RUNOFF MODELING USING GIS

APPLICATION IN TORRENTIAL BASINS IN THE APUSENI MOUNTAINS

Matei Domnița

Table of Contents

Acknowledgements

I would like to express my deep gratitude to Professor Ionel Haidu and Professor Okke Batelaan, my research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

1. Introduction

1.1 General aspects of the study

Because of irrational land exploitation and lack of flood protection infrastructure, Romania was and is still vulnerable to frequent floods. In the recent years, the number of these events has risen and Romania needed the help of the European Comission. In the summer of 2008, the European Comission gave a financing of11.78 million Euro and in the spring and autumn of 2005 another 71.2 million Euro were offered. According to the European Comission website, the help of EU was used mainly for the costs of emergency intervention and remake of the important flood protection infrastructure.

As evidenced by recent dramatic events, the frequency and severity of extreme weather phenomena, such as droughts, floods and associated landslides, storms, heat waves, forest fires and cold snaps, have increased over the last few decades. Impacts associated with these events appear to be due to several factors, including climate change and global warming, the growth of urban development and population density in exposed areas, and a higher concentration of assets and values at risk (Monti 2009). For the central and western areas of Romania different studies show an increase in rainfall events that create torrents and flash floods (NWS, 2006; Haidu and Bilaşco, 2007, Sorocovschi and Haidu 2003).

The "Romanian Waters" National Administration requested the Regional Water Branches to strenghten the communication with the county committees for emergency situations to help monitor and prevent the unwanted effects. The dams for lakes were also secured and rebuilt where needed and structures for controlling the discharge were built (ANAR 2011)

During a project in "risk mitigation and emergency preparedness in the event of natural disasters" the areas which have accumulation lakes or other discharge control structures and the lower courses of rivers in Romania were reinforced or rebuilt to strenghten the flood protection infrastructure. But most mountain areas still have flooding problems due to little or no maintenance of flood protection infrastructure in these areas. Besides these problems, Romania was also affected by uncontrolled deforestation without the removal of vegetation leftovers from the deforested areas (Fig. 1.1). The massive deforestation in the last decades raised the runoff coefficients and reduced the infiltration and retention, so a higher volume of rainfall becomes runoff which concentrates as flash floods in these areas. Therefore, a high percent of mountain

settlements are affected by torents and flash floods.

Fig. 1.1 *Deforestation in the Romanian Carpathians, Aug. 2006*

The path of runoff and occurence of torrents depends on the movement of the convective cell and the variation of rainfall intensity during the storm. Therefore the elaboration of a method to anticipate the effects of torrential rainfalls based on statistic data (Haidu, 2003; Sohn 2005) or dynamic models (Ebert 2005) is needed.

The impact of torrential rainfalls in small and very small basins is very high. These basins do not usualy have discharge controlling structures or other flood protection infrastructure, and their dimension allows the rapid concentration of rainfall. Drobot (2007) considers that a flood produces by torrential rainfall can be considered a flash flood in the following conditions: the catchment area varies from a few square kilometers to 200 km^2; the concentration time is less than six hours; the rainfall duration is smaller than the concentration time (maximum three hours); the flood is caused by a rainfall depth over 100mm.

A number of hydrology studies were carried out on the Apuseni Mountains and most of them were PhD theses. These studies include the works of Buta (1967), Iacob Ersilia (1971); Aniţan (1974) or more recent ones by Şerban (2004), Patko (2007), Arghiuş (2008), Bilaşco (2008), Crăciun (2011).

The surface runoff is the runoff that appears on the surface of the land in streams or as a

thin sheet of water flowing over the landscape. The main reason for studying surface runoff is the importance of this phenomenon in the occurence of flash floods. The runoff that appears flows over the land and quickly accumulates in the nearest watercourse downstream from the source. The fast concentration of water from a large surface makes the receiving stream unable to convey the great quantity of water and causes flash floods (Fig. 1.2).

The flash flood is a flood that follows the causing event (storm) in a very short time and manifests like a sudden increase in the water level and flow speed. The term „flash" is used to suggest the short time between the start of the rainfall and the maximum discharge of the flood, usually between some minutes up to some hours from the event, leaving a very short time for preparations and intervention. Usually, a threshold of six hours is used to distinguish a flash flood from a normal flood which has a slow increase in water level (Mogil et al, 1978; Georgakakos, 1986a; Gruntfest and Huber 1991). Most of the flash floods occur in basins with a drainage area of less than 100 km^2 (Kelsch 2001). These basins have a fast response to torrential rainfall due to the steep slopes, impermeable surfaces, saturated soils and human impacts (deforestation, fires) that cause modifications of natural drainage.

©Domnița Matei, 2009

Fig. 1.2 *Flash flood in the Gârda Seacă river, Jan. 2009*

The subject of this work is the development of a model that simulates the flash flooding

caused by runoff when a torrential rainfall occurs. In time the rainfall exceeds the infiltration rate of soil and runoff appears and starts flowing downstream to the nearest stream. There are several factors influencing the flow path of the water that reduce the effects of runoff: a part of water is lost because of evapotranspiration, another part may be temporarily stored in micro-topographic depressions and some of it may contribute to subsurface flow due to infiltration. The water that is not lost in these processes flows downstream to the nearest receptor like a river, lake, estuary or ocean.

A model will be developed to take the presented factors into account, it will be able to anticipate the quantity of water available for runoff and route it through the catchment in order to estimate the variation of the discharge that appears. The main purpose of this model is to obtain, knowing the landscape characteristics, the antecedent precipitation and the precipitation forecasted for a certain day, the quantity of water which will generate the flash flood and its distribution over time. The result of applying the model is the runoff hydrograph generated by a specific spatially distributed rainfall event that can damage an inhabited mountainous area. Vulnerability maps for flooding will be generated based on the model results.

The model will use a Digital Elevation Model (DEM), soil maps and land use maps in digital format. The precipitation data and trends are used to model the surface runoff. The model will be especially focused on ungauged basins, which is rather typical for the small mountainous catchments of the Apuseni Mountains. The model will therefore have to be realistic with regard to expected availability of data. A validation strategy for the physically based GIS rainfall-runoff model will be worked out by comparison with other published models and available discharge data in some gauged catchments.

The work consists of six main parts corresponding to the following stages of research: determination of the objectives and the actual level of research in the field – the study of possibilities in modeling the phenomenon - the construction of the database – the development of an automated algorithm to discretize the study area – the development and automation of the model to generate the forecasts on flash floods generated by torrential rainfall – the validation of the model using measured discharge from the field.

The first chapter consists of a presentation of the objectives and the geographic location of the study area used to apply the GIS models that I create. The actual stage of the rainfall-runoff modeling research in the world is presented as a starting point for the model implemented in this

work.

Chapter 2 presents some concepts related to spatial and data models, the data model that can be used for hydrologic modeling in GIS and the possibilities of implementing spatial process models in GIS. The last part of the chapter presents some functions that will be used in this study and the approach taken on model implementation.

Chapter 3 presents the database created for the study and the construction of this database. The database includes GIS datasets related to the topography of the terrain and its hydrological characteristics and the methods of obtaining and using these data are presented. The databases available for land use, terrain, soils and climate data in Europe are presented and an algorithm is shown for the processing of each one of these datasets.

Chapter 4 presents the implementation of some algorithms of spatial discretization of the study area as GIS modules. These algorithms are used to create a topological structure of the spatial distribution of the basins that can later be used for runoff routing and discharge modeling.

Chapter5 presents two conceptual models for floods generated by runoff. The first model is based on the time-area method (which implies the determination of the travel time and concentration time for runoff, the determination of discharge generated in different sections of the basins and the generation of the runoff hydrograph by linear routing and accumulation of the discharge towards the outlet). The discharge for each cell is calculated using the SCS Curve Number method for determining the runoff depth and runoff coefficients at cell level. The results include GIS datasets for runoff depth, runoff coefficients, runoff volume, travel time, time-area diagrams for runoff for each basin and finally hydrographs of discharge obtained by integrating the runoff in different sections of the catchment. The second model is based on a runoff routing based on the Saint-Venant equations for shallow water flow and automatic generation of the discharge tables used for plotting the hydrographs. The first model is implemented in the form of Python scripts for each module and can be used to forecast the discharge generated by torrential rainfalls.

Chapter 6 shows some application examples for the model in catchments from the Apuseni Mountains. The applications shown here are in catchments without discharge measurements and gages so they could not be validated from the mathematical perspective. This applications show the possibilities in using the model and the types of results that can be obtained from it.

Chapter 7 shows the validation approach taken in this study and the reasons for taking this

approach. Some general aspects of model validation are presented and the way the validation will be done in this case is shown.

Chapter 8 shows the validation of the model using catchments from the Apuseni Mountains where measured discharge and rainfall data was available and a comparison of the results from this model with results from another model developed in the Faculty of Geography in Cluj-Napoca. The conclusions that arise from the validation results are shown in chapter 9 and a discussion is made on the accuracy of the model and possible further developments.

1.2 Motivation and objectives of the research

Surface runoff is the main factor for causing flash floods in small mountainous catchments. The characteristics of vegetation in the alpine area and the thin soil layer create the conditions for a high runoff coefficient and a low infiltration rate. The changes in land use also affect the response of mountainous areas to high rainfall.

Drobot (2007) presents the main factors that influence the flash floods as the following:
- Natural factors
  - Initial humidity of soil in the catchment
  - Soil erosion represented by rills, gullies or torrents
  - Types of rocks present in the catchment
- Artificial factors
  - The lack of erosion and flow control structures
  - Excessive unplanned deforestation with no regard to forest management practices
  - Bad agricultural practices
  - buildings and deposits created very close to the streams

Among these factors, the erosion and flow control structures are nonexistent in small mountainous catchments and the deforestation or bad agricultural practice effects are becoming more frequent. Therefore, the threat caused by the occurrence of flash floods is getting higher.

The lack of data (for example soil properties or cross-sections of the streams) does not allow the usage of more complex models and the lack of long-term measurements does not allow

the application of stochastic methods. The mountainous areas lack discharge and water level measurements and some areas are hard or impossible to access in order to take these measurements. Therefore, an indirect estimation model is needed that can be used without complex measurement campaigns in the field. This model will mostly use data freely available on the internet, the main data sources being the DEM, soil maps (for the infiltration rate) and land use datasets.

