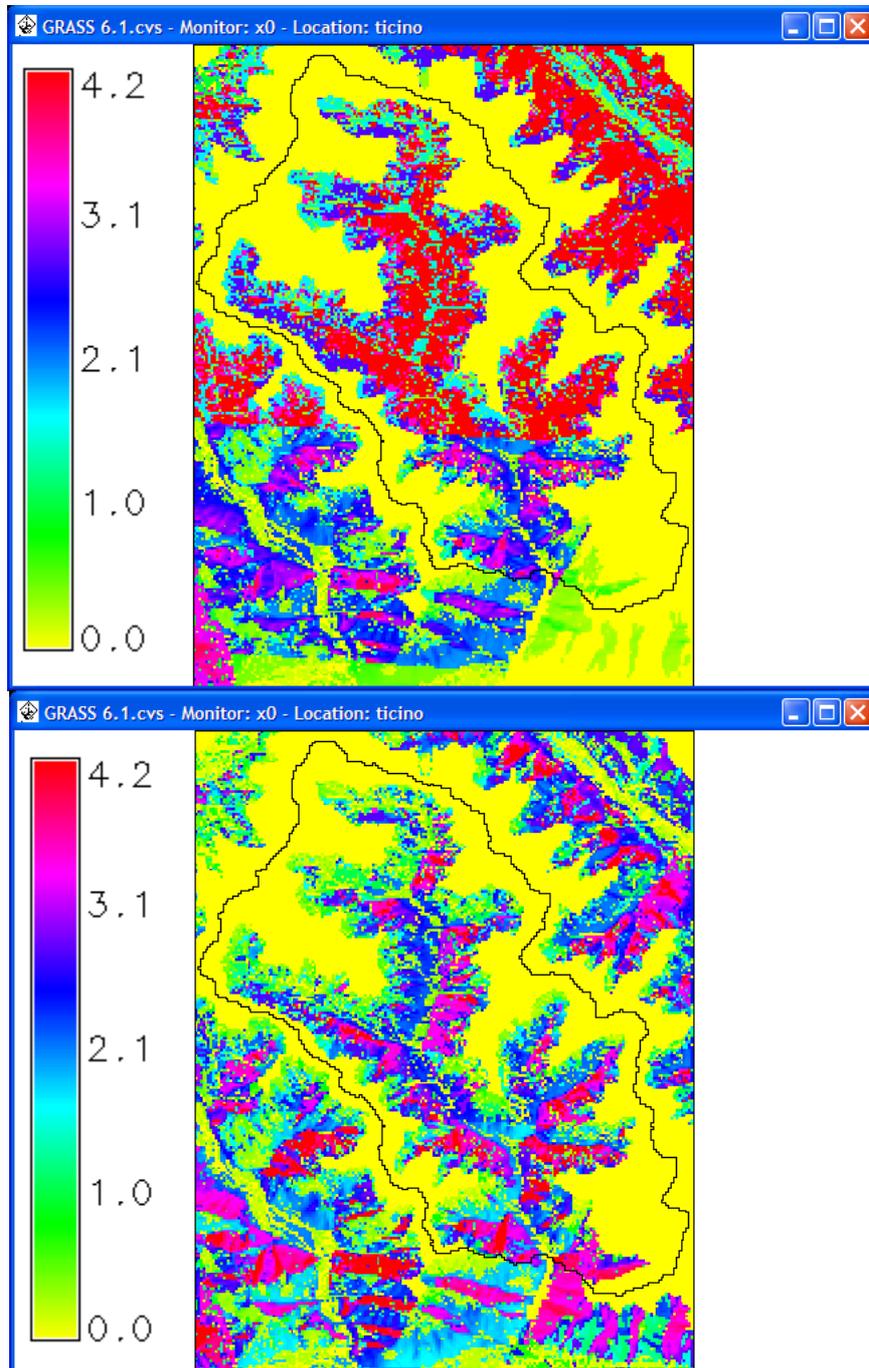


Figure 6.40 – canopy storage level (mm) at 24 hour distance (above 17<sup>th</sup> Aug 2004 at 12:00, bottom 18<sup>th</sup> Aug 2004 at 12:00)

Chapter 6 – Case study: the Verzasca basin

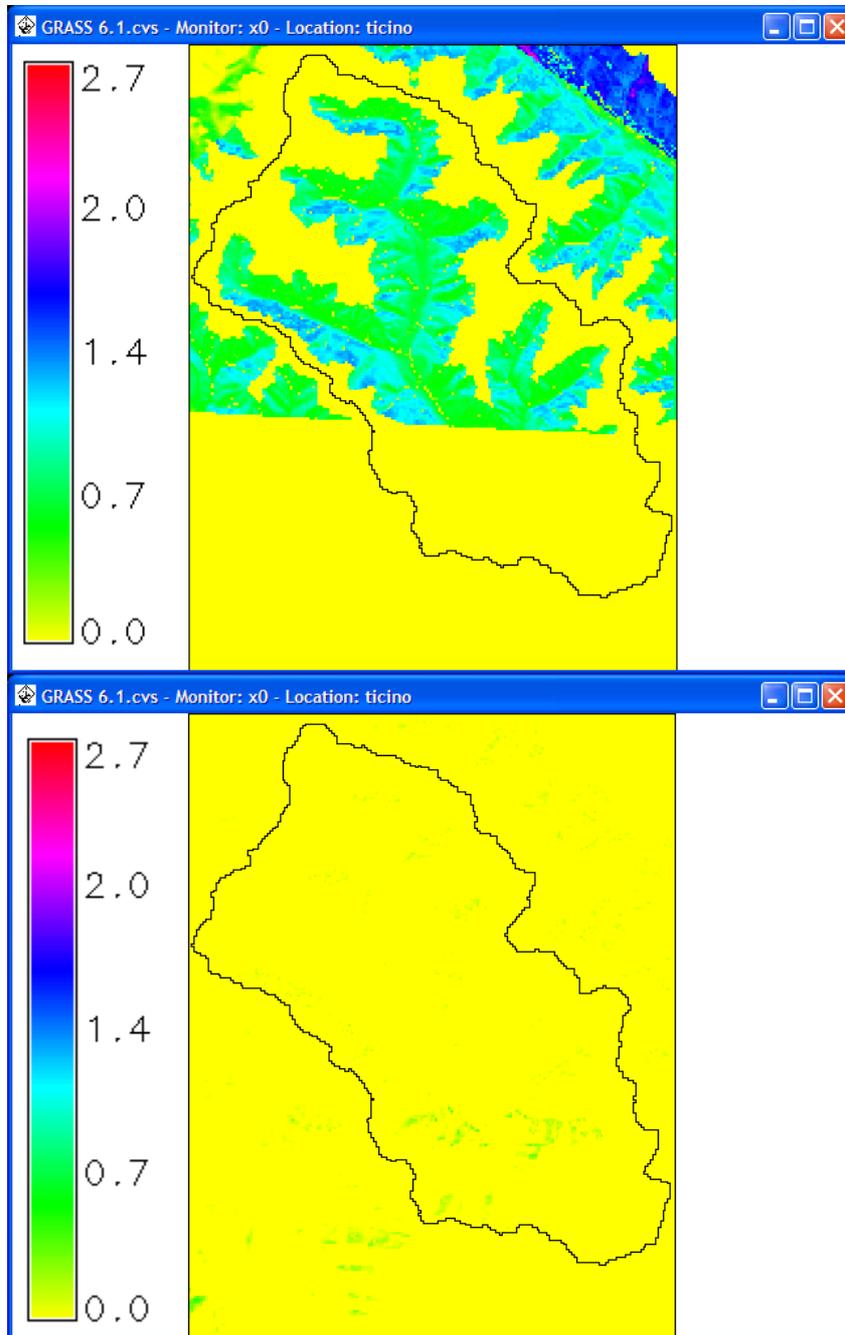


Figure 6.41 – interception drainage (mm/h) at 24 hour distance (above 17<sup>th</sup> Aug 2004 at 12:00, bottom 18<sup>th</sup> Aug 2004 at 12:00)

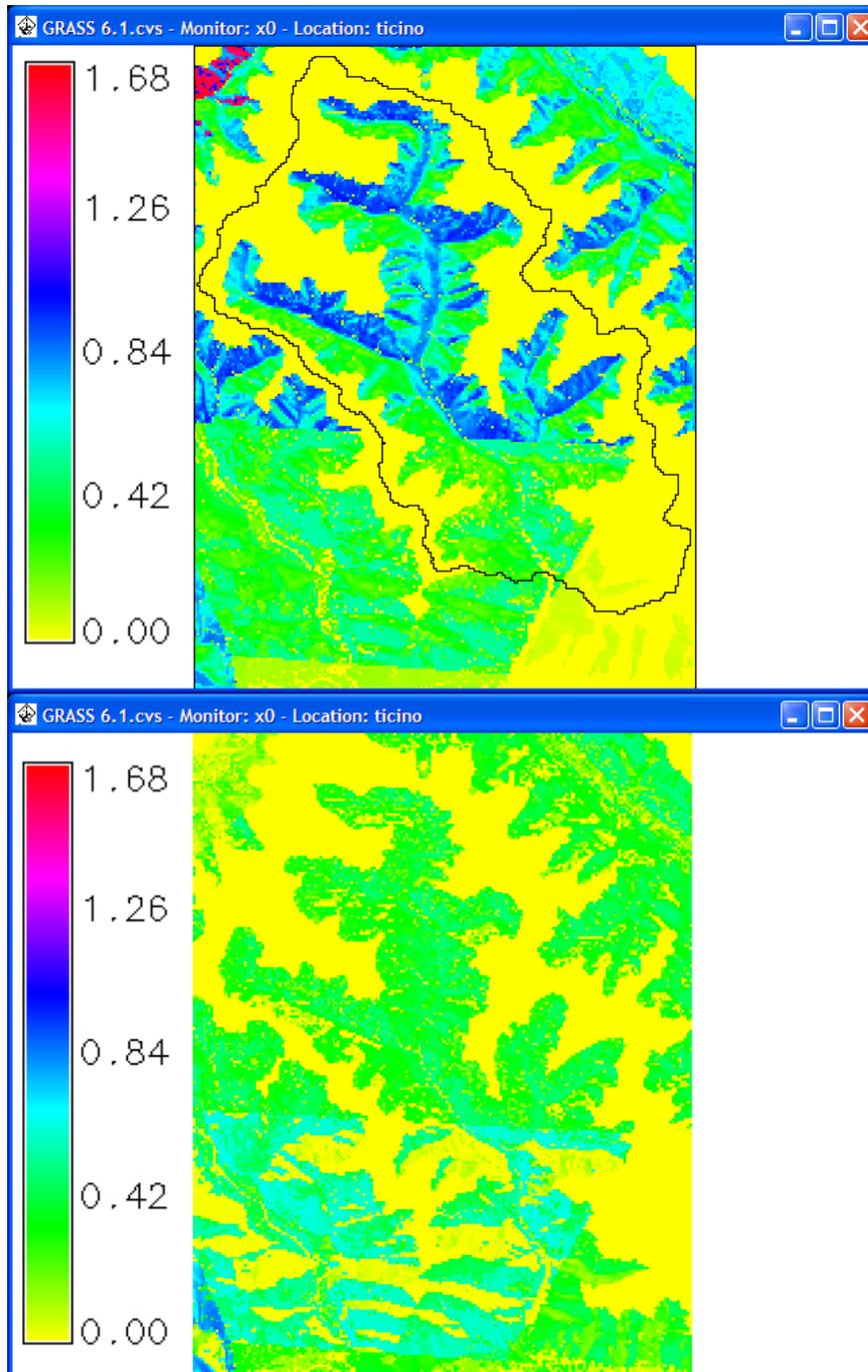


Figure 6.42 – interception losses (mm/h) at 24 hour distance (above 17<sup>th</sup> Aug 2004 at 12:00, bottom 18<sup>th</sup> Aug 2004 at 12:00)