The result of applying the model is the runoff hydrograph generated by a specific spatially distributed rainfall event that can damage an inhabited mountainous area. Vulnerability maps for flooding will be generated based on the model results.

These vulnerability maps should provide authorities with an accurate picture of the evolution of runoff in case of a known or predicted rainfall event and can help minimize losses caused by runoff.

Also, such a model can contribute to development projects in local water resources by the simulation of extreme natural phenomena, such as torrential rains or thunderstorms. The development in these areas can be executed to be able to restrain events that are expected to occur in the future.

Although a large number of rainfall-runoff models exist, not many of these are created for small basins or ungauged basins. The alternative models or specific software dedicated for modeling the rainfall-runoff processes are difficult to calibrate in areas where no historical measurements or gauges exist.

Alternative commercial solutions (e.g. MIKE SHE, SLURP, Hydra) are usually dedicated to more complex simulations in large basins (MIKE SHE) or urban areas (HYDRA). The free and Open Source solutions (IHACRES, Kineros, TOPMODEL) are difficult to calibrate and lack the facilities needed for applying them in areas without historical measurements or gauges.

1.3 Geographic location of the study areas

The Carpathians are a mountainous chain, part of the European central mountainous system. The Carpathians between the Wien Basin (separating it from the Alps) and the Timok passage (separating it from Stara Planina in the Balkan Peninsula) form a 1500 km long and 130 km width arch. They unfold on 6° on the latitude and 10° on the longitude. They cross seven

countries: Austria, Czech Republic, Slovakia, Poland, Hungary, Ukraine, Romania, Serbia (Fig. 1.3). The Romanian Carpathians are part of the eastern mountainous system, well defined by the general direction of the main peaks, by altitude, by massiveness and structure. The direction of their formation is from north-north-west to south-south-west, imposed by the resistance of the Russian Platform. This direction is modified towards the west.
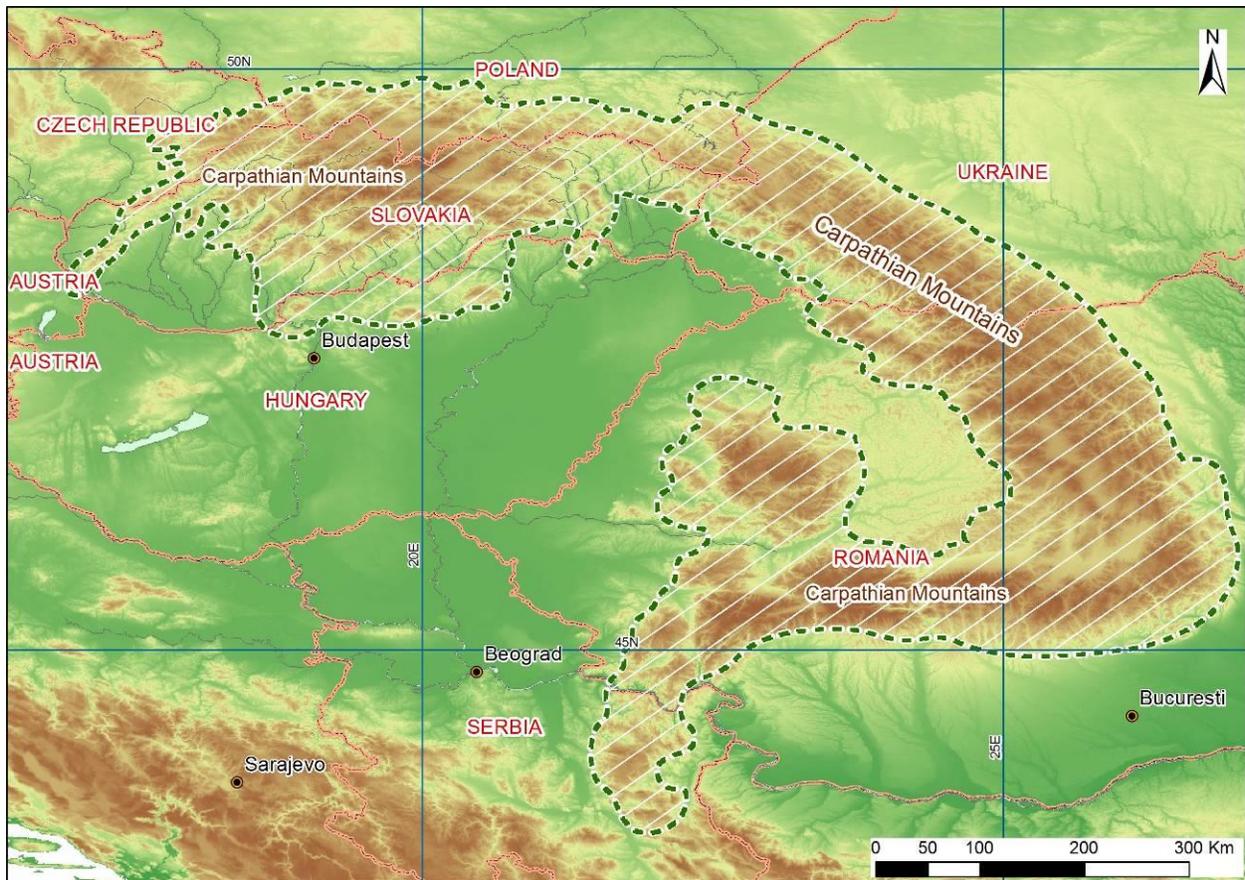
Fig. 1.3 The Carpathian Mountains' extent

Although they are massive mountains, the valleys of some rivers that cross through them (Danube, Jiu, Olt, Mureş, Someş) have separated them into well defined units.

The Romanian Carpathians are divided in three major groups:

– the Eastern Carpathians, between the northern border of Romania and the Prahova Valley

– the Southern Carpathians, between the Prahova Valley in the East and the Timiş-Cerna corridor in the West

14

– the Western Carpathians, between the Danube Valley in the south and the Someș Valley in North

The Western Carpathians comprise a large block of mountain country between the Pannonian Plain and Transylvanian Plateau, roughly 170 km in east–west extent (Șiria in Arad County to Turda in Cluj County) and 150 km north – south (Zalău in Sălaj County to Deva in Hunedoara County) (Fig 1.4). The mountains are of no great height except for a central area comprising the Bihor, Muntele Mare and Vlădeasa Mountains where there are several peaks higher than 1800 m.

The Apuseni Mountains have the largest extent (cc. 10750 km^2), are the highest and most complex mountainous sector of the Western Carpathians and have developed with the shape of a palm. They have been eroded in time, and they are the most fragmented group of the Romanian Carpathians because of the corridors and gulf-depressions that cross them. The altitudes of the Apuseni mountains reach more than 1800m in the central area and between 700-1000m in their northern and southern parts.

The limits of the Apuseni Mountains are:

– Mureș Corridor on the south

– the Transylvanian plateau on the east

– the Western Hills on the west

– Barcău valley on the north

They are composed of the Bihor Mountains in the central part (1849m), the Muntele Mare in the east and the Vlădeasa Mountains in the north. In the eastern prolongation of the Bihor Mountains there is the Găina Mountain and in the north-east lay the Gilău Mountains. In the south-east lay the Metaliferi Mountains and the Trascău Mountains, and in the west the Criș Mountains: Zarand, Codru Moma, Pădurea Craiului and Plopiș (Fig 1.4).

The North-Eastern area of this group of mountains is made up of Gilăului Mountains, part of the Vlădeasa Massif and part of Bihor Massif. The central-North-West of the Apuseni Mountains is dominated by a natural park, the **Apuseni Natural Park** (Fig. 1.4), which covers a massive part of Bihor to the South and Vlădeasa in the North, on the territory of three counties (Cluj 40 %, Bihor 32%, Alba 28%) (Oancea 1987)
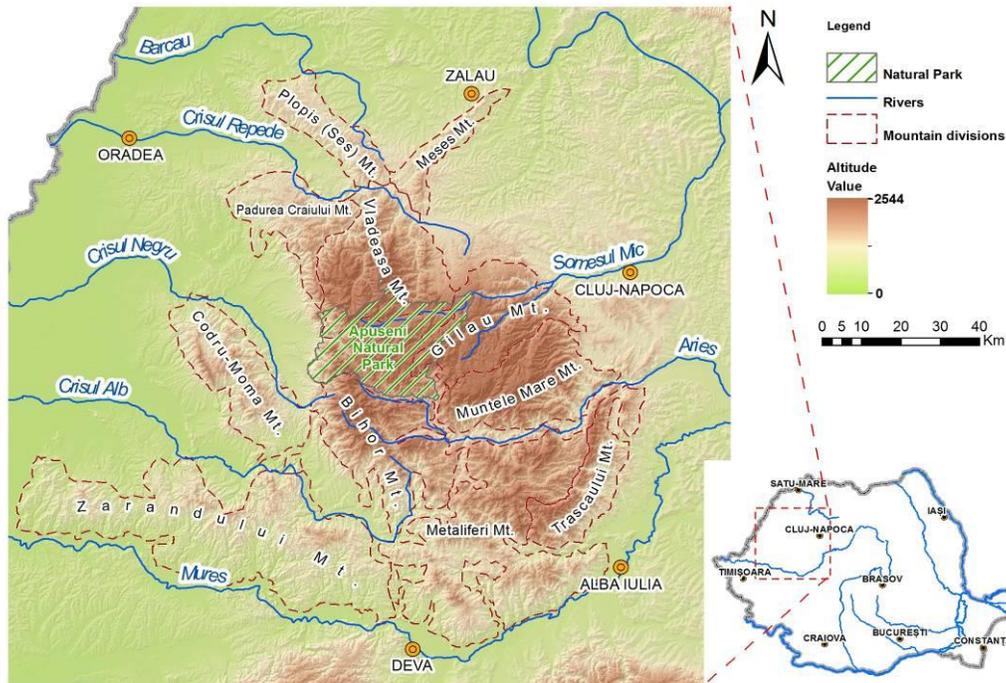
Fig. 1.4 *Apuseni Mountains and the Apuseni Natural Park*

The geological formation of the Apuseni Mountains consists of volcanic rocks (Metaliferi mountains), crystal schists (Bihor, Vlădeasa, Gilău, Găina Mountain), and most of all of limestone, thus the numerous karst phenomena (caves, gorges, hillocks). They are mineralized, rich in metals and nonmetals.

The Bihor Mountains are the main source for the hydrographic system in the area, providing the springs of major rivers, such as the Arieș, the Criș rivers, the Someș and Mureș rivers. Many of the smaller valleys are wild and difficult to reach, therefore hydrologic measurements in the area raise problems. Landslides that block roads and serious floods affected the Apuseni Mountain areas several times in the last years. Extensive deforestations have been made near the high villages in the mountains, causing a rise in the runoff coefficients and the possibility of flooding events.

The Apuseni Natural Park is a protected area and deforestation should be made without a justified reason, but massive deforestations are made in the villages outside the natural park. This study was made focusing on small basins in the Cluj and Alba counties included in this area. Basins in the natural park were also used as study areas because the discharge data availabile in the natural park are more complete. The basins where the applications will be made are presented in the next section.

16

1.4 Characteristics of the basins used in the applications

The applications of the created model will be made on several basins spread through the Apuseni Mountains (Fig. 1.5).
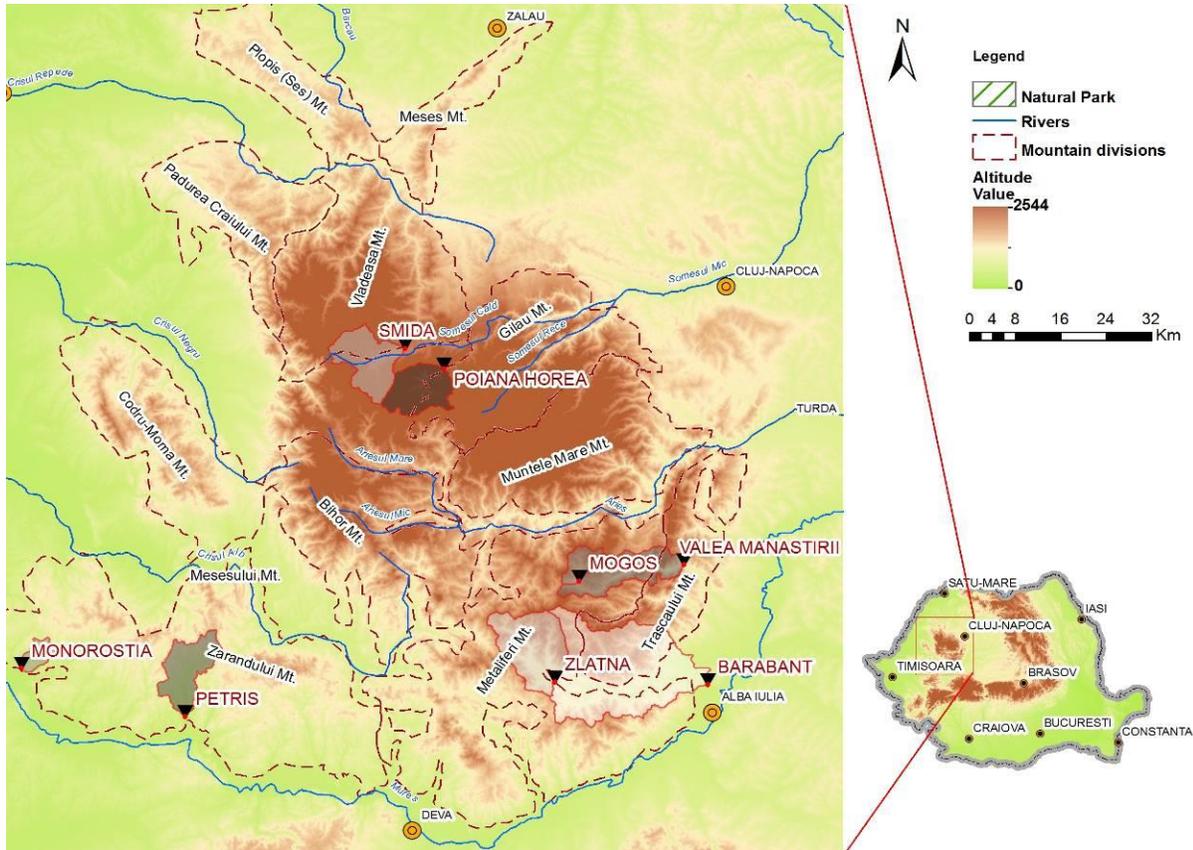
Fig. 1.5 *Study areas in the Apuseni Mountains for the model*

The first two basins are located in the higher altitudes of the mountains, in the National Park. The basin corresponding to the Smida measuring gage is located on the Someșul Cald river in the Vlădeasa Mountains and the basin corresponding to the Poiana Horea measuring gage is located on the Beliș creek in the Gilău Mountains.

The other basins are located in the southern part of the Apuseni Mountains at lower altitudes and correspond to the following measuring gages:

– Mogoș, on the Geoagiu creek

– Valea Mănăstirii, on the Geoagiu creek

– Zlatna, on the Ampoi creek

– Barabanț, on the Ampoi creek

–   Petriş, on the Petriş creek

–   Monorostia, on the Monorostia creek.

The locations of the basins in the Apuseni Mountains can be seen in Fig. 1.5 and their altitudes can be seen in Fig. 1.6. These characteristics will be presented for each study basin later in this chapter.
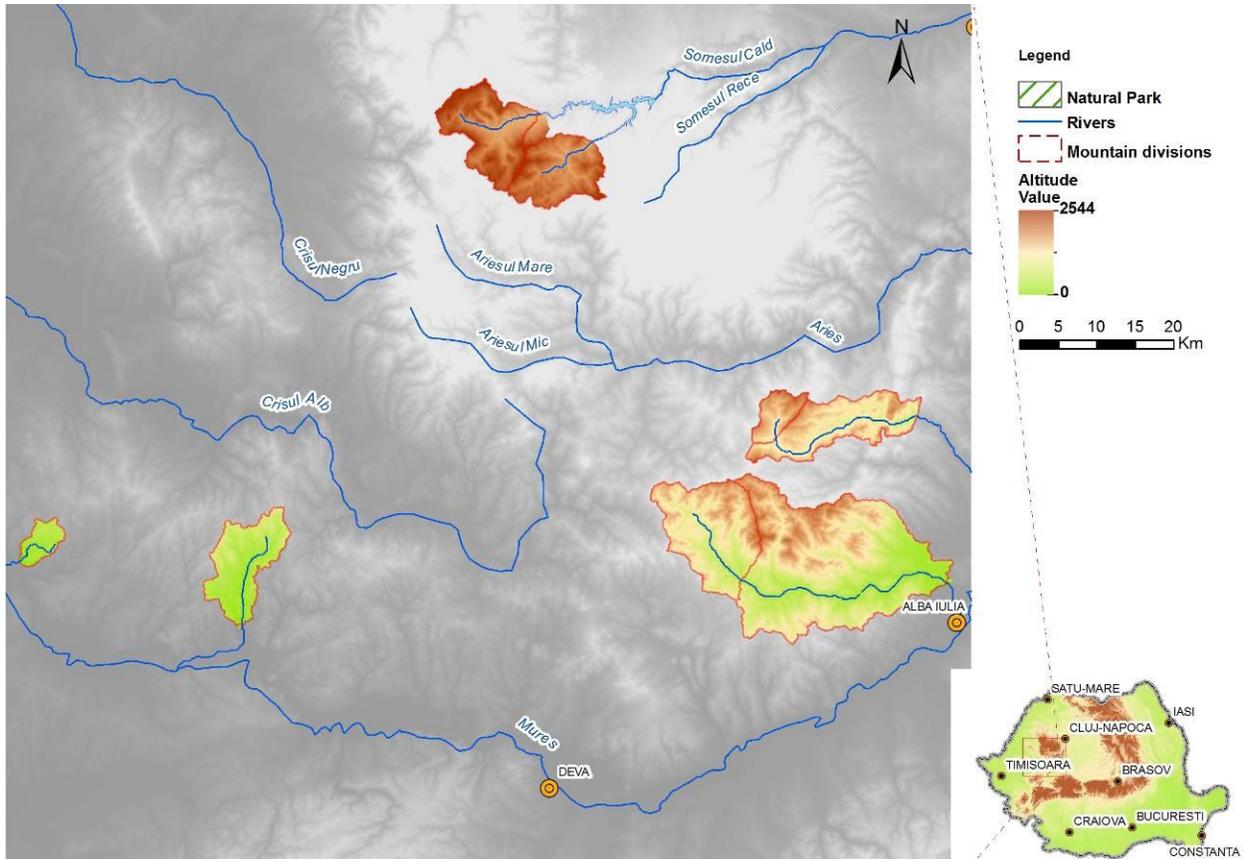
Fig. 1.6 *Study basins and main rivers in the area*

The **Beliş** basin is located in the northern half of the Apuseni Mountains and has an area of about 85 km^2. Its main stream is a right tributary of the Someşul Cald river. The basin extends along the boundary between the Gilău Massif in the east and the Padis plateau in the west. The outlet of the basin corresponds to hydrometric station Poiana Horea, near the river mouth where the river feeds Lake Fântânele-Beliş. (Fig. 1.6)

The watershed delimiting the Beliş basin upstream from Poiana Horea crosses the following peaks: Chicera Negrului (1496,8 m), Colţăul Vârfului (1652,6 m), Vf. Pietroasa (1564

m), Vf. Clujului (1399,3 m), Vf. Roşu (1568,5 m), Vf. Ţiclău (1566,8 m), Vf. Drăgoiasa (1538,8 m), Vf. Scoruşeţu (1407,5 m), Piatra Fulgerată (1428,5 m) etc. Two more important peaks in terms of altitude are found inside the basin: Vf. Sturu (1475,4 m), on the watershed between Pietroasa and Beliş streams: Vf. Dealu Calului (1452 m), on the watershed between Ţiclău and Beliş streams.

Neighboring basins are small right tributaries of the Someşul Cald River (Bătrâna Creek, Barni Creek, Şimonul I, II, III, Giurcuţa Valley - north and northwest; Răcătău Basin – East; Gârda Valley and Albac Valley - south The basin includes only one locality (Poiana Horea) developed along the Beliş and Apa Caldă streams. Flow monitoring is carried out only in the Poiana Horea hydrometric station.