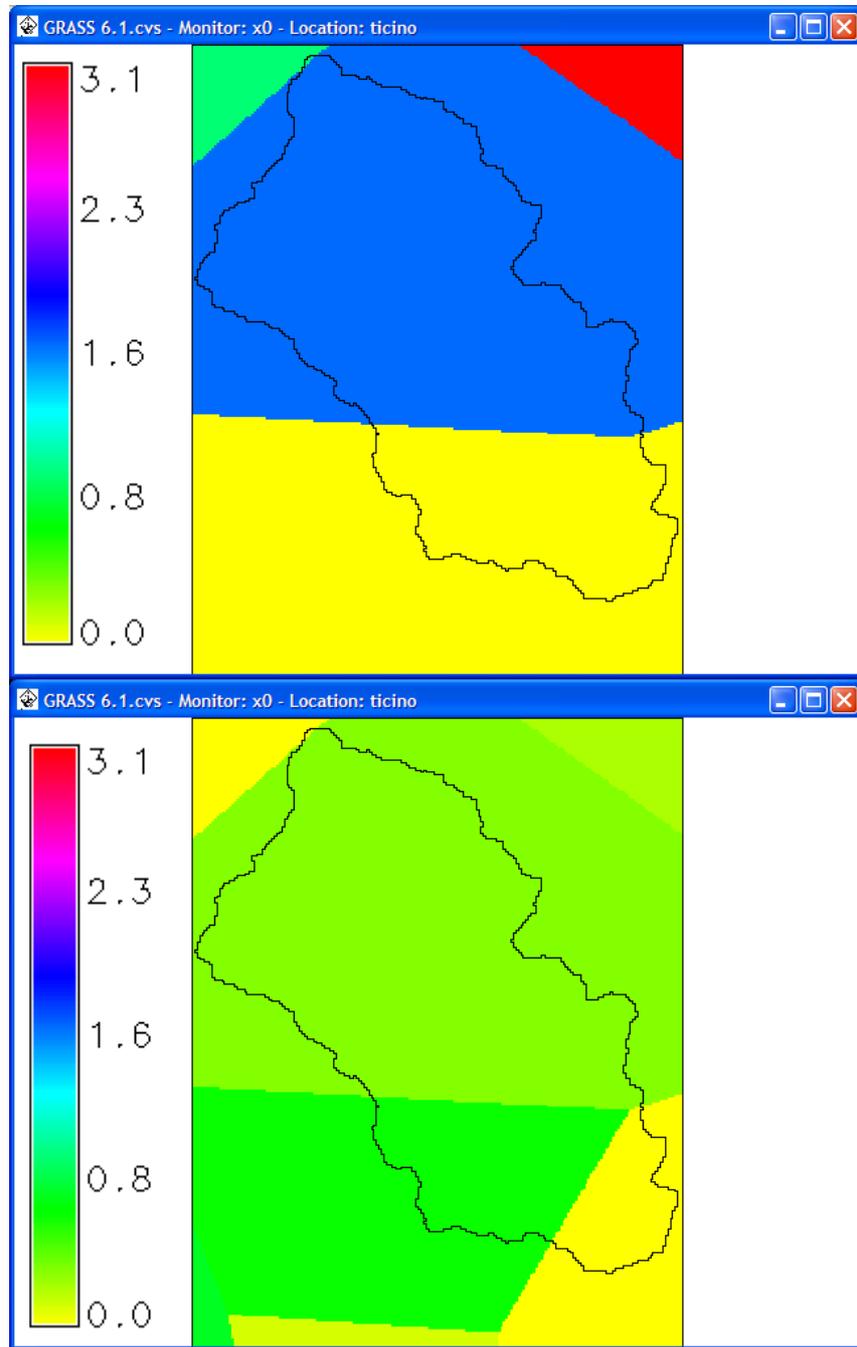


Figure 6.43 – rainfall intensity (mm/h) at 24 hour distance (above 17<sup>th</sup> Aug 2004 at 12:00, bottom 18<sup>th</sup> Aug 2004 at 12:00)

## Chapter 6 – Case study: the Verzasca basin

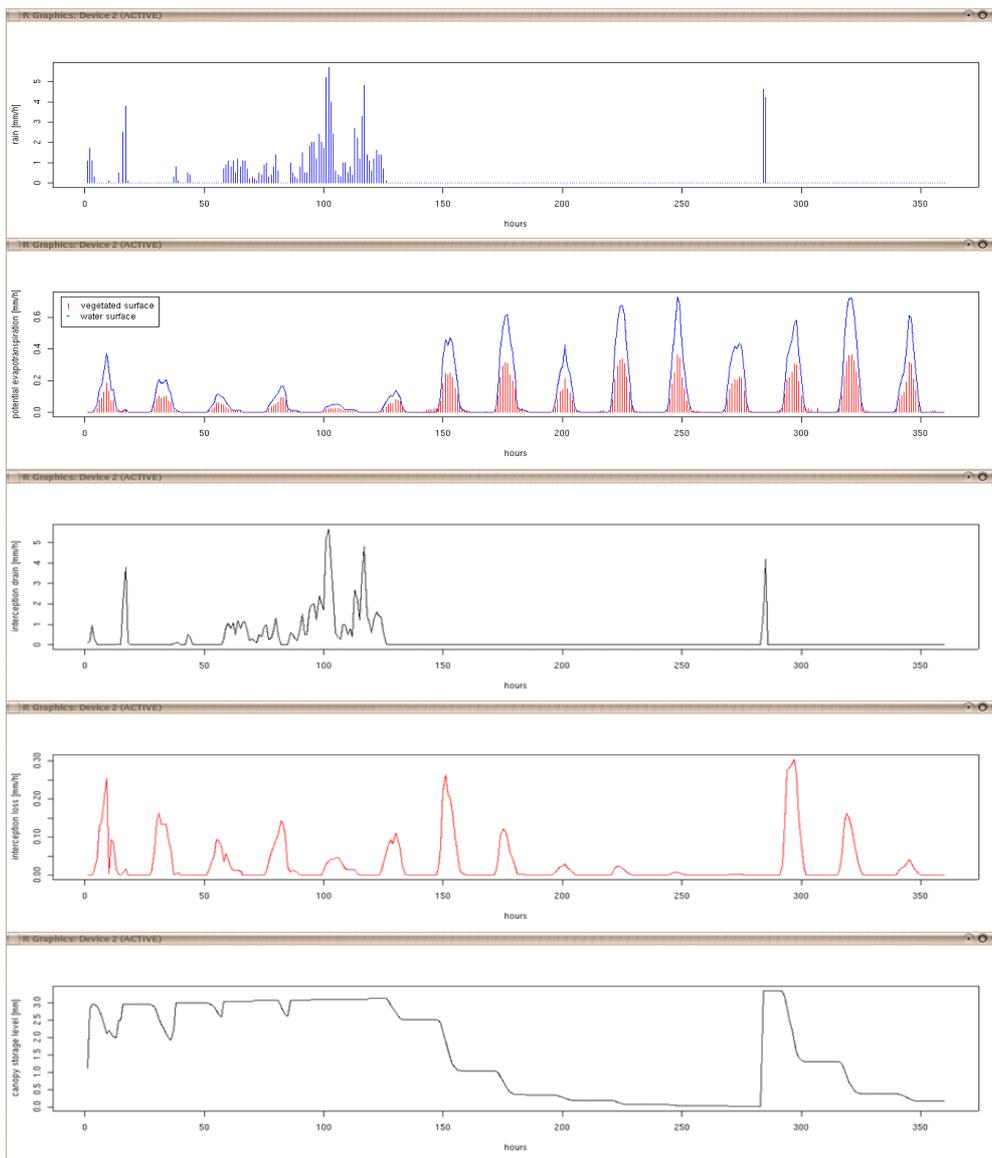


Figure 6.44 – from top to bottom: rainfall (mm/h), evapotranspiration (mm/h), interception drainage (mm/h), interception losses (mm/h), canopy storage level (mm).

### **6.15 Runoff calibration**

As already mentioned in section 4.8, in physically based distributed hydrological models, various parameters must be calibrated. This can involve tremendous effort, especially when the number of parameters is large and the applied hydrological model is computationally expensive. Automatic parameter estimation tools can significantly facilitate the calibration process, hence the inverse automatic calibration model UCODE-2005 (Poeter et al., 2005) has been combined with the developed distributed hydrological model HydroFOSS.

Using the developed UCODE-HydroFOSS linkage, a first sensitivity analysis was performed in order to extract information and to identify the most important parameters to be calibrated for reproducing a good fitting between observations and simulations.

Usually, in order to run effective calibration a period that, depending on the basin characteristics, can vary from 3 up to 12 months is used (Montanari, 2006). This long period allows to include all different types of events that can occur (e.g. short summer shower, and long winter precipitation), and to let the soil moisture reach the real condition.

This operation can easily ask for a couple of working months because:

- the high number of processes and calculus required for each cell of the grid, the simulations result expensive in terms of time and computational power;
- the calibration runs one complete running for each parameter perturbation;
- multiple calibration process (with different adjustable parameters with different starting values) has to be run to reach the desired solution.

With the aim to validate the HydroFOSS model, and considering the few time available to conclude the PhD, a first calibration has been performed on the only month of August 2004.

Obviously this calibration can only reveal the general correctness

of the model and the parameter values derived from this calibration could not be valid for a good fitness in other periods.

For kindly concession of the CSCS (Swiss National Supercomputing Center, [www.cscs.ch](http://www.cscs.ch)) of Lugano, the implemented commands, the GRASS GIS, and the UCODE-2005 program has been installed on a HP computer with 4CPU Itanium/2 (IA-64) under a Linux (Suse distribution) environment. Such a system allows the running of one month simulation in approximately one and half hour.

Considering the short period considered, with the aim to reduce the warm-up period where soil the water content is not correctly simulated, a new parameter has been introduced: the starting soil saturation percentage (sat).

### *6.15.1 First sensitivity analysis*

Considering the soiltypes and the landuses present in the Verzasca basin, the significance 36 parameters to the discharges simulation has been analysed:

- the Manning coefficient for 11 different landuses identified by the symbols  $n_X$ , where X represent the soiltype category index ( $n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8, n_9, n_{10}, n_{11}$ );
- the soil conductivity for 6 different soiltypes identified by the symbol  $ks_X$ , where X is the soiltype category index ( $ks_2, ks_3, ks_4, ks_7, ks_8, ks_9$ );
- the saturated soil water content for 6 different soiltypes identified by the symbol  $ts_X$ , where X is the soiltype category index ( $ts_2, ts_3, ts_4, ts_7, ts_8, ts_9$ );
- the residual soil water content for 6 different soiltypes identified by the symbol  $tr_X$ , where X is the soiltype category index ( $tr_2, tr_3, tr_4, tr_7, tr_8, tr_9$ );
- the soil thickness for 6 different soiltypes identified by the symbol  $L_X$ , where X is the soiltype category index ( $L_2, L_3, L_4, L_7, L_8, L_9$ );
- The starting soil saturation percentage SAT.

Parameters were chosen that likely would have the greatest effect

on channel discharges. From a physical point of view, those parameters directly influence the water flows into the subsurface, overland and streams.

Parameters directly responsible for the amount water available for the runoff driving the interception drainage, the evapotranspiration and the snowmelt were not included in the recharge-parameter analysis. The amount of water available for the runoff was understood to have the greatest effect on the channel flow; therefore, those parameter sensitivities could overwhelm other sensitivities and mask subtler effects of the runoff flow processes.

Scaled sensitivities were calculated by perturbing (increasing and decreasing) the “a priori” parameter values of the model and calculating the derivative of the simulated value (discharge), with respect to the parameter. The size of the perturbation is calculated as a user-specified factor. All parameters by 25 percent.

The importance of different parameters to the calculation of a observed value  $i$  and the importance of different observations to the estimation of a single parameter  $b_j$  can be detected by the dimensionless scaled sensitivity ( $DSS_{ij}$ ) index.

The dimensionless scaled sensitivity, equation, is calculated as (Hill, 1998):

$$DSS_{ij} = \left( \frac{\partial y'_i}{\partial b_j} \right) \cdot b_j \cdot \varpi_i^{1/2} \quad [\text{eq. 6.2}]$$

where:

- $i$  identifies one of the observations,
- $j$  identifies one of the parameters,
- $y'_i$  is the simulated value associated with the  $i$ -th observation (in this case it is the hourly measured discharge at the two limnigraph of figure 6.19),
- $b_j$  is the  $j$ -th estimated parameter (in this case they are the 36 analysed parameters),
- $\partial y'_i / \partial b_j$  is the sensitivity of the simulated value associated with the  $i$ -

## Chapter 6 – Case study: the Verzasca basin

th observation and is evaluated at the final parameter values, and

$\omega_i$  is the weight for the i-th observation (in this case it is assigned equally to each observation).

In figures 6.45 and 6.46 the obtained DSS relative to station 1 and station 2 has been respectively plotted.

Keeping in mind that for one parameter the observation in correspondence of higher DSS have higher importance for its estimation, and that for one observation the parameter with higher DSS has higher importance in the simulation of the correspondent simulation, these figure can be analysed and same consideration can be done.

The figures show that in the first half month (from 0<sup>th</sup> to ~330<sup>th</sup> hours) the starting saturation percentage (SAT) is the most important parameter for the best fitting of the simulation, while the Manning coefficient for the river landuse (n\_11) results in being quite important for all over the period.

Other sensitive parameters are the soil conductivity and the soil thickness for the regosol eutric predominant soiltype (ks\_2 and L\_2), which is the most diffuse soiltype in the catchment.