Someşul Cald basin upstream from the the **Smida** hydrometric station is located in the upper basin of Someşul Mic and has an area of about 110 km^2.

The Someşul Cald Basin upstream of the Smida hydrometric station is situated in the central-northern part of the Apuseni Mountains. Its left hillslope is part of the Vlădeasa Massif, and the right one is part of the Padiş Plateau. The following peaks are situated on its watershed: Miclau (1640m), Drăgoiasa (1538m), Roşu (1568m), Biserica Moţului Peak (1466m), Brăiesei Peak (1692m), Britei Peak (1759m). The basins in vicinity are Draganului Basin (in the north-west), Săcuieului Basin (in the north-east), Arieşului Basin (in the west and south), and Belisului Basin (south-east). The main hydrographical course is Somesul Cald, which springs at 1365m altitude (under the Piatra Arsă Peak, 1550m, from a limestone area of a triassic-jurassic age) and flows for 15 km to the Smida hydrometric station. The karst phenomena are present from the springing point. The river disappears through a ponor with a cave and it reappears in the Rădeasa cave; from here it crosses Cetatea Rădesei along 250m, through a tunnel cave. After the junction with Pârăul Ars (its first left affluent from the Tunelul Mic cave) Somesul Cald enters a sector of gorges with abrupt limestone walls and altitude difference of over 100m (Ujvari I., 1972). After exiting this sector of gorges Somesul Cald gets more water from three other small affluents: Alun, Alunul Mic, Ponor. The most important affluents from the right side are: Bătrâna and Izbucul (with its affluent Călineasa).

The catchment area is symmetrical, because Somesul Cald collects substantial quantites of water both from its left affluents of the southern hillslope of Vlădeasa Mountains and from its right affluents descending from the higher areas of the Padiş Plateau.

The hydrologic regime is of Carpathic Transylvanian type, characterized by reduced winter runoff and over 20% higher runoff during summer.

Studies conducted by Ujvari I., (1972) on the Someș Basin show the following quantitative characteristics of the annual hydrological survey: mean annual precipitations are around 1075mm; mean discharge more than 600mm; evapotranspiration around 440mm; groundwater flow , close to 200mm; and the values of the mean runoff coefficients are around 0,5-0,7.

Other two study basins are the Geoagiu basin upstream of the Mogos hydrometric station and the Geoagiu basin upstream of the Valea Mănăstirii hydrographical station.

The Geoagiu basin upstream from **Mogoș** has an area of about 29 km^2. The basin extends on the boundary between the Trascău and the Metaliferi mountains and its altitudes range from 775 to 1436 m.

The **Valea Mănăstirii** hydrometric station is located on the Geoagiu brook downstream from Mogoș. The basin has an altitude range from 450 to 1436 m and an area of about 139 km^2.

The main stream is the Geoagiu brook, a right affluent of Mureș with the confluence near Teiuș. The watersheds of the two subbasins that separate them from the Galda Basin (in the south) and Arieș (in the west and north) cross the following mountain peaks: Geamana (1366m), Vf. Căpăţâna (1173 m), Piatra Crisnicului (1349 m), Vf. Poieniţa (1436 m), Dl. Goşa (1052 m), Vf. Pleşii (1250 m), Piatra din Chei (1167 m), Vf. Prisecii (1150 m), Vf. Măgulicea (1128 m), Negrileasa Mogoşului (1364 m). The territory of the two basins is characterized by a high density of rural localities, placed mostly on the bottom of the valleys. These are a few of them: Valea Mănăstirii, Cheia, Onceşti, Măcăreşti, Bărleşti-Cătun, Cojocani, Valea Barnii, Bărleşti, Valea Tupilor, Bârzogeni, Tomeşti, Buteşti ş.a. The runoff regime belongs to the western Carpathic type and Carpathic Transylvanian, after exiting the mountainous area.

Another study basin in the lower ranges of the mountains is the Ampoi basin. The total surface of the Ampoi Basin is around 576km, and the watershed is defined by the following altimetric benchmarks: Vf. Petriceaua (1144 m), Vf. Brădoi (1236 m), Vf. Vilcoi (1348 m) Buza Măgurii (1264 m), Vf. Muncelului (1282 m), Vf. Piatra Arsă (1306 m), Vf. Dealurilor (843 m), Dl. Mamutului (765 m), Vf. Gorganu (877 m), Vf. Namasoaia (934 m), Vf. Breaza (1121 m), Vf. Grohas (1120 m), Dl. Brătianului (1032 m). The main hydrographic course that drains this basin is the Ampoi, right affluent of Mures, of which juncture is in the Alba Iulia Depression, near by Alba Iulia city. The length of the river is of 60km, with springs situated near by Abrud, at the foot

of the Dosu (1040m) and Petriceaua (1144m) hills, and it crosses the mining region of Zlatna (*Ujvari I., 1972*).

The basin is asymmetrical, developed more on the left side. As a consequence, the most important affluents of Ampoi are on the left side, on the direction of the last prolongations of the karst peaks of Trascau Mountains. (*Ujvari I., 1972*). These affluents are Fenes, Ampoita and Ighiu, Ighiu with gorges on the springs karst areas (in this area there is also the karst lake of Ighiel). The Ampoi river is also a water supply for the city of Alba Iulia.

The rural settlements on the Ampoi basin are highly concentrated on the lenght of the Ampoi valley. Some of these are: Budeni, Boteşti, Izvoru Ampoiului, Zlatna, Podu lui Paul, Suseni, Pătrângeni, Galaţi, Presaca Ampoiului, Meteş, Tăuţi.

The hydrological survey conducted by *Ujvari I. (1972)* on Romanian rivers show the following mean values of hydrological assessment in the Ampoi basin: the mean annual precipitation quantities: 790mm; the mean discharge: 283mm; evapotranspiration: 502mm; groundwater flow: 70mm. The mean annual runoff coefficient that can be deduced out of these values is around 0,35. The hydrological characteristics presented above have suffered some changes during 40 years of hydrological monitoring of the basin but they still remain important bench-marks in the hydrological research of a hydrographic basin. The runoff regime belongs to the western carpathian type and carpathic transylvanian type, after exiting the mountainous area, until the juncture with the Mures river. The basins studied in this paper are upstream the hydrometric station Zlatna (its area is around 104km^2) and the Barabant hydrometric station (its area is around 528km^2).

The Petriş basin is situated on the south of the Apuseni Mountains and it springs at more that 700m altitude, under the Omeag peak (749m) and is a right affluent of Mures river. Their junction is downstream the Pojoga settlement. The work focuses only on the basin area that is upstream the Petriş hydrometric station (around 92 km^2). The watershed of this basin delimits it from other basins, such as Crisul Alb or Săvârşin and intersects the following peaks: Vf. Omeag (749 m), Dl. Pietroasa (708 m), Vf. Burlan (501 m), Măgura (512 m), Vf. Dăescu (711 m), Dl. Pleşu (726 m). The shape of the basin is asymmetrical and it is more extended on its right side by small prolonged subbasins of which temporary water courses junction with Petriş near Roșia Nouă, Corbești, Petriș

The Monorostia basin is also part of the southern area of the Apuseni Mountains, and it is

the smallest basin among those selected for applications (20km^2). The main water course is a right affluent of Mures with which it junctions at its exit from the Monorostia settlement. The watershed that delimits this basin from the others intersects the following altitudes: Vf. Moghilă (391 m), Blidu Tomii (517 m), Capu Vucinei (307 m). This basin is also more extended on its right hillslope.