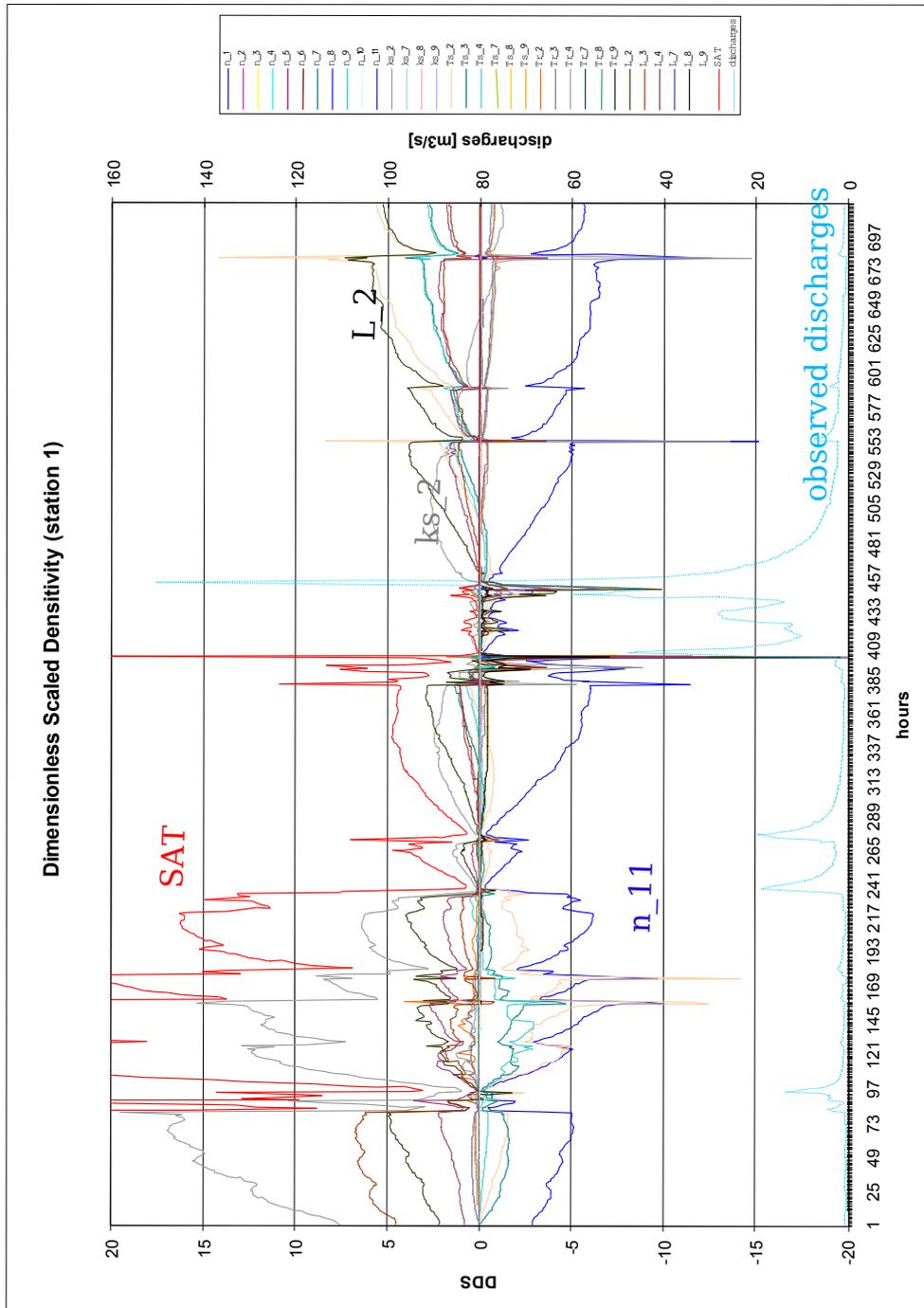


Figure 6.45 – Dimensionless Scaled Sensitivity of the 36 model parameters, and the relative observed river discharge at station 1.

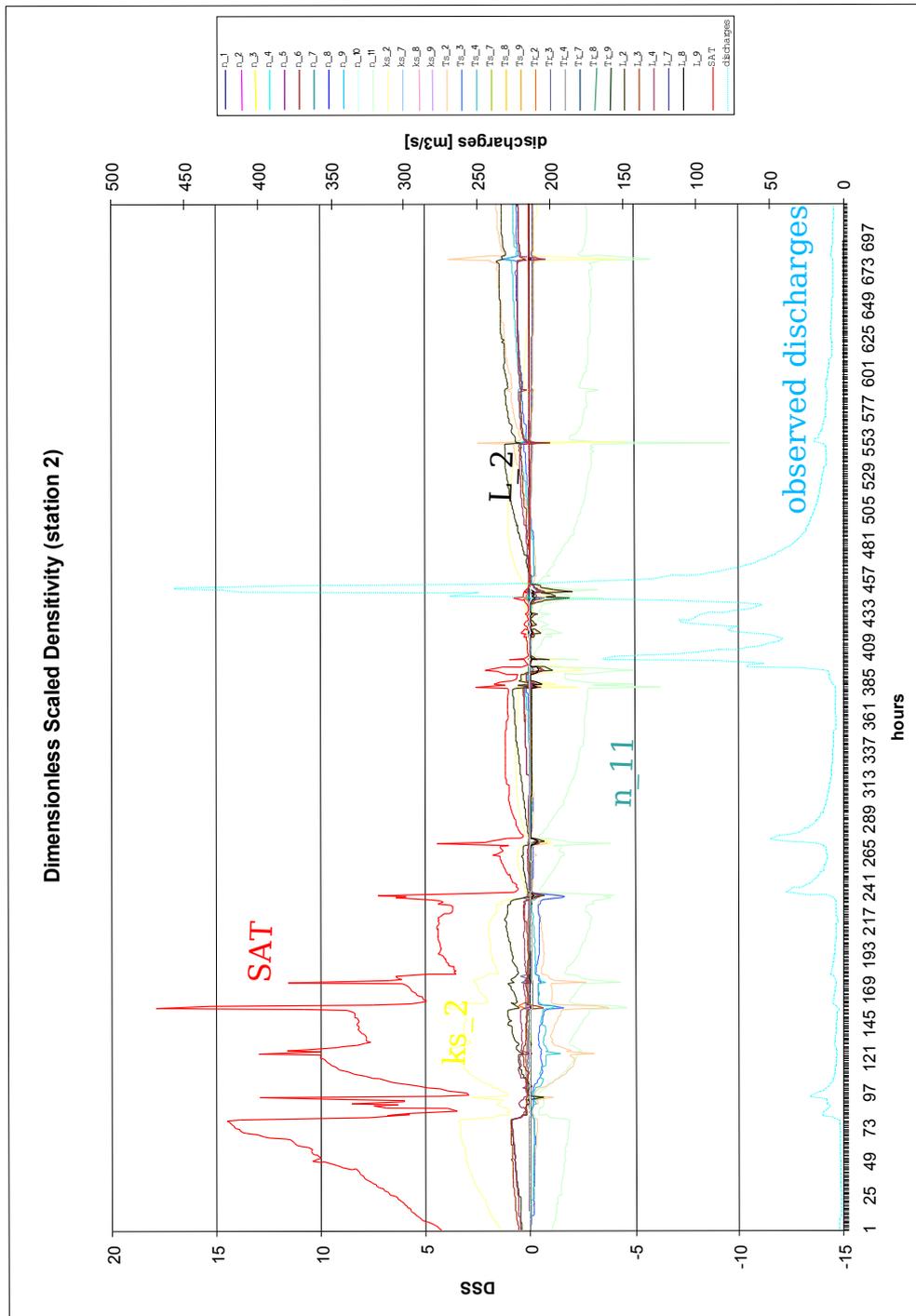


Figure 6.46 – Dimensionless Scaled Sensitivity of the 36 model parameters, and the relative observed river discharge at station 2.

A better view of the parameters influence can be achieved by

plotting the maximum DSS for each observation with information on the parameter type, as shown in figure 6.47: this plot highlight which parameter is the most important for the correctness of the fitting along the time axis.

Four different periods of influence can be detected:

- a first period of warm-up where SAT is driving the fitting (0 – 330);
- two periods of low flow where the n\_11 is driving the fitting (330-430 and 480-720);
- one period of rising flow where L\_2 is driving the fitting (430 – 450);
- one period of recession flow where ks\_2 is driving the fitting (450 – 480).

These influence periods can be explained as follows:

- the model need of a first period to well represent the real soil saturation level, in this phase the starting saturation level strongly affect it;
- the rising limb of the hydrograph is mainly due to the overland flow (fast flow component) that occur only when the soil is saturated: the soil thickness influences the soil water capacity and therefore regulates the water excess (amount of water available for overland flows);
- the recession limb of the hydrograph is mainly due to the subsurface flow (slow flow component) whose velocity, and therefore the amount of water that reach the streams at each hour, is driven from the soil conductivity parameter;
- in low flow condition the amount of water in the river and its velocity is driven from the roughness coefficient of Manning.

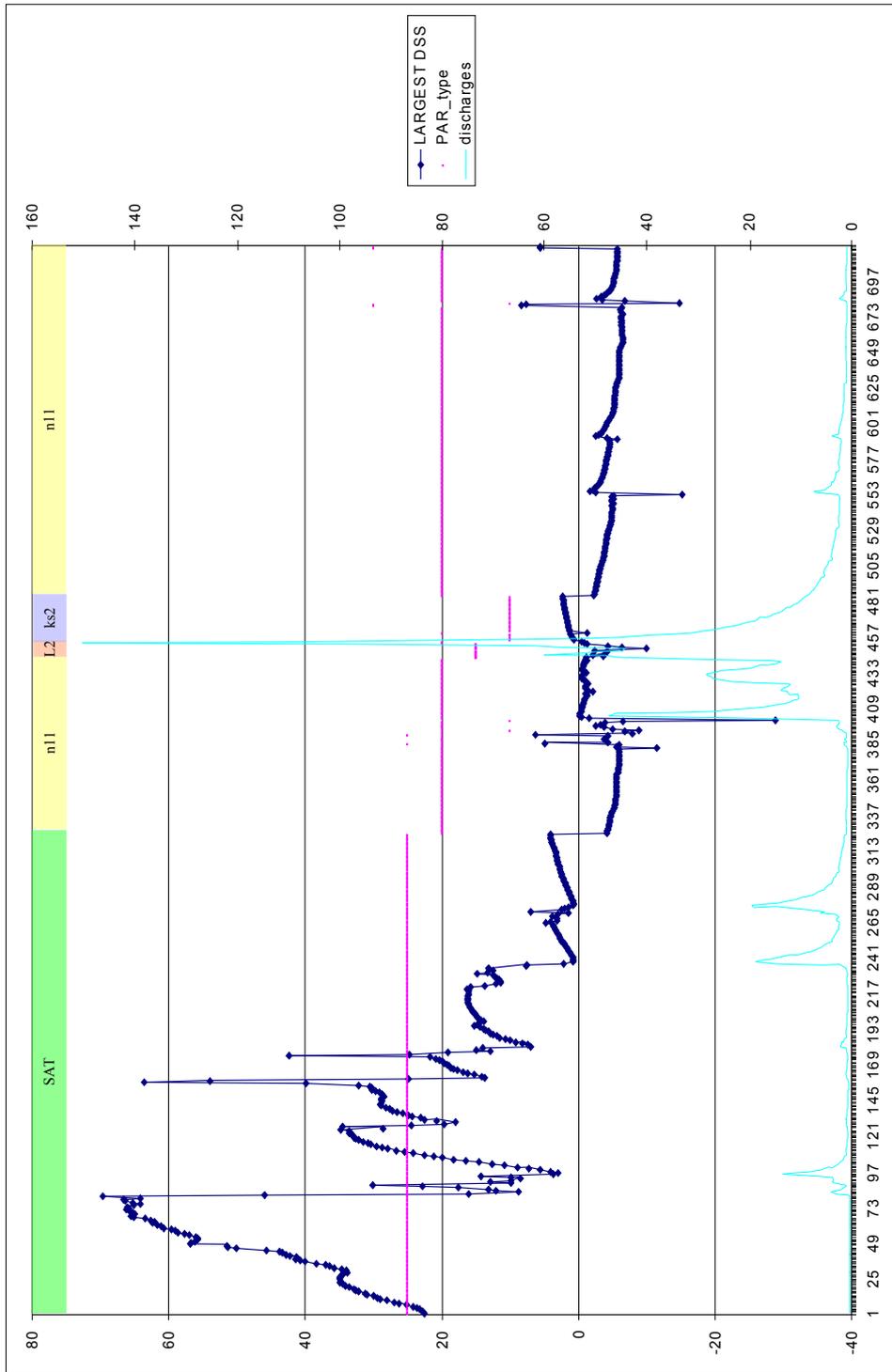


Figure 6.47 – largest calculated DSS for station 1

## Chapter 6 – Case study: the Verzasca basin

The total amount of information provided by the observations for the estimation of one parameter  $b_j$  can be calculated by means of the Composite Scaled Sensitivity (CSS) index.

CSS summarize all the sensitivities for one parameter. CSS are calculated for each parameter using the dimensionless scaled sensitivities for all observations. Because they are dimensionless, CSS can be used to compare the amount of information provided by different types of parameters. Model simulation results will be more sensitive to parameters with large CSS relative to those for other parameters. The CSS for the  $j$ -th parameter,  $CSS_{ij}$ , is calculated as (Hill, 1998):

$$CSS_j = \sum_{i=1}^{ND} \left[ \left[ (DSS_{ij})_{\mathbf{b}_i}^2 / ND \right] \right]^{1/2} \quad [\text{eq. 6.3}]$$

where:

- ND is the number of observations (discharges measured at station 1 and 2) being used in the regression fitting,
- $\mathbf{b}_i$  is a vector that contains the parameter values at which the sensitivities are evaluated (the 36 parameters), and
- DSS<sub>ij</sub> equals the scaled sensitivities of equation 6.2.

Figure 6.48 showing a plot of the calculated CSS for the different parameters, highlights, the overall importance of the Manning coefficient of the river ( $n_{11}$ ), of the conductivity of the regosol eutric predominant soiltype ( $ks_2$ ), of the soil thickness and of the saturated soil water content of the most diffuse soiltypes ( $L_2$ ,  $L_3$ ,  $L_4$ ,  $ts_2$ ,  $ts_3$ ,  $ts_4$ ). Of course the starting saturation percentage of the soil results in being the most important due to the already mentioned warm-up period, when it drives the goodness of the fitting.

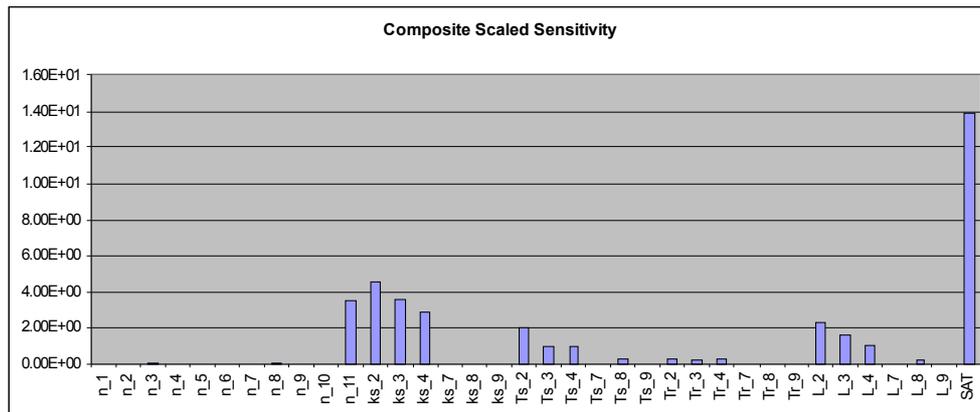


Figure 6.48 – Composite Scaled Sensitivity of the 36 parameters analysed.

The influence of a given observation on the simulation fitting due to its location in the space of the inputs can be evaluated by means of the LEVERAGE statistics. Looking at figure 6.49 where the LEVERAGE calculated for the two station, and the discharges for station 2 are plotted, can be noted that, not taking in account the first warm-up period, the most important observations for a good fitness are the one positioned at the beginning of the first rising limb of an event. Thus those observation should have a large weight in order to force the simulation to closely estimates them and therefore make the model having a good fitting with observations.

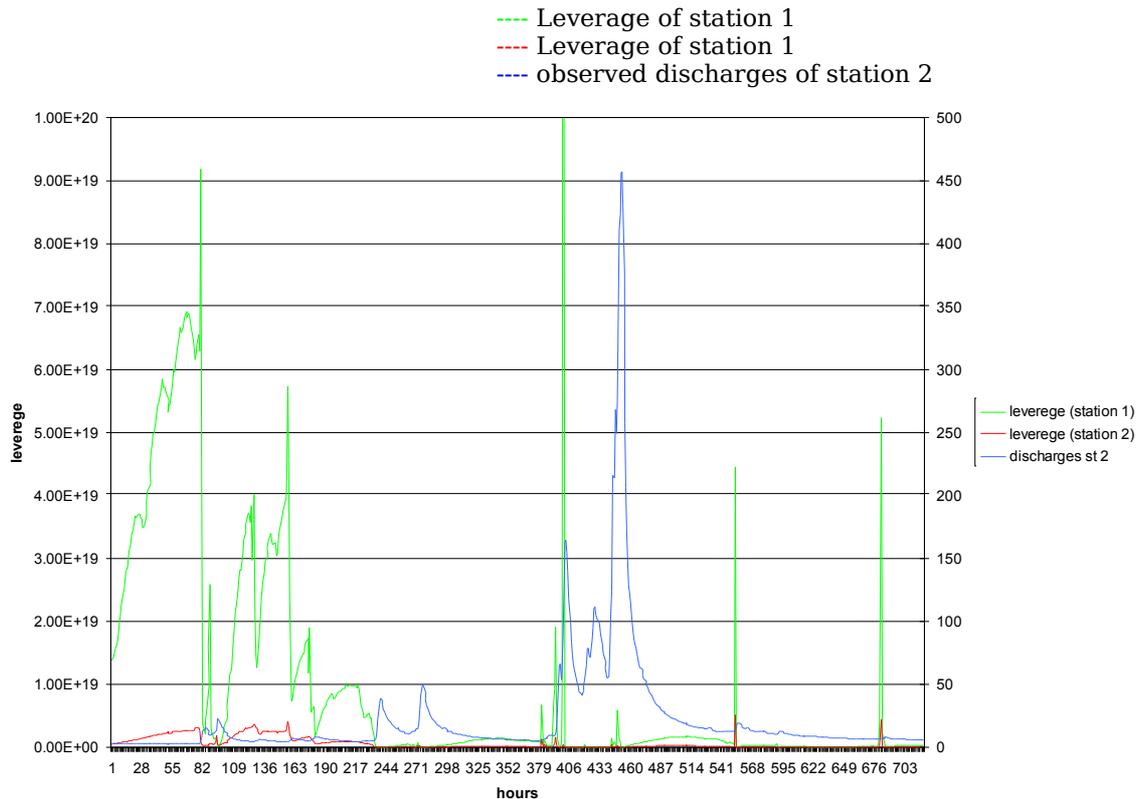


Figure 6.49 – LEVERAGE statistics for station 1 and 2.

### 6.15.2 First model calibration

The calibration was run accordingly to section 4.8.3 by means of the UCODE calibration inverse model. Inverse modelling, or parameter estimation modelling, is an automated calibration technique that works by finding parameter values (e.g., hydraulic conductivities, roughness coefficients) that minimize the sum of the squares errors, also called the objective function, for a given model configuration.

As a result of the sensitivity analysis carried out in the previous section (6.14.1), it has been decided to focus the calibration process upon 8 parameters:  $n_{11}$ ,  $ks_2$ ,  $ks_3$ ,  $ks_4$ ,  $L_2$ ,  $L_3$ ,  $L_4$ , and SAT.

The starting values adopted during the calibration (shows in table 6.9) where chosen as found in literature (Liu and Todini, 2002), (Foglia at al., 2006) where the TOPKAPI model was applied.

In order to reduce the iterations, only the most important parameter of  $ks$  and  $L$  set was directly calibrated. The others parameters were

## Chapter 6 – Case study: the Verzasca basin

proportionally modified at each iteration following the equation:

$$P_i = P_j * P_j' / P_i' \quad [\text{eq. 6.4}]$$

where  $P_i$  is the parameter  $i$ ,  $P_j$  the reference parameter  $j$ ,  $P_j'$  and  $P_i'$  are the starting values of  $P_j$  and  $P_i$  at the beginning of the calibration.

Therefore the introduced relations were:

$$\begin{aligned} ks_3 &= ks_2 * 0.000592 / 0.000654 \\ ks_4 &= ks_2 * 0.000592 / 0.000620 \\ L_3 &= L_2 * 0.2 / 0.6 \\ L_4 &= L_2 * 0.2 / 0.2 \end{aligned} \quad [\text{eq. 6.5}]$$

in this way the calibration were actually performed only on 4 parameters:  $n_{11}$ ,  $ks_2$ ,  $L_2$ , and SAT.

The physical meaning of the results has been controlled by applying upper and lower constrains to the starting values and maximum fractional parameter change allowed between parameter iteration.

n_1	0.210000	Ts_2	0.680000
n_2	0.100000	Ts_3	0.600000
n_3	0.350000	Ts_4	0.520000
n_4	0.250000	Ts_7	0.525000
n_5	0.250000	Ts_8	0.400000
n_6	0.250000	Ts_9	0.450000
n_7	0.300000	Tr_2	0.080000
n_8	0.400000	Tr_3	0.120000
n_9	0.150000	Tr_4	0.120000
n_10	0.150000	Tr_7	0.140000
n_11	0.040000	Tr_8	0.020000
ks_2	0.000600	Tr_9	0.125000
ks_7	0.000055	L_2	0.400
ks_8	0.000001	L_7	0.250
ks_9	0.000130	L_8	0.375
SAT	40.00000	L_9	0.080

Table 6.9 – starting values for the calibration: in coloured cells the calibrated parameters.

## Chapter 6 – Case study: the Verzasca basin

The calibration result shows the following values:

- n_11 (river Manning coefficient)	=	2.3E-02	[s/m <sup>1/3</sup> ];
- ks_2 (regosol soil conductivity)	=	0.678E-03	[m/s];
- L_2 (regosol soil thickness)	=	0.374	[m];
- SAT (starting soil saturation)	=	7.6	[%].

Consequentially the other scaled values are:

- ks_3	=	0.614E-03	m/s
- ks_4	=	0.647E-03	m/s
- L_3	=	0.125	m
- L_4	=	0.374	m

The calibrated values have physical meaning. A partially low value of the river Manning coefficient could be explained with the presence of a rocky river bed for a long part of the Verzasca.

The simulation generated with the calibrated parameters has been plotted against the observations for the two different stations in figures 6.50 and 6.51.

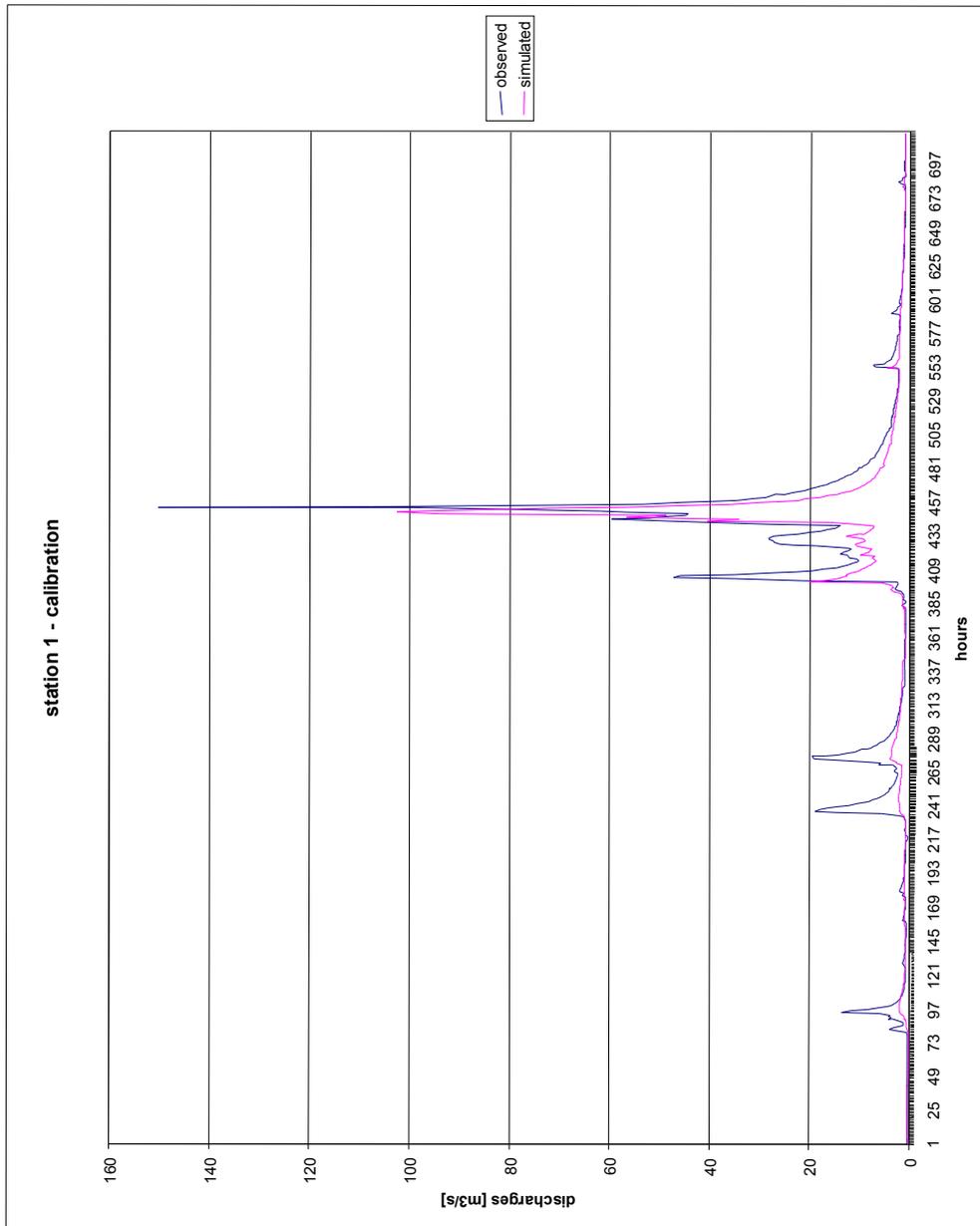


Figure 6.50 – Observed and post calibration simulated discharges at station 1.