The altitudes and main points on the watersheds of all basins can be seen in Fig. 1.7
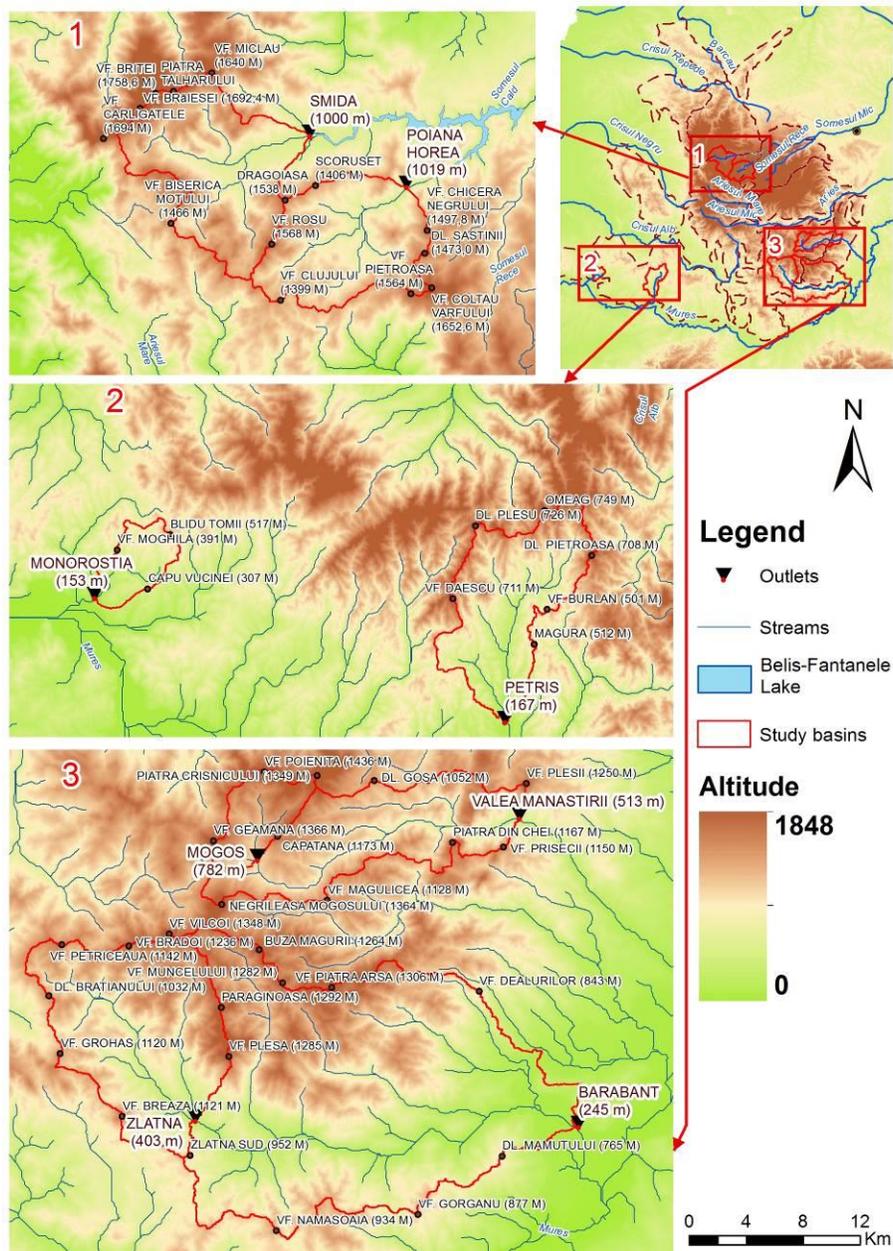
***Fig. 1.7** Catchment altitudes and main points on the study watersheds*
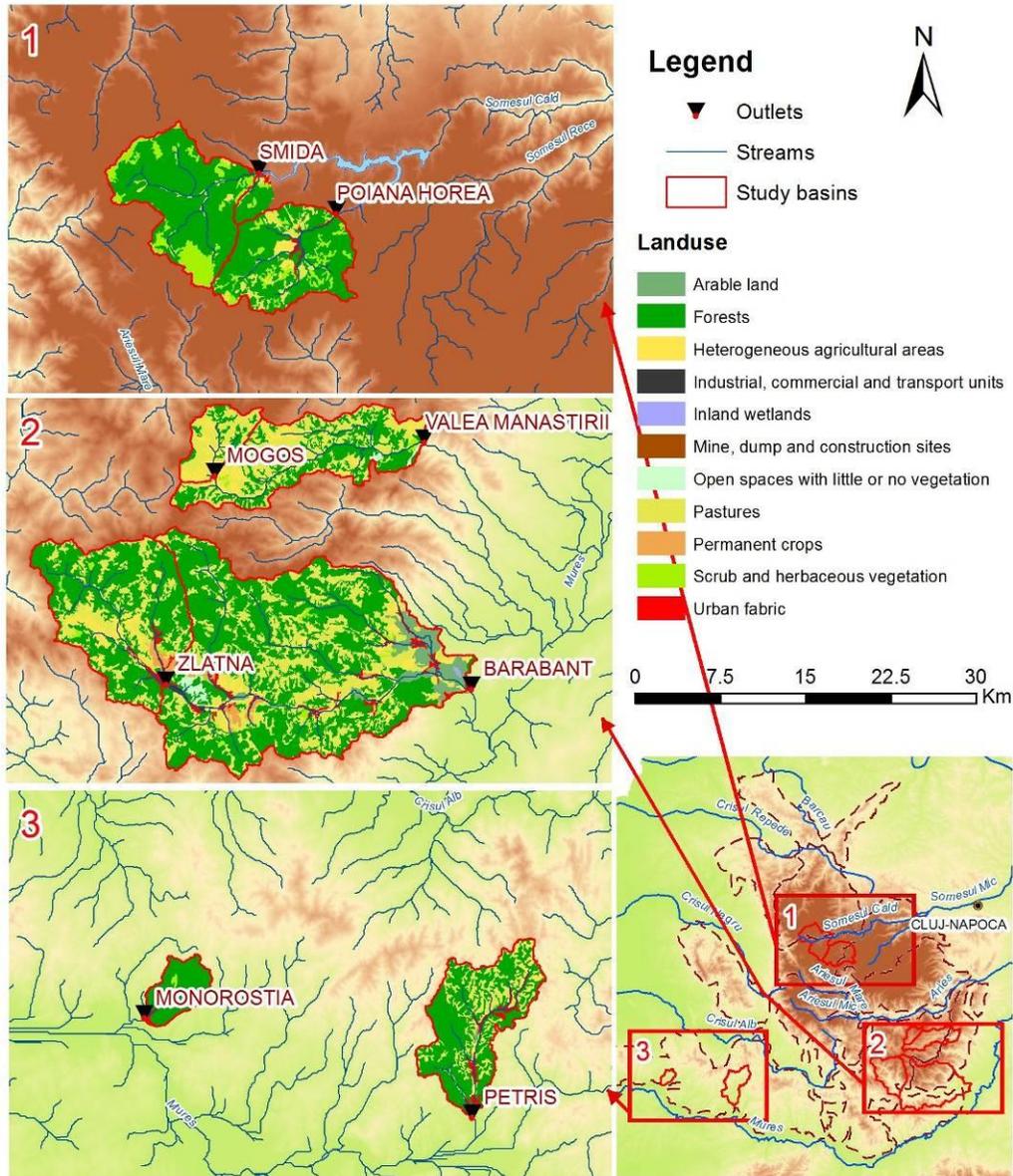
Fig. 1.8 *Land use of the study basins and main rivers in the area (CORINE land cover database)*

The highest values of the terrain slope appear in basins belonging to karst areas in the eastern part of the Apuseni mountains, and also in the Somesul Cald basin. In these basins the slopes reach 50-70% values. This is especially the case of gorges sectors in these basins, characterized by high values of fragmentation depth. The smaller slopes correspond to valley stream areas, where the process of accumulation of sediments resulted from hillslope runoff is dominant. Smaller values of slope are also present in interbasin areas, which appear in the relief as flat surfaces (see the Somesul Cald and Belis basins, which correspond mostly to the Padiș Plateau subunit).
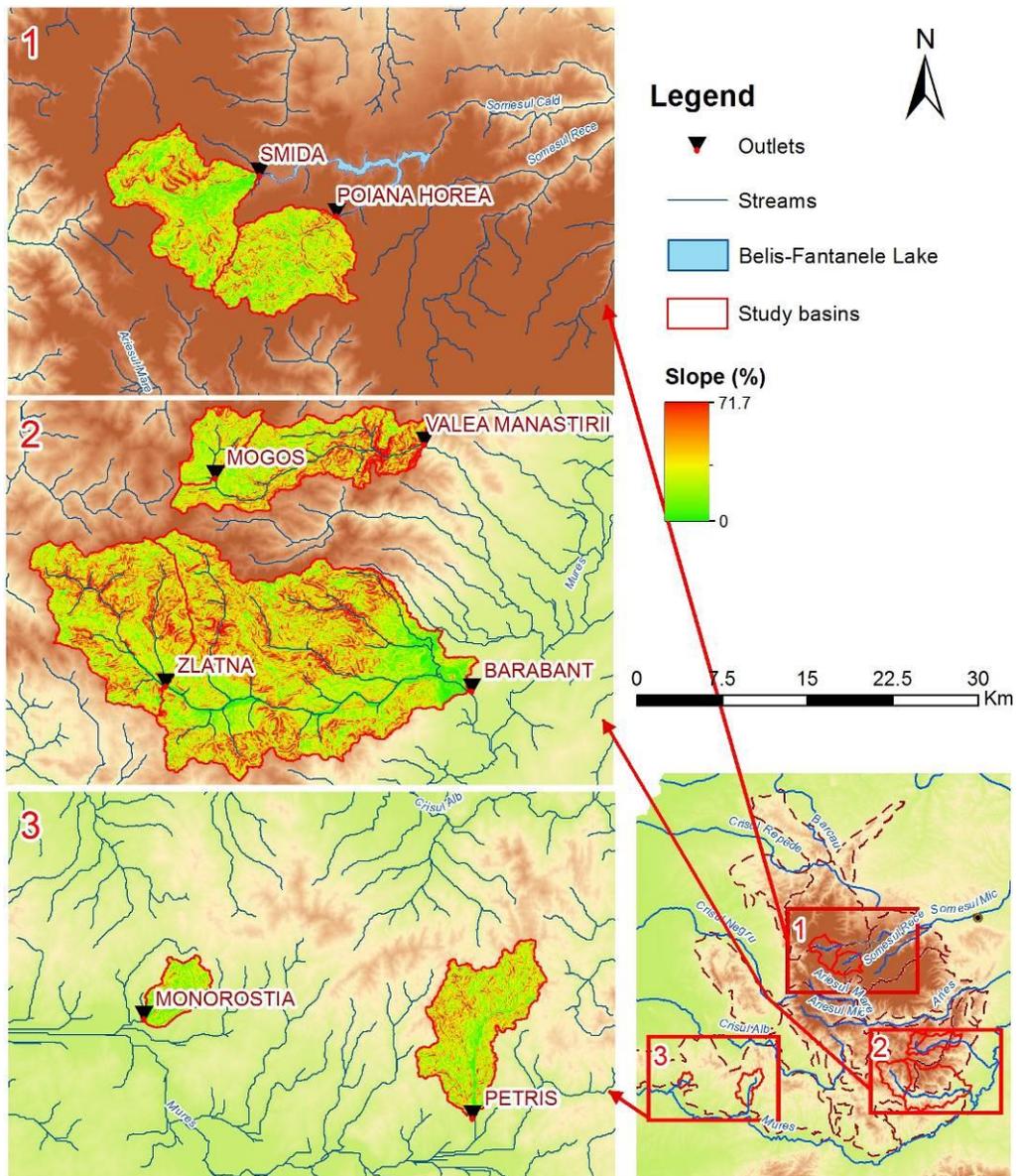
Fig. 1.9 *Slopes of the study basins*

The hydrologic soil groups of the basins according to the USDA-SCS classification can be seen in Fig. 1.10. The USDA-SCS soil classification is presented in detail in section 3.3.

The soil analysis according to its infiltration capacity shows that there is a dominance of the B soil hydrological group, characterized by clay or sandy clay soil, with an average capacity of infiltration. These kinds of soils are present in an up to 100% proportion in Valea Manastirii basin, Monorostia and Petris basins.

The C soil hydrologic soil group, characteristic to clay soils with smaller capacity of infiltration, is dominant in the Beliş basin and Someşul Cald basin. The D hydrologic soil group, specific to loamy clay and loamy soils, with very low capacity of infiltration, is dominant in the

Ampoi basin, but is also present in the south-eastern extremity of the Someșul Cald and Beliș basins.