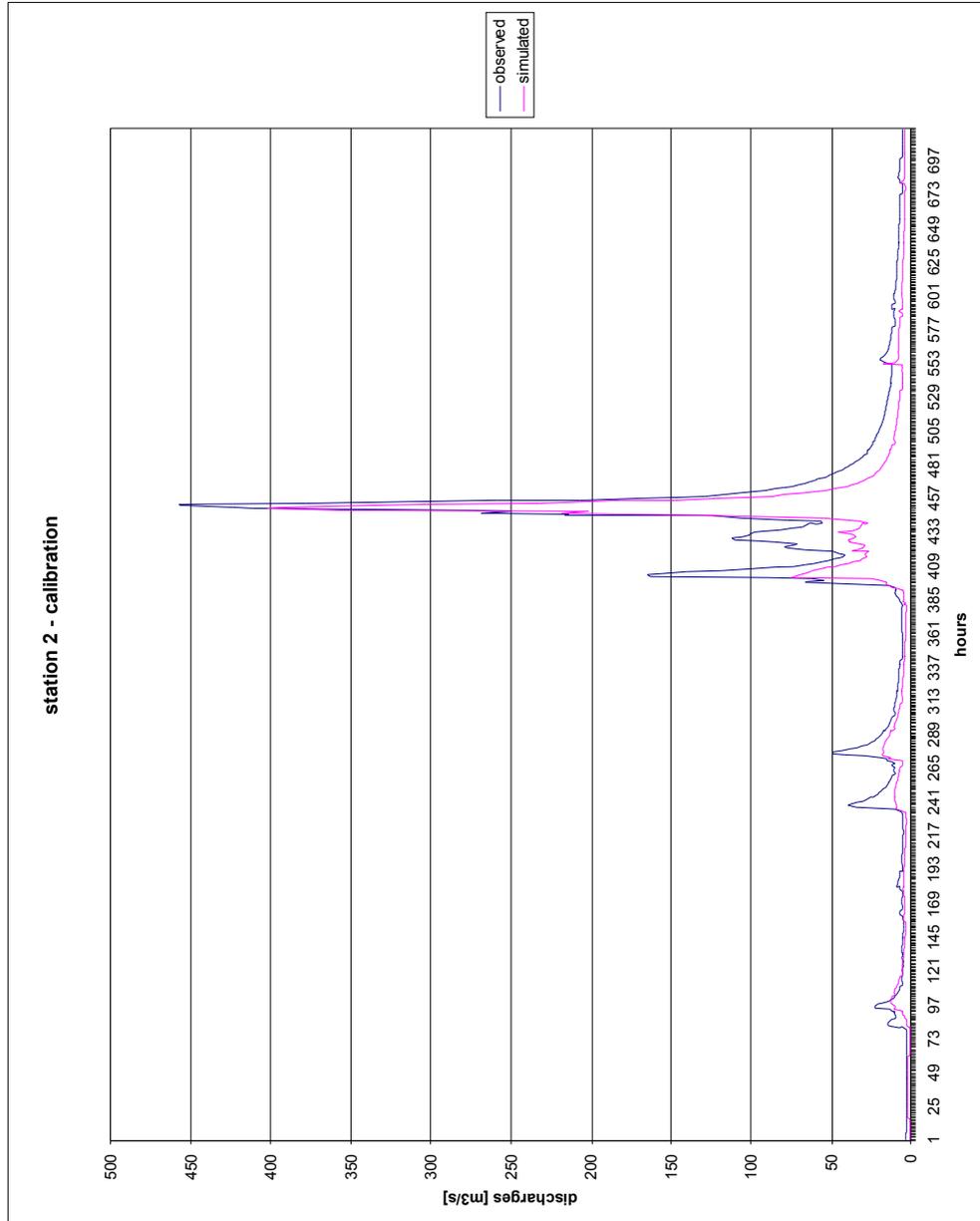


Figure 6.51 – Observed and post calibration simulated discharges at station 2.

## Chapter 6 – Case study: the Verzasca basin

Looking at the figures 6.50 and 6.51 the hydrograph behaviour is well represented, and large peaks occurred at the 457<sup>th</sup> hour appear to be well fitted. This is particularly important, in fact, in floods alert application, be able to reproduce high discharge can help in save life and properties.

The quite good low flows fitting could be partially explained by the nature of the basin, that does not present significant base flows (process not simulated in the model), and by the presuming well simulated snowmelt process.

Calibration improvement are needed to better simulate the recession limb of the hydrograph.

The total error produced by the model simulation, by means of the total discharge in August, is the 37%.

This values, has to be considered satisfactory because this is just the result of a first calibration study on a short period and because the large uncertainties in distributed hydrological modelling hardly lead to much better results.

The validity of the runoff modules can be detected looking at figures 6.52, 6.53, 6.54, 6.55 where the channel water discharge, the overland water discharge, the soil water discharge, and the soil water content are represented before, during, and after the intense rain event of the 20<sup>th</sup> August as an example.

In the figure 6.53 not zero values are present only where the is the river, the values grown accordingly with flow direction, and they rise and decrease following the rain event evolution (from A to C). Therefore the validity of the spatial distribution and behaviour is assured.

The overland flows are generally not present when there is not an heavy or persistent rain (the saturation excess is null), then when the event occurs the soil get saturated and the overland discharges start. The first cell where the overland flow occurs are the ones with higher flow accumulation values, in fact they are saturate from the water directly coming from the rain and from the flows (overland and subsurface) coming from the up-land cells. The figure 6.54 well

## Chapter 6 – Case study: the Verzasca basin

represent the spatial distribution of the overland discharge and its temporal behaviour.

The flows in the soil is slow (subsurface component) is slower than the other two, and the decreasing of the discharges need longer time to be significant. Moreover the magnitude of the discharge depends on the hydraulic head, thus soil types with lower thickness can intake less water and will produce lower discharge values, while cells at the bottom of the valley, that receive water from all the upland ones, will produce higher discharges. In figure 6.55 this spatial distribution and temporal behaviour can be found. The low subsurface discharge values in correspondence of the river are because those the cells are not fully covered by soil, resulting in reduced discharges.

The soil water content rise until a maximum capacity, that depends on the soil type, is reached. As already mentioned the filling of the soil can be done by direct rainfall or upland discharges, thus higher values can be observed at the bottom of the valley or where an heavy rain is present. This component is emptied by evapotranspiration and subsurface flows, thus the water content decrease slowly. Figure 6.56 show the described spatial and temporal behaviour. Fully distributed colors in the map, represent saturated area.

In all these maps, the absence of heavy rain, in the lower east part of the basin, is correctly represented in all the components. Here the channel flow does not increase, no overland flow occurs, low soil water discharges are simulated, and the soil water content do not increase.

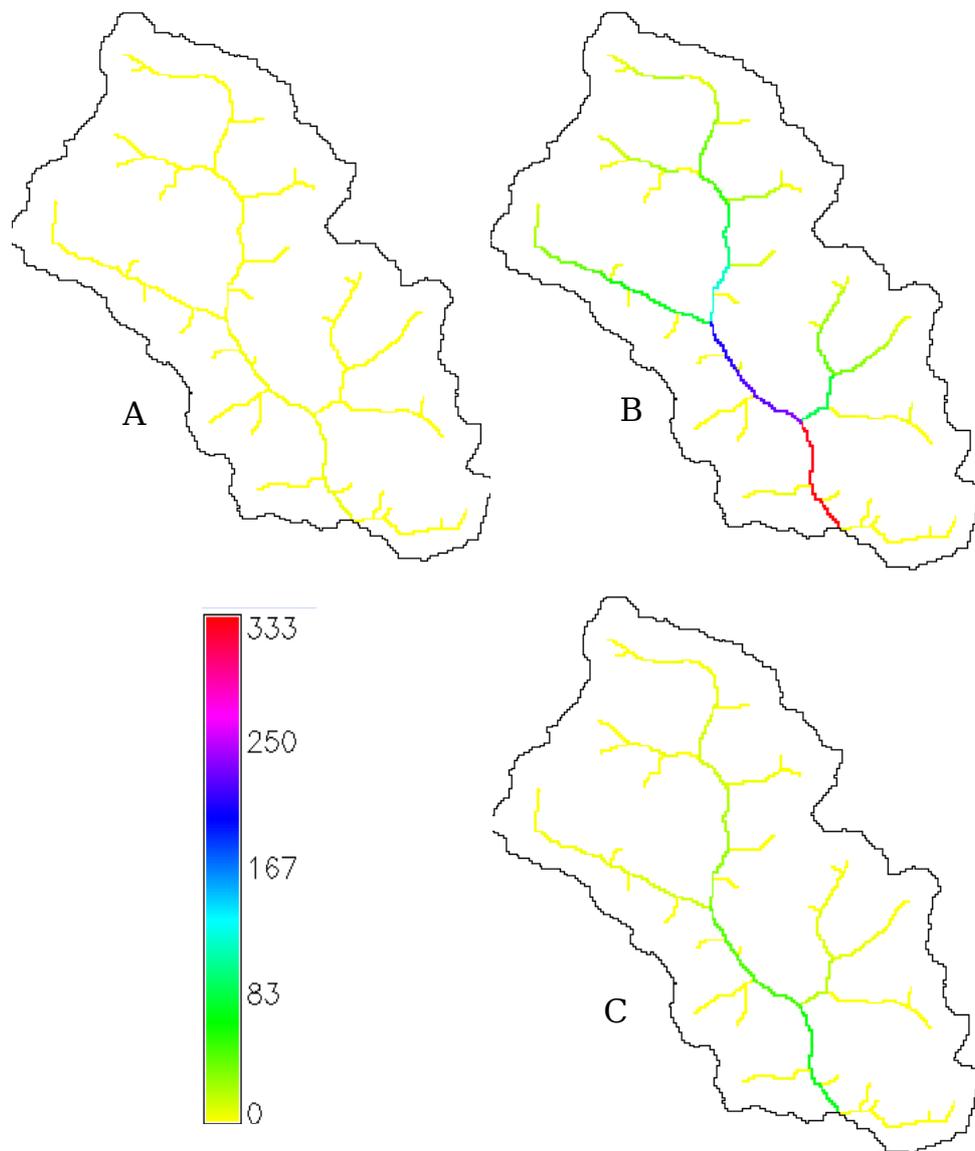


Figure 6.52 – channel water discharge maps [m<sup>3</sup>/s].

A: before an event (17<sup>th</sup> August 2004, 16:00).

B: during the peak (20<sup>th</sup> August 2004, 06:00).

C: during the recession (20<sup>th</sup> August 2004, 16:00).

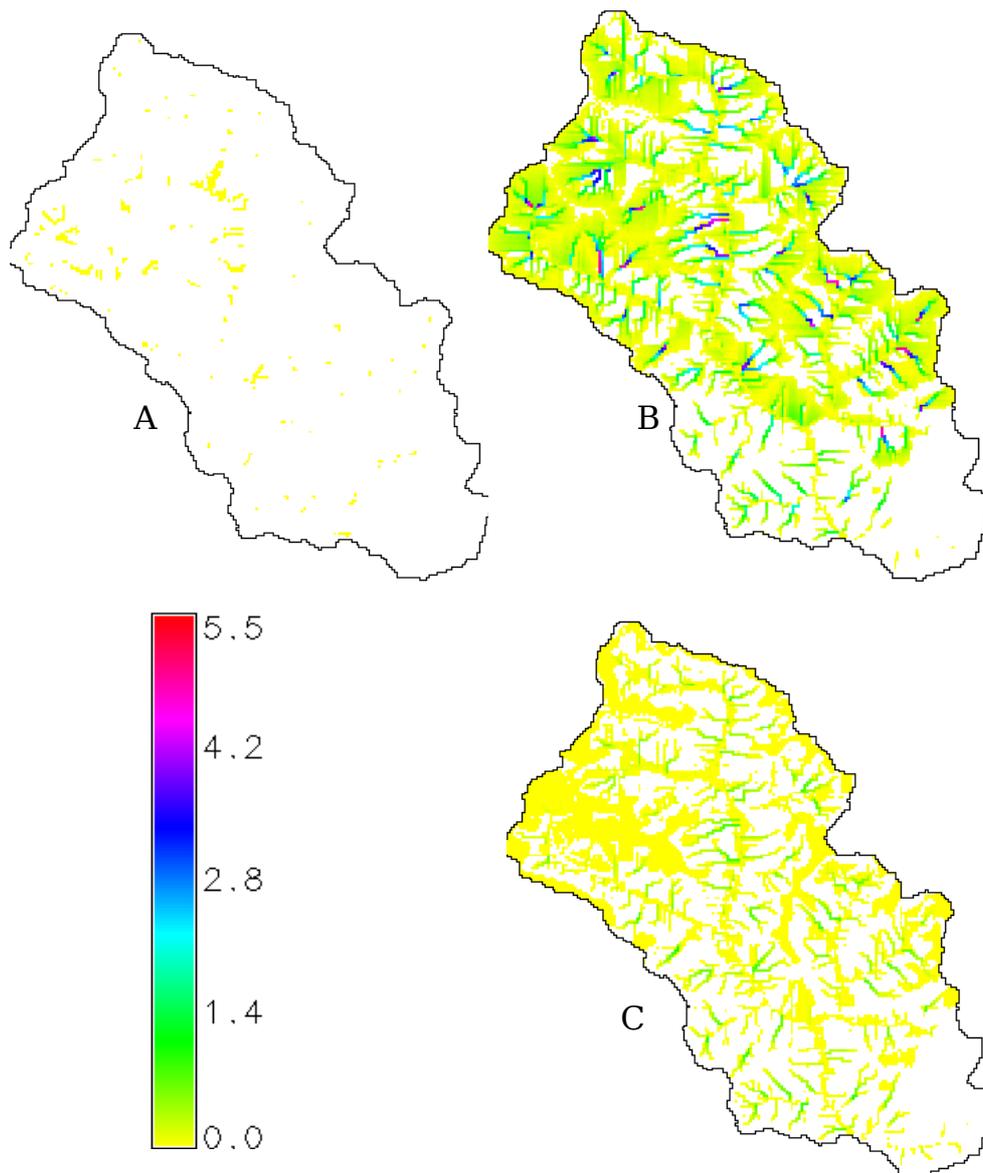


Figure 6.53 – overland water discharge maps [ $\text{m}^3/\text{s}$ ].