These latter two hydrologic soil groups, corroborated with lands with a small coefficient of forestation are the most favorable for surface runoff processes, as they contribute to increasing the runoff speed, decreasing the travel time and concentration time, increasing the volume of the runoff.
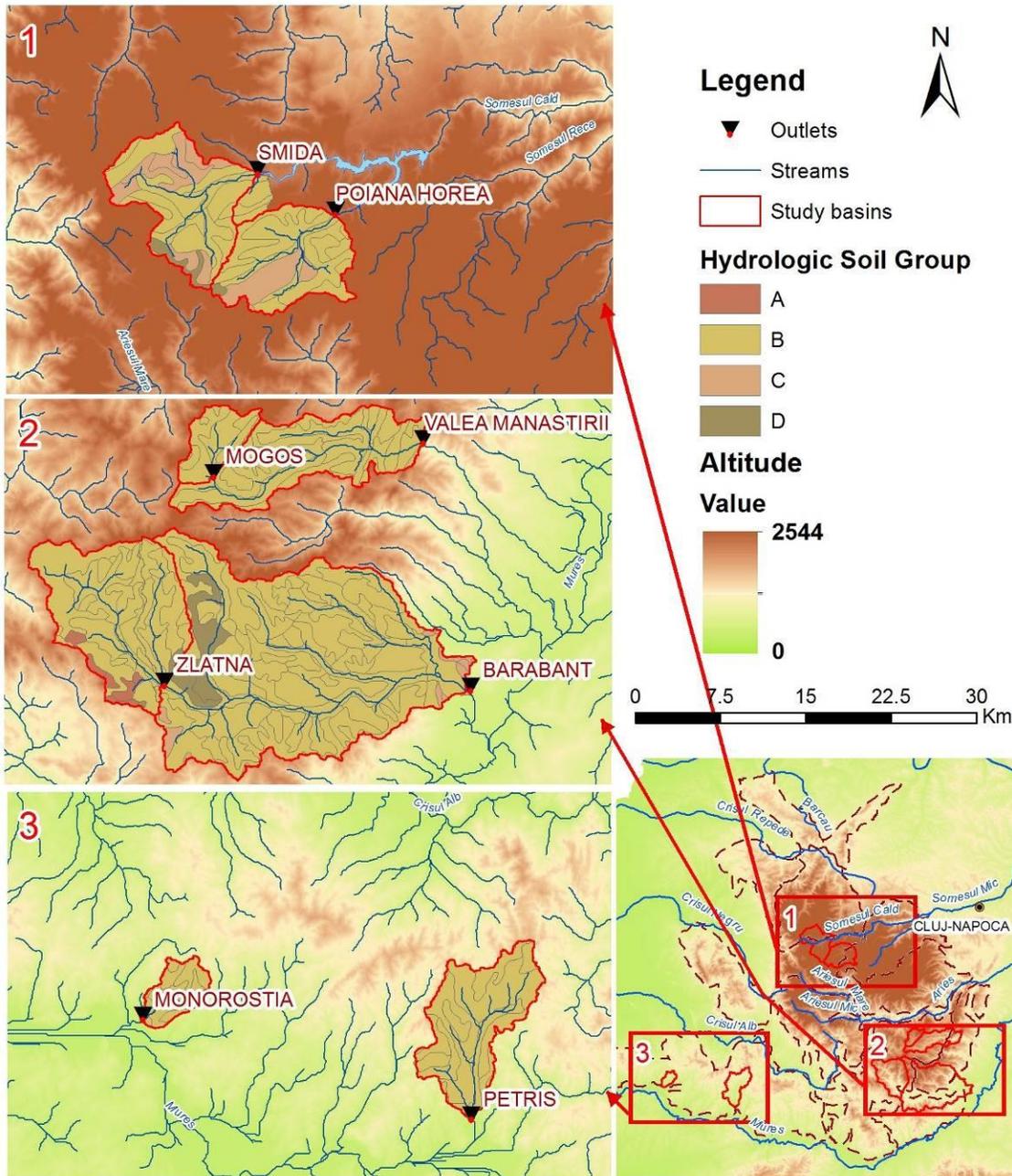
Fig. 1.10 *Soil groups of the study basins*

1.5 Evolution and stage of research in rainfall-runoff modeling using GIS

The rational method

Probably the first try to create a model that estimates maximum discharge was made over 150 years ago. Hydrologists of that time did not have modern calculation methods and their capacity of solving mathematical problems was limited. The Irish engineer Thomas James Mulvaney (1822-1892) published a simple equation to determine the maximum discharge for a quantity of rainfall in a catchment. The equation (1.1) published by Mulvaney in 1851 is the following:

$$Q = C \cdot A \cdot R \tag{1.1}$$

The equation only predicts the hydrograph peak, not the whole hydrograph. The maximum discharge Q is predicted by the equation for a rainfall with the duration at least equal to the concentration time of the catchment.

The input variables are:

- Catchment Area A
- Maximum catchment average rainfall intensity R
- An empirical coefficient C. The coefficient is the proportion of rainfall that contributes to runoff.

The method shows the way maximum discharge is presumed to increase with the catchment area and rainfall intensity. The calculation of this maximum discharge is made in a rational way, hence the name of the method. The rational method is a popular, easy to use technique for estimating peak flow in any small drainage basin having mixed land use. It generally should not be used in basins larger than 1 square mile.

The empirical runoff coefficient C is used to show that not all the water falling to the ground as rainfall contributes to discharge. The coefficient is the proportion of rainfall that contributes to runoff. Usually tables are used with values of this coefficient corresponding to different land uses. In basins having a significant nonhomogeneity of land use, an average coefficient can easily be determined by multiplying the percentage of each land use in the basin by its appropriate coefficient and determining a mean runoff coefficient for the catchment.

The rational method was adapted and used in Romania by different authors

One of the applications of the method presented by Magyari (2008) was formulated as this equation (1.2):

$$Q = 16.67 \cdot \alpha \cdot i \cdot F \qquad (1.2)$$

where,

Q – maximum discharge [m³/s]

α – adimensional runoff coefficient [-]

i – mean rainfall intensity (rainfall duration is at least equal to the concentration time) [mm/min]

F – catchment area [km²]

The constant 16.67 appears from the transformation of mm/min and km² in m³/s in the following way (1.3):

$$\frac{mm}{min} \cdot km^2 = \frac{10^{-3} m}{60 s} \cdot 10^6 m^2 = \frac{10^3}{60} \cdot \frac{m^3}{s} = 16.67 \cdot \frac{m^3}{s} \qquad (1.3)$$

Although this method does not offer the possibility of obtaining an entire hydrograph but just its peak (maximum discharge) it is enough for a lot of hydrological engineers that design bridges or dams that can sustain the maximum discharge.

The rational method was adapted for GIS use and used in Romania by different authors (Șerban, Diaconu, 1995; Păcurar, 2006; Bilașco Șt., 2008; Magyari Saska Zs. 2008).

Methods based on time and area

The first try to create a spatially distributed hydrologic model was made by Ross in 1921. His idea was simple: split the catchment into zones depending on the time needed for water to flow to the catchment outlet (isochrones). Considering a time interval t, the area of the catchment is split in the following way: the first zone is the one from where the water flows in the interval of time [0,t], the second zone is the one from where the water flows in the interval [t,2t] and so on. If the quantity of runoff produced in each of these intervals can be calculated, the runoff hydrograph can be obtained by integrating the runoff from each of these intervals in the outlet according to

flow time.

Using this concept he hydrologist obtains a **time-area diagram (TAD)** (the area from which water flows in each time step) (**Fig. 1.11**). The time-area diagram represents the lag in time needed for water from each area of the catchment to reach the outlet.
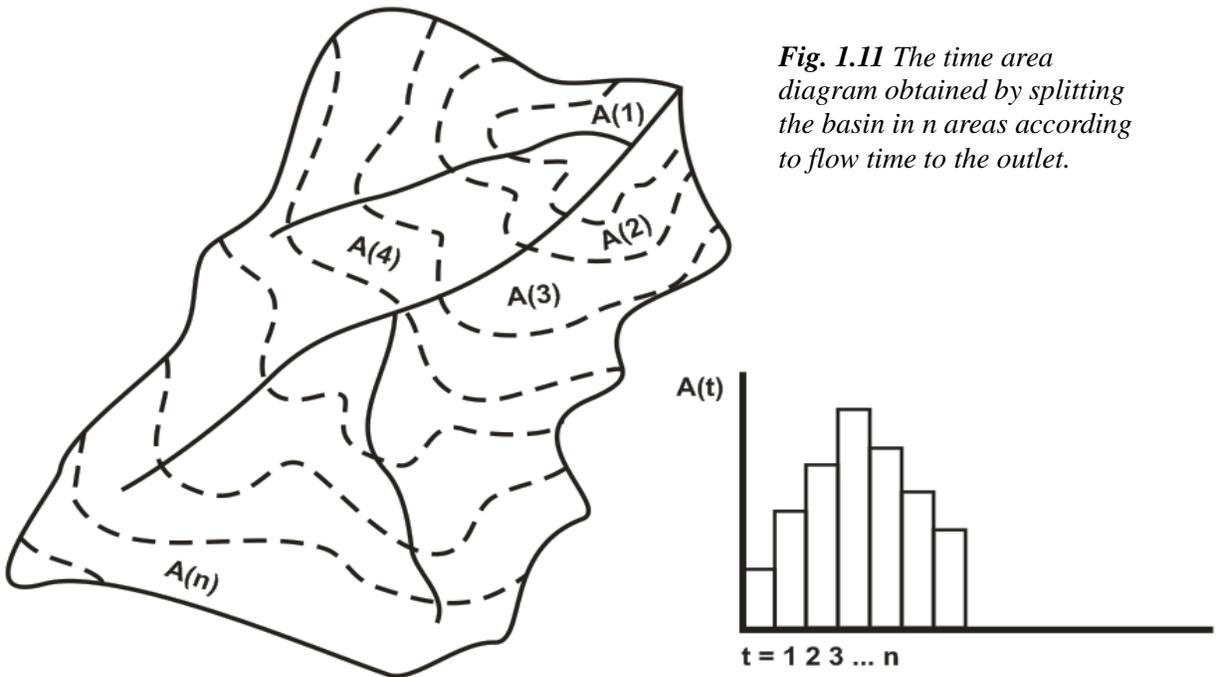
Fig. 1.11 *The time area diagram obtained by splitting the basin in n areas according to flow time to the outlet.*

The time-area concept was used in the USA by Zoch (1934) and Clark (1945), in England by Richards (1944). The idea is the base of some newer distributed hydrologic models. Kull and Feldman (1998) adapted Clark's method for usage with radar data obtained from the NEXRAD system.

Clark's method was also adapted and implemented in GIS by different researchers (ex: Chandramohan et al., Sarangi et al., Changbin et al.). Also, concepts from this method were used in different applications for runoff modeling (ex: Domniţa et al. 2009).

The main element of the time-area method is the TAD that represents the relationship between runoff travel time and the portion of the watershed area that contributes to runoff during that time period (McCuen, 1989). The travel time required for a rain drop falling on any location in a sub-area is the same as that for any other drop falling on the same sub-area. An isochrone is a contour line connecting all points that share the same travel time. Due to their construction, isochrones can not cross one another, can not close, and can only originate and terminate at the watershed boundaries (Dooge, 1959).

The fundamental assumption of the traditional time-area method is that it accounts for translation and does not account for storage effects. Neglecting the storage effect results in lack of attenuation and therefore predicts higher peak flow rates (Ajward, 1996). However, neglecting storage effect can be overcome by adjusting the hydrograph resulting from the time-area method by routing it through a linear reservoir with the appropriate storage coefficient (Bedient and Huber, 1992).

If the rainfall intensity is high and the runoff speed can be estimated with good accuracy, the time-area method is an easy way to obtain a flood hydrograph using little input data. The model proposed in this study will use a linear routing method based on time-area concepts similar to the ones from method.

Many modules for determination of runoff speed through a catchment exist as components of GIS products nowadays. Even if they differ in principles and implementation, the user has many choices of algorithms in calculating runoff speed and flow routing.

The development of the Unit Hydrograph

The time-area concept is a good method that has been used in hydrologic models since its creation, but it has a big problem: the accurate determination of the contributing areas for each time step, because the overland and groundwater flow speed are not that easy to determine with the needed precision for each flow path.

Sherman tried to avoid this problem in 1932 by trying to present the lag time needed for water to reach the catchment outlet as a time distribution without a direct link to catchment areas. His idea was the use of mathematical functions to obtain the answer of the basin to a single unit of effective rainfall uniform over the catchment. He called this a hydrograph, but the name changed to **Unit Hydrograph** after that.

The unit hydrograph is the runoff hydrograph generated by an excess precipitation of one unit (1 mm) that flows towards the catchment outlet. The hydrograph can be determined for a river, a catchment or in a certain point of a catchment with regard to the excess rainfall duration and distribution over the catchment.

The unit hydrograph is probably the most common hydrologic technique and is very simple to understand and apply. Practically, the hydrograph is just a transfer function for the rainfall to reach the outlet of a catchment without considering the spatially distributed characteristics of a

catchment.

The response of the catchment to rainfall is considered linear in case of the unit hydrograph, and therefore the principle of superposition can be applied. Watershed response to a given amount of excess precipitation is just a multiplier of the unit hydrograph. Two units of excess rainfall will produce a discharge two times higher than one unit, with the same temporal distribution. A number of hydrographs calculated for successive time intervals can also be distributed in time with the corresponding delay and then summed to obtain a total hydrograph for a catchment.

Usually a unit hydrograph is derived from historical rainfall and runoff data. The volume of water produced by the storm (total area under the hydrograph curve) divided by the area of the watershed equals depth of excess precipitation. The ordinates of the storm hydrograph are divided by this depth to obtain the unit hydrograph.

If no historical rainfall and runoff data is available the hydrograph can not be obtained in the way presented above. The other methods for the determination of the unit hydrograph that appeared are called synthetic unit hydrographs.

The two widely known methods for determination of the unit hydrographs are the Snyder method (1938) and the USDA SCS (United States Department of Agriculture Soil Conservation Service) method.

Snyder's synthetic unit hydrograph is based on the creation of the hydrograph from characteristics of the rainfall and the catchment.

Using the equations proposed by Snyder five characteristics of the Unit Hydrograph can be obtained (Chow et al. 1988): the peak discharge per unit of watershed area, q_{pR}, the basin lag, t_{lR}, the base time, t_b, and the widths, W (in time units) of the unit hydrograph at 50 and 75 percent of the peak discharge. (Fig. 1.12).

Fig. 1.12 *Snyder's synthethic unit hydrograph*
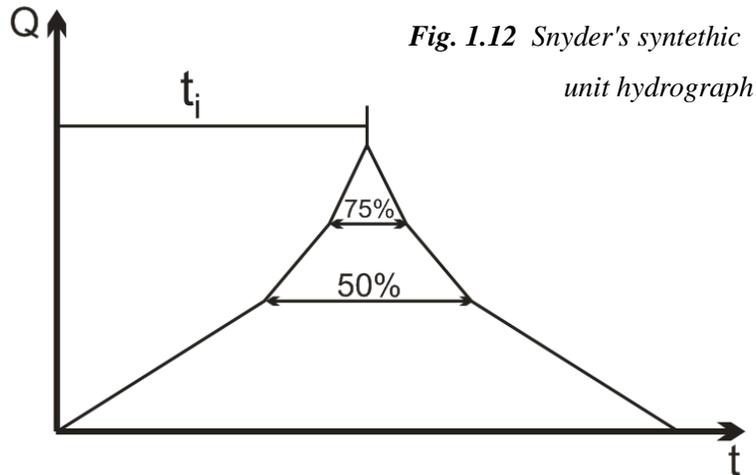
The parameters for the Snyder hydrograph calculation are:

- A peak flow factor (Cp) coefficient derived from gauged watersheds in the same region that represents variations in watershed slopes and storage characteristics
- A lag factor (Ct) coefficient derived from gauged watersheds in the area, and represents the effects of retention and storage
- The length of flow to the outlet from the most distant point in the basin
- The main channel length from the outlet to a point closest to the centroid of the basin

Another method to determine a hydrograph is the method proposed by SCS (Schwab et. al 1993). This method is used to create an adimensional hydrograph with ordinate values expressed in a ratio Q/Q_P (flow/peak flow) and containing the values of the ratio t/t_p (time/time to peak) on the abscissa. The dimensionless unit hydrograph can be used later to determine a watershed specific unit hydrograph knowing some characteristics of the watershed.

The data needed to apply the method are: the area of the catchment A, the time of concentration T_c and the duration of the unit excess rainfall D. The concentration time can be calculated using different formulas like the Kirpich/Ramser formula.

Other unit hydrograph calculation methods also exist along the two well known methods presented here.

Even if the unit hydrograph is a widely used technique, some problems arise in determining the quantity of rainfall that forms runoff (effective rainfall). This is more or less the same problem that appears in the rational method when the runoff coefficient is determined and requires a good knowledge of the hydrological processes in the catchment.

The unit hydrograph does not need the spatially distributed characteristics of the catchment but a large number of models that calculate maximum discharge are still based on this technique.

First digital computer models: the Stanford model

In the 1960s along with the availability and rise in popularity of digital computers the idea of implementing hydrologic models on these computers for a rise of performance in calculations. The hydrologic models that appeared in this period could only use a very limited storage capacity and little memory so they could not become very complex. The hydrologic models that appeared with the development of computers had a very similar structure at the start: a series of reservoirs representing different processes in the basin and mathematical functions representing the flux between these storage units.

One of the first models of this kind was developed at Stanford University by Norman Crawford and Ray Linsey and published as Crawford's PhD thesis in 1962. The model has evolved in the Hydrocomp simulation program and was widely used after its initial development. The idea of the model was a continuous representation of most hydrologic processes in a catchment (infiltration, soil moisture, evapotranspiration, runoff) and their inclusion in a single structure.

The Stanford model had five reservoirs: surface water, soil water, interception, active groundwater, passive groundwater (Şerban 1995) and a maximum of 35 parameters from which most could be determined based on the basin characteristics and only a few of them needed calibration.

Hydrologic models similar to the Stanford model are called ESMA (explicit soil moisture accounting) developed very fast and in a very large number. The majority of these models could give a reasonable runoff prediction after the calibration of a number of parameters. There was also the possibility of adding more components or parameters to these models to introduce different processes. Dawdy and O'Donnell (1965) created a generic structure for a this type of model that was similar to most of the models of this type existent at that time. (Fig. 1.13).

In the decade after 1962 the hydrologic research in Stanford was very active and consisted in a number of PhD theses published on subjects like soil erosion, sediment transport, snow accumulation and melt, water quality studies, stochastic generation of design storms, optimization of model parameters, full equations for routing using finite difference methods, and infiltration analysis. (Crawford and Burges 2004)
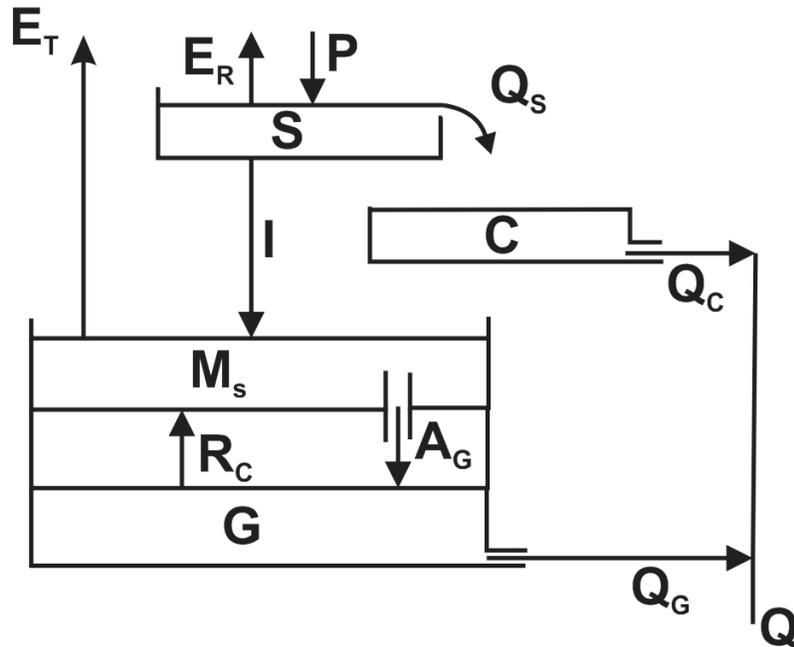
Fig. 1.13 *The structure of reservoir models (Dawdy and O'Donell)*

Where:

Processes:
R - Rainfall
E_T - Evapotranspiration
E_R - Evaporation from the soil
I - Infiltration
A_B – Alimentation of groundwater reservoir
R_C – Return of water through capillarity

Reservoirs:
S – Surface water
C – Channel water
(Surface water routing)
M_S – Soil Moisture
G - Groundwater

Outflow:
Q_S - Surface runoff
Q_C - Outflow from Surface water routing
Q_G - Outflow from groundwater
Q - Total outflow

The development of models similar to the Stanford model was not interrupted and there are some actual models based on this model. Singh (1993) presented some examples of these models including the HSPF. SSARR and Sacramento models (USA). the HBV model (Sweden). the Tank model (Japan), the UBC model (Canada), and the RORB model (Australia). Some similarities between the early versions of the Xinanjiang (Variable Infiltration Capacity) model and the Stanford watershed model also exist.