A: before an event (17<sup>th</sup> August 2004, 16:00).

B: during the peak (20<sup>th</sup> August 2004, 06:00).

C: during the recession (20<sup>th</sup> August 2004, 16:00).

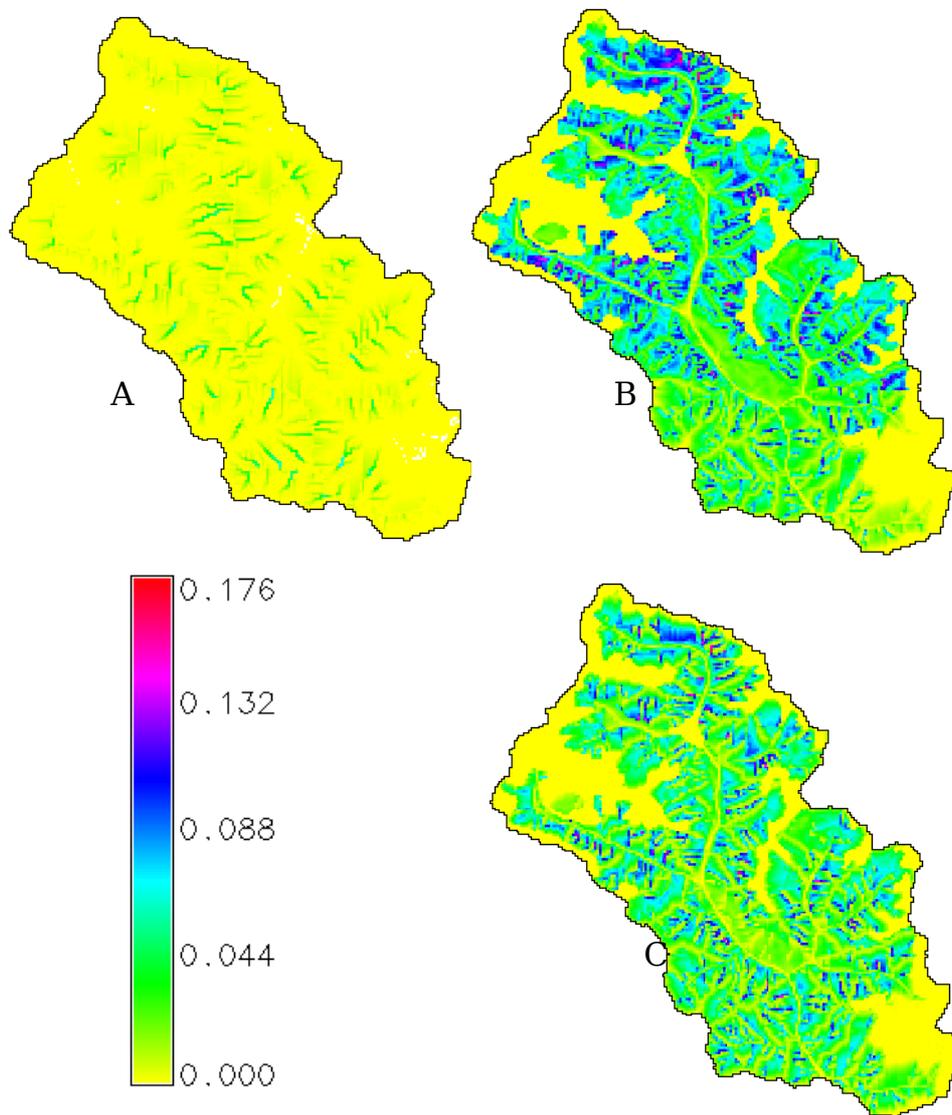


Figure 6.54 – soil water discharge maps [ $\text{m}^3/\text{s}$ ].

A: before an event (17<sup>th</sup> August 2004, 16:00).

B: during the peak (20<sup>th</sup> August 2004, 06:00).

C: during the recession (20<sup>th</sup> August 2004, 16:00).

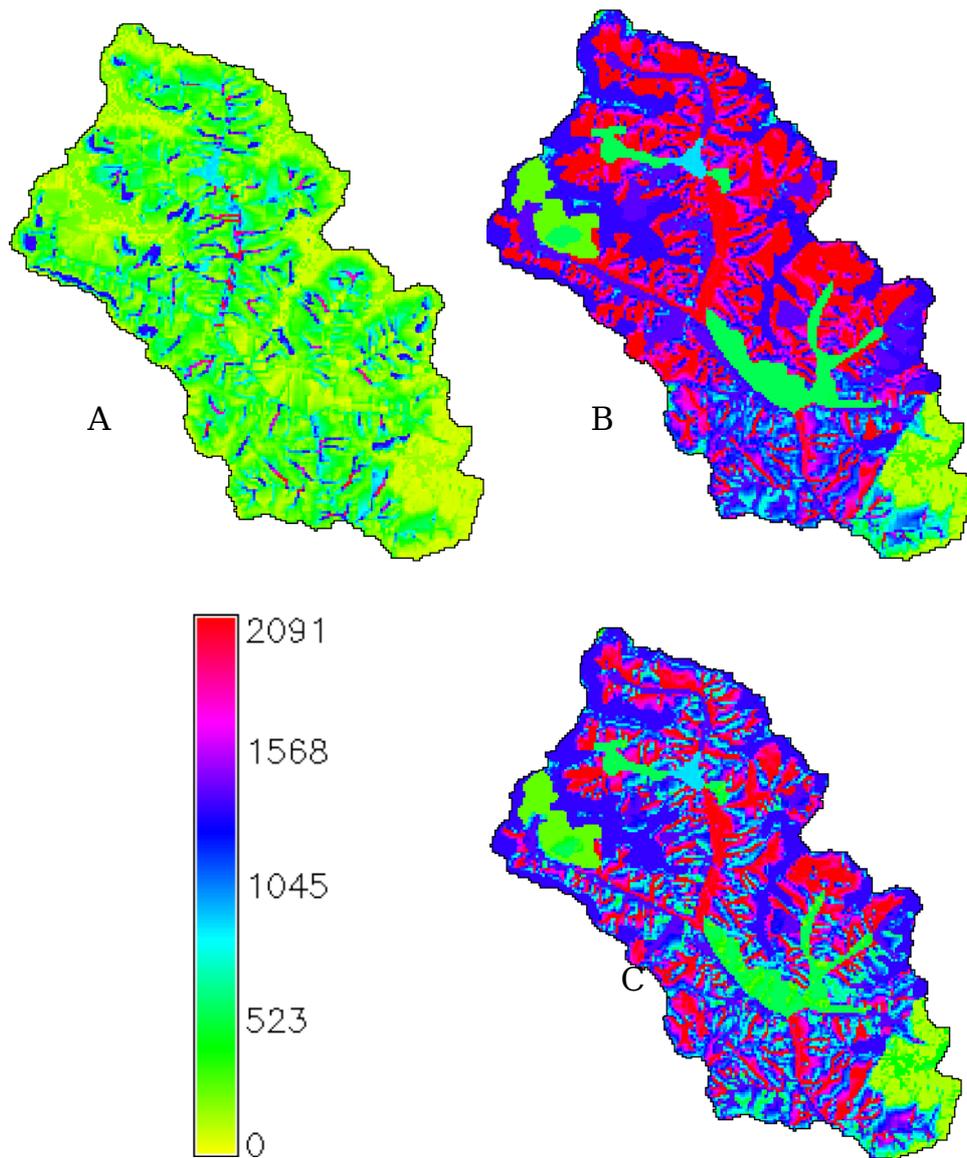


Figure 6.55 – soil water content maps [m<sup>3</sup>].  
A: before an event (17<sup>th</sup> August 2004, 16:00).  
B: during the peak (20<sup>th</sup> August 2004, 06:00).  
C: during the recession (20<sup>th</sup> August 2004, 16:00).

## 7. Conclusions and future works

Nowadays Geographical Information Systems are projected in the era of modelling in order to allow a conscientious environmental management of the resources. In agreement with this tendency, in this work the GIS capability to model the rainfall-runoff has been proven by means of the development of procedures and new commands simulating all the physical processes driving the water flows (HydroFOSS model). Distributed hydrological modelling, allows us to predict the impact of landuse changes and climatic change on water resource and on hydrological regimes of river basins.

Considering the natural capability of GISs to handle, store, represent and analyse spatial information, the embedding of such a kind of model into a GIS system establishes a valuable resource for solving numerous management problems (flood mitigation, storm drainage, bridges, sewerage and culverts design, water supply, irrigation, hydro power, navigation, pollution control, fresh water conservation, etc.).

The usage of a Free and Open Source (FOSS) GIS Software, largely facilitates the task of new procedures and model development due to the accessibility of its code, and availability of documentation.

Moreover, a GIS-embedded approach leads to better accuracy, less duplication, easier map storage, more flexibility, more data sharing and interoperability, timeliness, more efficiency, higher product complexity, and full control of the spatial component.

The followed modular approach, in conjunction with the standard GIS commands and the accessibility of the code, guarantees, to an expert user, the fully customization of the application.

During this work a new approach for raster time series data management, and a database to archive in real time observations of three meteorological station networks, have been developed.

## Chapter 7 – Conclusions and future works

This research has also pointed out the need of a new vector library design able to handle time series with a constant spatial component and a variable attribute component dynamically extracted from a database.

A procedure for input map generation has been developed.

The input climatic raster maps has been generally derived by means of Thiessen tessellation procedure; this “rough” techniques of interpolation is commonly adopted in hydrological modelling due to the lack of sufficient spatially distributed observation.

In order to improve the temperature field estimation, different methods were investigated by means of minimum MAE (mean absolute error). For the Canton Ticino region the linear gradient approach resulted the more appropriate technique.

Input maps of parameters or index characterizing the watershed are also required from distributed hydrological models; these maps has been spatialized by means of soil-landuse characteristics, current time of the year, and altitude.

Climatic and watershed characterizing maps, could be derived from remotely sensed data, although their validity, in term of magnitude of the estimates, has to be proven.

It was carried on a first study of the rainfall estimations derived from ground meteorological radar observations. The results of this study showed that the spatial information is correctly estimated, while the quantitative information is not jet accurate.

More researches in this area, and some application test will be argument of future works.

New commands simulating the snow accumulation and melting, potential evapotranspiration, canopy interception processes, and new procedure to automatically calculate the net radiation have been developed. A first validation has proven their validity, in terms of spatial representation of the phenomena, and mean quantitative values.

## Chapter 7 – Conclusions and future works

A new command for the simulation of the runoff process has been generated and included in a main shell script, that automatically calls, in the appropriate order, all the required procedures and commands to simulate the entire rainfall-runoff process.

Although the model is entirely physically based, its parameters (with a physical meaning), are not directly measurable on the field. In fact they are mean values referred to a cell area (100 m<sup>2</sup> in the case study) where high heterogeneities can be found.

This uncertainties in parameters values ask for calibration procedures aiming at estimate their correct values.

For this reason a linkage between the new developed HydroFOSS model and the FOSS program UCODE-2005 for automatic inverse calibration has been established.

The validity of the runoff module has been proven by means of a sensitivity analysis and calibration in the Verzasca basin. The generated maps are in full accord with all the expected temporal and spatial behaviour of the three runoff components.

Due to the lack of time and the high computational time and power requested for such a kind of model, the sensitivity analysis and the calibration was run on a short period of one month.

Beneath this limited period, the sensitivity analysis pointed out the importance of four different parameters:

- the starting soil saturation level, driving the warm-up period needed to represent the correct soil saturation state. Calibrating this value can result in reduced warm-up periods required.
- The soil conductivity, drives the drainage of the subsurface soil component which mainly characterize the recession limbs of the hydrograph. Thus its correct calibrated value allows a good fit of the simulations in recession periods.
- The soil thickness, drives the saturation excess trigger. The water exceeded quickly flows through the overland component to the river, where a flood peak is generated. Therefore this parameter regulates the goodness of the fit during the rising discharges

periods.

- The river Manning coefficient, being a roughness coefficient, has a friction function driving the river flow velocity. Thus it is important for the correct simulation of dry periods, where the available channel water is quite constant.

The calibration was carried out for the month of August 2004 by using the hourly observations of two limnigraph stations and reveals satisfactory results. The hydrograph behaviour is well represented and the peaks are well fitted.

Sensitivity analysis and calibration depend on the starting parameter values and on the observation values. Therefore different observed periods or catchments, could lead to different calibrated values and model responses. In order to extend the validity of the simulation through time and space more calibrations are needed.

Three main future activities of research have been detected:

- Sensitivity analysis and calibration, on longer periods and different basins. In this way different rainfall event types (thunderstorm, shower, drizzle, etc.) and different field conditions (slopes, aspects, landuses, soiltypes, etc.) can be considered.
- Single module validation. Although the model calibration indirectly accomplishes all the others developed modules, and a first validation of each one has been carried out, a validation executed with direct measurements (not available at this time) will greatly improve the validity of each modules.
- Remotely sensed map investigation. The ability of this technique to well describe the spatial distribution of the phenomena could result in a big improvement for distributed hydrological modelling once quantitatively verified. In particular, for the rainfall intensity, it could be interesting to evaluate the importance of the gain in spatial information (derived from the use of radar data) versus the gain in magnitude measurements

## Chapter 7 – Conclusions and future works

(derived from the use of rain gauges) in hydrological modelling.

# Acknowledgment

This research has been possible due to the support of the IST (Istituto Scienze della Terra, Institute of Earth Sciences), belonging to the DACD (Dipartimento di Architettura, Costruzioni e Design, Department of Architecture, Constructions and Design) of the SUPSI (Scuola Universitaria Professionale della Svizzera Italiana, University School of Applied Sciences of the Southern Switzerland).

A special thanks to Prof. F. Sansò, Prof. M.A. Brovelli and to all the people working at the “Laboratorio di Geomatica” (Geomatic Laboratory) of the “Como Campus” of the “Polytechnic of Milan” that help me in the project development.

Thanks to Helena Mitsova, from the Dept. of Marine, Earth and Atm. Sciences of the North Carolina State University in Raleigh (<http://skagit.meas.ncsu.edu/~helena/>): here notes and suggestions help me in correcting this document and making its content better.

Last but not least, a big thanks to all my family and people I love that allows me to get this PhD degree.

# REFERENCES

- Abbott, M. et al., 1986a, An introduction to the European Hydrological System – Système Hydrologique Européen, SHE; 1. History and philosophy of a physically based distributed modeling system, *Journal of Hydrology*, Vol. 87, 45-59.
- Abbott, M. et al., 1986b, Physically based distributed modelling of an upland catchment using the Système Hydrologique Européen, *Journal of Hydrology*, Vol. 87, 79-102.
- Allen, R.G., L.S. Pereira, D. Raes, and M. Smith. 1998. Crop Evapotranspiration: Guidelines for computing crop water requirements. Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, Rome, pp. 300
- Antenucci, J. C., Brown, K., Croswell, P. L. and Kevany, M. J., 1991. Geographic Information Systems: A Guide to the Technology.
- Beasley, D. B., Huggins, L.F., and Monke, E.J. 1980. ANSWERS: A model for watershed planning. *Trans. of the ASAE* 23(4):938-944.
- Beck, M. B., 1991. Forecasting environmental change, *J. Forecast.*, 10, 3-19.
- Beven, K. J., 1981, Kinematic subsurface stormflow, *Water Resources Research*, 17(5): 1419-1424.
- Beven, K. J., 1982, On subsurface stormflow: prediction with simple kinematic theory for saturated and unsaturated flows, *Water Resources Research*, 18(6): 1627-1633.
- Beven, K. J., 1989, Changing ideas in hydrology: the case of physically-based models, *Journal of Hydrology*, 105: 157-172.
- Beven, K. J., 1997, TOPMODEL: a critique, *Hydrological Processes*, vol. 11, 1069-1085.

## References

- Beven, K. J., 2001, Rainfall-runoff modeling – The Primer. Chichester, John Wiley & Sons Ltd, 360
- Beven, K. J., and Kirkby, M.J., 1979, A physically based, variable contributing area model of basin hydrology, *Hydrological Science Bulletin*, 24(1): 43-69.
- Beven, K.J., Lamb, R., Quinn, P., Romanowicz, R., and Freer, J., 1995, TOPMODEL, in *Computer Models of Watershed Hydrology*, Singh V.P. (ed.), Water Resources Publications, 627-668.
- Bian, L., Sun, H., Blodgett, C., Egbert, S., Li,W., Ran, L., and Koussis, A., 1996. An integrated interface system to couple the SWAT model and ARC/INFO. Proc., 3rd Int. Conf./Workshop on Integrated Geographic Information Sys. and Envir. Modeling., UC Santa Barbara.
- Blake, G.J., 1975. The interception process. In: *Prediction in catchment hydrology*. Australian Academy of Science, Canberra, Australia, pp. 59-81.
- Borga M., Tonelli F., Moore R.J. and Andrieu H., 2002. Long-term assessment of bias adjustment in radar rainfall estimation, *Water Resour. Res.*, Vol. 38, No. 11, pp. 1226-1235.
- Chorley, R. J., and Haggett, P., 1967. *Models in geography*. Methuen, London.
- Chow, V .T., 1973. *Open Channel Hydraulics*, McGraw-Hill, New York.
- Chumchean S, Sharma A. and Seed A., 2003. Radar rainfall error variance and its impact on radar rainfall calibration, *Phys. Chem. of the Earth*, Vol.28, pp. 27-39.
- Ciarapica, L., Todini, E., 2002. TOPKAPI: a model for the representation of the rainfall-runoff process at different scales, *Hydrological Processes* 16, 207-29.

## References

- Collins F.C.Jr., Bolstad, P.V., 1996. A Comparison of Spatial Interpolation Techniques in Temperature Estimation. Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling. January 21-25, 1996, Santa Fe, New Mexico, USA
- COST Project 72, 1985. European Commission Report EUR 10353.
- Crain, I.K. And MacDonald, C.L. ,1984. From land inventory to land management - the evolution of an operational GIS. *Chartographica* 21: 39-46.
- Crawford, N.H. and Linsley, R.K., 1966. Digital simulation in Hydrology: Stanford Watershed Model IV, Technical Report n. 39, Stanford University, Palo Alto, California.
- Dawdy, D.R. and O'Donnel, T., 1965. Mathematical models of catchment behavior, *J. Hydraul. Div., Am. Soc. Civ. Eng.*, 91 (HY4), 123-127.
- De Roo, A.P.J., Wesseling, C.G., Van Deursen, W.P.A., 2000. Physically-based river basin modelling within a GIS: The LISFLOOD model. *Hydrological Processes* 14, 1981-92.
- DeBarry, P. A., 1991. GIS: Applications in nonpoint source pollution assessment. Proc., ASCE 1991 Nat. Conf. on Hydr. Engrg., ASCE, Reston, Va.
- Doe, W. W., Saghafian, B., and Julien, P. Y., 1995. Land-use impact on watershed response: The integration of two-dimensional hydrologic modeling and GIS. *Hydrological Processes*, 10, 1503-1511.
- Dunne, T. and Black, R.D., 1970. Partial area contributions to storm runoff in a small New England watershed, *Water Resources Research*, 6(5), 1296-1311.
- Dunne, T., 1978. Field studies of hillslope flow process. In M.J. Kirkby (Ed.), *Hillslope Hydrology*, Wiley, New York, 227-293.

## References

- Dunnie, T., Moore, T.R., Taylor, C.H., 1978. Recognition and prediction of runoff-producing zones in humid regions., *Hydrological Sciences Bulletin* 20(3): pp. 305-327.
- Fetter, C.W., 1994. *Applied Hydrogeology*, Third Edition, Chapter 10, Macmillan.
- Foglia, L., Mehl, S.W., Hill, M.C., Perona, and Burlando, P., 2006. Evaluation of information criteria, linear statistics, cross validation, and predictions using alternative groundwater models, submitted to *Groundwater*.
- Freeze, R.A. and Harlan, R.L., 1969, Blueprint for a physically-based digitally simulated hydrologic response model, *Journal of Hydrology*, n. 9, 237-258.
- Freeze, R.A. and Harlan, R.L., 1969, Blueprint for a physically-based digitally simulated hydrologic response model, *Journal of Hydrology*, n. 9, 237-258.
- Gjersten U., Salek M. and D.B. Michelson, 2003. Gauge-adjustment of radar-based precipitation estimates – a review. Cost Action 717, Working Document WDD\_02\_200310\_1.
- GRASS Development Team, 2006. GRASS 6.0 Users Manual. ITC-irst, Trento, Italy. Electronic document: [http://grass.itc.it/grass61/manuals/html61\\_user/](http://grass.itc.it/grass61/manuals/html61_user/)
- Green, W.H. and Ampt, C.A., 1911. Studies on soil physics: 1. Flow of water and air through soils. *Journal of agricultural Science*, 4, 1-24.
- Hamlett, J. M., Miller, D. A., Day, R. L., Petersen, G. W., Baumer, G. M., and Russo, J., 1992. Statewide GIS-based ranking of watersheds for agricultural pollution prevention. *J. Soil and Water Conservation*, 47, 300–404.
- Harr, R.D., 1981. Some characteristics and consequences of snowmelt during rainfall in western oregon, *Journal o f Hydrology*, nr. 53, pp. 277-304.
- Harris, T. M. and Elmes, G. A., 1993. The Application of GIS in Urban and Regional Planning: a review of the North American experience. *Applied Geography* 13(1), 9-27.

## References

- Henderson, F.M., Wooding, R.A., 1964. Overland flow and groundwater flow from a steady rainfall of finite duration, *J. Geophys. Res.*, 69 (6), 1531-1540.
- Hill M.C., 1998. Methods and guidelines for effective model calibration. US Geological Survey; Water resource investigation report, 98-4005.
- Hill, M.C. and C.R. Tiedeman, 2005. Effective calibration of environmental models, with analysis of data, sensitivities, predictions, and uncertainty, in press.
- Hill, M.C., Banta, E.R., Harbaugh, A.W., and Anderman, E.R., 2000. MODFLOW-2000, The U.S. Geological Survey modular ground-water model, User's guide to the observation, sensitivity, and parameter-estimation processes: U.S. Geological Survey Open-File Report 00-184, 209 p.
- History of ESRI. Accessed 01/17/06: <http://www.esri.com/company/about/history.html>
- Hofierka, J., Suri, M., 2002. The solar radiation model for Open source GIS: implementation and applications. Manuscript submitted to the International GRASS users conference in Trento, Italy, September 2002.
- Horton, R.E., 1933, The role of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union*, 14:446-460.
- INTERA TYDAC Technologies Inc. 1992. SPANS Reference Manual. Ottawa.
- Joss J., Schädler B, Galli G., Cavalli R., Boscacci M., Held E., Della Bruna G., Kappenberger G., Nespor V. and Spiess R., 1997. Operational use of radar for precipitation measurements in Switzerland, [www.meteoswiss.ch](http://www.meteoswiss.ch).
- Kasten F., 1996. The Linke turbidity factor based on improved values of the integral Rayleigh optical thickness. *Solar Energy*, 56 (3), 239-244.

## References

- Knisel, W. G. 1993. GLEAMS: Groundwater loading effects of agricultural management systems, University of Georgia, Coastal Plains Experiment Station, Biological and Agricultural Engineering Department, Publication No. 5, pp. 260
- Knyazikhin, Y., Glassy, J., Privette, J.L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J.T., Votava, P., Myneni, R.B., Nemani, R.R., Running, S.W., 1999. MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document, <http://eosps0.gsfc.nasa.gov/atbd/modistables.html>, 1999.
- Krajewski W.F., 1987. "Co-kriging radar-rainfall and rain gauge data", J. Geophys. Res., Vol. 92 , pp. 9571-9580.
- Liu, Z, and Todini, E., 2002, Towards a comprehensive physically-based rainfall-runoff model, Hydrology and Earth System Science, 6, 859-881.
- Liu, Z., Martina, M. L.V., and Todini, E., 2005, Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian catchment, Hydrology and Earth System Sciences, 9(4), 347-364.
- Lombardo F. and Staggi L., 1998. Verifica e taratura dinamica della strumentazione pluviometrica finalizzata alla valutazione degli errori per intensità di pioggia elevata, Proc. XXVI Conv. Idr. Costr. Idrauliche, Genova, Vol. 2, pp. 85-96.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, Journal of Hydrology, 235 (2000), 276-288.
- Madsen, H., 2003. Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives, Advances in water resources, 26, p. 205-216.

## References

- Maidment, D. R., 1991. GIS and hydrologic modeling. Proc., 1st Int. Symp./Workshop on GIS and Envir. Modeling.
- Mark, D. M., Chrisman, N., Frank, A. U., McHaffie, P. H. and Pickles, J., 1997. The GIS History Project. Retrieved 15 January 2001 from the World Wide Web: [http://www.geog.buffalo.edu/ncgia/gishist/bar\\_harbor.html](http://www.geog.buffalo.edu/ncgia/gishist/bar_harbor.html)
- Menzel L. Modellierung der Evapotranspiration im System Boden-Pflanze-Atmosphäre, 1997. PhD Dissertation, Federal Institute of Technology Zurich, Geographical Institute Book No. 67 (in german).
- Menzel, L., 1997. Modelling canopy resistances and transpiration of grassland. Physics and Chemistry of the Earth, special issue: Atmospheric and Hydrological Processes and Models at the Soil Vegetation Atmosphere Interface.
- Meteorological Organization (WMO), 1981. Technical Note N° 172, WMO-No. 554, Geneva, Switzerland, pp. 121-123.
- Michelson D.B., 2003. Quality control of weather radar data for quantitative application, PhD thesis, Telford Inst. Env. System., Salford University, 281 pp.
- Mitas, L. and Mitasova, H., 1999. Spatial Interpolation. In: P.Longley, M.F. Goodchild, D.J. Maguire, D.W.Rhind (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications, Wiley.
- Mitasova, H. and Mitas L., 1993. Interpolation by Regularized Spline with Tension: I. Theory and implementation. Mathematical Geology, no. 25, pp. 641-655.
- Mitasova, H. and Mitas, L., 2001, Modeling Physical Systems, In: Geographic Information Systems and Environmental Modeling, Parks B., Crane M. and Clarke, K eds., Prentice Hall, 189-210.

## References

- Montanari A. and E. Toth, 2006. Calibration of spatially-distributed hydrological models: possible strategies and analysis of the effect of data availability. Presentation at the EGU General Assembly, Vienna, Austria, 02-07 April 2006. ([www.costruzioni-idrauliche.ing.unibo.it/people/alberto](http://www.costruzioni-idrauliche.ing.unibo.it/people/alberto), 2006).
- Morris, E.M., 1991. Physics-based models of snow. In: Bowles, D.S., O'Connell, P.E. (Eds.), *Recent Advances in Modeling of Hydrological Systems*, Kluwer Academic, Dordrecht, pp. 85-112.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., King, K.W., 2002, *Soil and Water Assessment Tool Theoretical Documentation*, version 2000, Texas Water Resources Institute, College Station, Texas, TWRI Report TR-191.
- Neteler M., 2005. GRASS 6 Programmer's Manual, GRASS Development Team, Published under GNU Free Documentation License (GFDL), <http://mpa.itc.it/markus/grass61progman/>
- Nizeyimana, E., 1997. Quantification of NPS pollution loads within Pennsylvania. Publ. No. ER97-08, *Envir. Resour. Res. Inst.*, Pennsylvania State University, University Park, Pa.
- Ogden, F. L. 1998 . *CASC2D version 1.18 reference manual*. Dept. of Civil & Environmental Engineering Rep. U-37, CT1665-1679, Univ. Connecticut, Storrs, Conn.
- Pandey, A. and Sahu, A.K., 2002, *Generation of curve number using Remote Sensing and Geographic Information System, GIS Development*, <http://www.gisdevelopment.net/application/nrm/water/watershed/watws0015.htm>
- Penman, H. L. 1948. Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London*, A193, pp. 120-146.
- Poeter E.P. and Hill, M.C., 1996. Unrealistic parameter estimates in inverse modeling: a problem or a benefit for model calibration? *Proceedings ModelCARE Conf.*, p. 227.

## References

- Poeter E.P. and Hill, M.C., 1997. Inverse models: A necessary next step in ground-water modeling, *Ground Water* 35(2), 250-260.
- Poeter, E.P. and Anderson, D.R., 2005. Multi-model ranking and inference in groundwater modeling, *Ground Water*, 43, no. 4, p. 597-605.
- Poeter, E.P., Hill, M.C., Banta, E.R. 2005. MMRI, A computer code for Multi-Model Ranking and Inference, a JUPITER API Application Code, U.S. Geological Survey, Water-Resources Investigations report.
- Rango, A. and Martinec, J., 1995. Revisiting the degree-day method for snowmelt computations, *Water Resources Bulletin*. Vol. 31, no. 4, pp. 657-669.
- Refsgaard, J.C., 1997. Parametrisation, calibration and validation of distributed hydrological models, *Journal of Hydrology*, 198: 69-97.
- Renschler, C.S. 2003. Designing geo-spatial interfaces to scale process models: The GeoWEPP approach. *Hydrological Processes* 17, p.1005–1017.
- Rewerts, C. C., and Engel, B. A., 1993. ANSWERS on GRASS: Integration of a watershed simulation with a geographic information system. Abstracts, Proc., 8th Annu. GRASS GIS User's Conf. and Exhibition, Conf. Agenda and Listing of Abstracts.
- Rosenthal, W. D., Srinivasan, R., and Arnold, J., 1993. A GIS-watershed hydrology model link to evaluate water resources of the lower Colorado river in Texas. Proc., Application of Advanced Information Technol.
- Sansò F., 1996. Elementi di teoria della probabilità, 1st edition, Città Studi Edizioni, pp. 111-119.
- Savabi, M. R., Flanagan, D. C., Hebel, B., and Engel, B. A., 1995. Application of WEPP and GIS-GRASS to a small watershed in Indiana." *J. Soil and Water Conservation*, 50(50), 477–483.

## References

- Schmid, F.; Fraefel, M.; Hegg, C., 2004. Unwetterschäden in der Schweiz 1972-2002: Verteilung, Ursachen, Entwicklung. - Wasser Energ. Luft 96, 1/2: 21-28.
- Schulla J. and Jasper, K., 1999. Model description WaSim, IHW-ETH, Zurich
- Schulla J., 1997. Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen, PhD. Thesis ETH, Zurich
- Seo D.J., 1998. Real-time estimation of rainfall fields using radar rainfall and rain gage data, J. Hydrol., Vol. 208, pp. 37-52.
- Shuttleworth, W. J., 1988. Evaporation from Amazonian Rainforest, Proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 233, No. 1272 (Apr. 22, 1988) , pp. 321-346
- Shuttleworth, W. J., 1992. Evaporation, Chapter 4 in Handbook of Hydrology, Edited by D. R. Maidment, McGraw-Hill, New York.
- Singh, V.P. and Woolhiser, D.A., 2002. Mathematical Modeling of Watershed hydrology, Journal of Hydrologic Engineering, 7:4 (270-292).
- St-Onge, B.A. and Achaichia N., 2001. Measuring forest canopy height using a combination of lidar and aerial photography data, International Archives of Photogrammetry and Remote Sensing, Volume XXXIV-3/W4 Annapolis, MD, 22-2r4 Oct. 2001.
- Strahler, A.H.,Muller, J.-P., MODIS Science Team Members, 1999. MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document Version 5.0, [http://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod09.pdf](http://modis.gsfc.nasa.gov/data/atbd/atbd_mod09.pdf), 2006.
- Tarboton, D. G. and C. H. Luce, 1996. Utah Energy Balance Snow Accumulation and Melt Model (UEB): Computer model technical description and user's guide, Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station.

## References

- Tarboton, D. G., R. Woods, C. Pearson, U. Shankar, R. Henderson and R. Ibbitt, 1999. Distributed Hydrologic Modeling using GIS and Topmodel, Presentation at 19th Annual AGU Hydrology Days, Fort Collins, Colorado, August 16-20
- Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2): 309-319.
- Thorntwaite, C.W., and Mather, J.R., 1955. The water balance, Publ. in *Climatology*, 8(1), Lab. of Climatology, Centerton, N.J.
- Todini, E., 1988. Rainfall-runoff modeling – past, present and future, *Journal of Hydrology*, 100, 341-352.
- Todini, E., 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, *Hydrol. Earth System Sci.*, Vol. 5, No. 2, pp. 187-199.
- Todini, E., 2005. Multi-Sensor Precipitation Measurements, Integration, Calibration and Flood Forecasting (MUSIC), Final Report, EU Project EVK1-CT-2000-0058.
- Todini, E., and Ciarapica, L., 2001, The TOPKAPI model, (chapter 12), in: *Mathematical models of large watershed hydrology*, V.P. Singh et al. (Eds), Water Resources Publications, Littleton, Colorado.
- Wesseling, C.G., D. Karssenbergh, W.P.A. van Deursen, and P.A. Burrough., 1996. Integrating dynamic environmental models in GIS: The development of a Dynamic Modelling language, *Transaction in GIS* 1, 40-48.
- Wigmosta, M. S., Vail, L. W., and Lettenmaier, D. P. A., 1994. Distributed hydrology vegetation model for complex terrain, *Water Resources Research*, 30(6), 1665-79
- WMO, 1986. Intercomparison of Models of Snowmelt Runoff, Operational Hydrology Report No. 23, World Meteorological Organization - No. 646.

## References

- Wooding, R.A., 1965. A hydraulic model for the catchment-stream problem: I. Kinematic-wave theory; II. Numerical solutions; III. Comparison with runoff observations, *Journal of Hydrology*, Volume 191, Number 1, April 1997, pp. 371-371(1)
- Xu, C.Y. and Chen, D., 2005. Comparison of seven models for estimation of evapotranspiration and groundwater recharge using lysimeter measurement data in German, *Hydrol. Process.* 19, pp. 3717–3734
- Y. Knyazikhin, J. Glassy, J. L. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J. T. Morisette, P. Votava, R.B. Myneni, R. R. Nemani, S. W. Running, MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document, <http://eosps0.gsfc.nasa.gov/atbd/modistables.html>, 1999.
- Ye, Z., Maidment, D.R., McKinney, D.C., 1996. Map-based surface and subsurface flow simulation models: an object-oriented and GIS approach, University of Texas, Austin.
- Zollmann, C., 2000. Nachhaltige Bewässerungslandwirtschaft in semi-ariden Gebieten, Entwicklung einer Upscaling technik für die Hydrologische Modellierung, PhD. Thesis, ETH, Zurich.

# Appendix: Meteorological Network Stations Database

In this appendix a description of the Meteorological Network Stations Database (MNSD) tables is described.

## *Measurement\_points*

It archives the spatial component of the climatic stations, it has the following attributes:

- `feature_id` – the station identification number stored as long integer and set as primary key;
- `east`, `north`, `elev` – the coordinates stored as integer values;
- `descr` – the description of the station, stored as `varchar(40)`.

## *Measurements*

It stores the observed values, it has the following attributes:

- `feature_id` – the station identification number stored as long integer and set as primary key and as foreign key;
- `param_id` – the parameters identification number stored as integer and set as primary key and foreign key;
- `meas_time` – the time of the observation, stored as timestamp;
- `meas_value` – the observed value, stored as real;

## *Parameters*

It is the look-up table to associate informations with different parameters type, it has the following attributes:

- `param_id` – the parameters identification number stored as integer and set as primary key;
- `name` – the name of the observation type stored as `varchar(5)`;
- `descr` – the description for the observation type stored as `varchar(20)`;
- `unit` – the unit of the observation type stored as `varchar(10)`;

## Appendix: Meteorological Network Stations Database

### *Instruments*

It stores information about instruments installed at different measure points, it has the following attributes:

- `instr_id` – the unique identifier for each instrument stored as long integer and set as primary key;
- `feature_id` – the station identification number stored as long integer and set as foreign key;
- `owner_id` – the identification number of the organization owner of the instrument stored as integer and set as foreign key;
- `type_id` – the identification number of the instrument type stored as integer and set as foreign key;
- `serial` – the serial number of the instrument stored as `varchar(30)`;

### *Instr\_type*

it is the look up table to relate instrument identifier with its informations, it has the following attributes:

- `type_id` – the identification number of the instrument type stored as integer and set as primary key;
- `brand` – the name of the brand of the instrument stored as `varchar(20)`;
- `version` – the version of the instrument stored as `varchar(20)`;

### *Organizations*

It stores the owner of the different instruments, it has the following attributes:

- `owner_id` – the identification number of the organization owner of the instrument, stored as integer and set as primary key;
- `name` – the name of the organization owing the instruments stored as `varchar(20)` and set as primary key;
- `address` – the address of the organization stored as `varchar(50)`;
- `town` – the town of the organization stored as `varchar(15)`;
- `nation` – the nationality of the organization stored as `varchar(15)`;
- `phone` – the phone number of the organization stored as `varchar(20)`;

## Appendix: Meteorological Network Stations Database

- email – the contact email address stored as varchar(30);
- contact – the identifier name of the person of contact stored as varchar(30);

### *Maintenance*

It archives the maintenance operations done on the instruments, it has the following attributes:

- main\_id – the unique identifier of each operation stored as long integer and set as primary key;
- main\_time – the date and time of the operation stored as timestamp;
- descr – the description of the executed operation stored as varchar(50);
- instr\_id – the identifier of the instrument stored as long integer and set as foreign key;
- person\_id – the name of the person that executed the operation stored as long integer and set as a foreign key;

### *Persons*

It store information about the person involved in the technical maintenance of the network, it has the following attributes:

- pers\_id – the unique identifier of the person involved in the network maintenance, stored as long integer and set as primary key;
- name – first name of the person stored as varchar(20);
- last – family name of the person stored as varchar(20);
- phone – the phone number of the person stored as varchar(20);
- email – e-mail of the person stored as varchar(30);