Distributed models based on hydrologic processes

A more recent approach in hydrologic modeling is the creation of models based on

equations that describe different surface and underground hydrologic processes with regard to their spatial distribution. An example of such a model was presented in 1969 by Freeze and Harlan in a paper.

The equations they presented are based on known equations for different surface and subsurface hydrologic processes, and their work integrated all of these equations in a single hydrologic model in case of common conditions. The research of Freeze and Harlan is still the basis of spatially distributed rainfall-runoff models all over the world. The partial differential equations that they used could be solved by numerical methods but the equations they used in the model required, in all cases, some simplifying assumptions.(Beven 2003)

The first applications in the field were made in research catchments by Freeze and Stephenson in 1974 (Reynolds Creek, Idaho). The results were not very accurate but the two researchers could explain the reasons why the complexity of the catchment did not allow for a good estimation of all flow parameters. This is one of the first applied examples of a validation for a hydrologic model

Distributed models split the representation of the terrain in a mesh of points where the results are calculated. Each of these points started with its own parameter values and calculations were made at point level in the catchment representation. Today's GIS techniques use a similar type of representation for the raster layers, if we consider every cell to be a point in this mesh.

The problems that appeared with Freeze and Harlan's model include problems in representation of the catchment and the other parameters, problems in obtaining of the spatially distributed details on soil types or vegetation and problems in model calibration.

The distributed models have developed a lot in the recent period because of the availability of increased computing power and the ability to make a large amount of calculations on the digital model using GIS. Geographic data obtained through remote sensing using different systems is also a very important step in distributed hydrologic modeling because it eliminates the need to make field measurements for every application of the model. Another reason for the recent popularity of distributed models is the ability to offer distributed results, very important for some kinds of applications like flooded areas or pollutant transport.

Local changes in landuse (deforestation, new buildings or urban infrastructure) can also be taken into account using a distributed hydrologic model and the effects of these changes can be predicted in a clear and quite accurate spatial context.

Recent distributed hydrologic models include the SHE model (Système Hydrologique Européen), created in 1977 from the collaboration between three institutions: Institute of Hydrology (UK), the SOGREAH company in France and DHI (Danish Hydraulic Institute). The development of this model continued independent in these institutions and the actual model proposed by DHI is called MIKE SHE. The Institute of Hydrology in UK developed the IHDM (Institute of Hydrology Distributed Model) (Calver and Wood 1995), and a research team in Australia created the THALES model (Grayson et al. 1995) and the dynamic CSIRO TOPOG model (Vertessy et al., 1993; Zhang et al 1999). These are only some of the distributed models existing at this time and a lot others exist. The differences between distributed models at this time lie mainly in the discretization of the catchment or the simplifications in equations representing the processes. All these models have things in common with the original blueprint published by Freze and Harlan.

Distributed models based on distribution functions

The spatially distributed models presented before are quite inefficient in terms of computational power and data needs (number of parameters). The mathematical equations on which they are based are still a simplified representation of how the processes are believed to work in reality. Obviously, the idea of finding a simpler way to represent the spatial distribution of the catchment using a distribution function appeared.

The models built based on this idea use a distribution function to represent the spatial variability of runoff occurrence. This function can be a statistical function (Probability Distributed Model, Moore and Clarke 1981), a distribution of the catchment in hydrologic response units (SLURP model) or a function based on physical characteristics of the terrain that represents the hydrological similarity of certain areas (TOPMODEL).

The main limitations of these models are in this distribution function. This function shows that not all the catchment responds in a similar way to a rainfall input and the surface runoff volume that appears is different in different parts of the catchment. Because the model is not totally distributed like the model proposed by Freze and Harlan it is easier to determine the specific response to a rainfall input for each value of the distribution function but difficult to determine the exact areas where this values occur and their evolution in time. The result for the entire catchment can be correct, but the spatial distribution of the discharge generated is not known exactly.

TOPMODEL allows for a determination of the locations according to the values of the distribution function but this is not an exact result.

This is the reason why these models could be called semi-distributed models for runoff.

Recent development of models

Along with the development of computers, the distributed models are more and more complex and powerful. Recent models are coupled with the Geographic Information Systems (GIS) that they use at least for data input and result presentation. In this way, the models become easier to use and give results that are easier to interpret. On the other hand, the increased complexity of models leads to a larger number of parameters to be calibrated and along with them to more sources of error. These problems that arise in calibration can cause incertitude in prediction results.

From these reasons, models that are simpler and have a smaller number of parameters can be almost as good and as easier to use than more complicated models. In the case where the only need is the maximum discharge, lumped models like IHACRES can offer results as good as the spatially distributed models if the calibration is correct.

For spatially distributed predictions it is still unclear if a fully distributed model can offer better results than a distribution function based model like TOPMODEL in the case when the conditions of applying the model are correct.

Besides these considerations, the application of any model is also limited in terms of data available and knowledge in the hydrologic process domain. The forecasts are hampered by the limited data or measurement techniques available. Therefore, the models can be at most as good as the data used to apply, calibrate or validate them.

1.6 Spatially distributed hydrologic models

Spatially distributed runoff models. The SCS Curve Number method

The **SCS curve number** method is a simple method used on large scale for determination of the approximate runoff value corresponding to a certain rainfall quantity in a certain area. Although the method is designed for a single storm, it can be scaled to calculate the annual values for runoff in an area.

The SCS-CN method was developed in 1954 and it is documented in Section 4 of the National Engineering Handbook (NEH) published by Soil Conservation Service (now called the Natural Resources Conservation Service), U.S. Department of Agriculture in 1956. The origin of the method was probably based on the proposal of Sherman (1942 1949) on plotting direct runoff versus storm rainfall. The subsequent work of Mockus (1949) focused on estimating surface runoff for ungauged watersheds using in-formation on soil, land use, antecedent rainfall, storm duration, and average annual temperature. Andrews (1954) also developed a graphical procedure for estimating runoff from rainfall for combinations of soil texture and type, the amount of vegetative cover, and conservation practices.

All of these are combined into what is referred to as the soil-cover complex or soil-vegetation-land use complex (Miller and Cronshey 1989). Thus, the empirical rainfall-runoff relation of Mockus (1949) and the soil-vegetation-land use complex of Andrews (1954) constituted the basis of the SCS-CN method described in the Soil Conservation Service (SCS) National Engineering Handbook Section 4 (NEH4 1985).

The runoff curve number (also called a curve number or simply CN) is an empirical parameter corresponding to different soil-vegetation-land use combinations.

The SCS Curve number method only forecasts the quantity of runoff formed in any point of the catchment but does not model the flow routing or the distribution of runoff through time. Because of this reason the requirements of the method are quite low, only the rainfall depth and an empirical parameter named the Curve Number are mandatory. The Curve Number (CN) value can be obtained from the hydrologic soil group, landuse and moisture conditions of the soil, the last two values being more important.

The SCS-CN method is based on the water balance equation and two fundamental hypotheses (Mishra and Singh 2003).

The water balance equation states that:

$$P = I_a + F + Q \tag{1.4}$$

The first hypothesis states that the ratio of the actual amount of direct runoff to the maximum potential runoff is equal to the ratio of the amount of actual infiltration to the amount of the potential maximum retention:

$$\frac{Q}{P - I_a} = \frac{F}{S} \tag{1.5}$$

The second hypothesis states that the amount of initial abstraction is some fraction of the potential maximum retention.

$$I_a = \lambda S \tag{1.6}$$

Where:

P = total precipitation (mm) ;

I_a = initial abstraction (mm);

F = cumulative infiltration excluding Ia (mm);

Q = direct runoff (mm);

S = potential maximum retention or infiltration;

The current version of the SCS-CN method presented in NEH4 considers λ equal 0.2 for the usual practical applications. As the initial abstraction component accounts for factors like surface storage, interception and infiltration before runoff begins, λ can also take other values depending on the application. In theory, λ can take any value between 0 and ∞ (Mishra and Singh 1999) but most of the current applications use the suggested value of 0.2.

Combining equations (1.5) and (1.6), the main equations for the SCS Curve Number Method are obtained:

$$Q = \frac{(P - I_a)^2}{(P - I_a + S)} \tag{1.7}$$

$$I_a = 0.2 \cdot S \tag{1.8}$$

By replacing I_a in equation (3), an equation with only two parameters is obtained.

$$Q = \frac{(P - 0.2 \cdot S)^2}{(P + 0.8 \cdot S)} \tag{1.9}$$

The potential maximum soil retention, S, can be obtained according to the CN value.

$$S = \frac{25400}{CN} - 254 \tag{1.10}$$

The equations are based on the trends observed in data obtained from the study areas, so they are empirical equations rather than equations based on physical laws. The CN is a hydrologic parameter that relies implicitly on the assumptions of extreme runoff events and represents a convenient representation of the potential maximum soil retention, S (Ponce and Hawkins 1996). The Curve Number (CN) is used in the determination of S and values for the CN for different landuse, soil types and soil moisture conditions can be found in tables (table). The origin of the original CN array tables seems to be lost; Rallison (1980) and Fennessey (2001b) have published the only known papers indicating what watersheds the original data may have come from.

However, there also appears to be a misconception as to the scale of data that were actually used to develop the CN array table, or the CN's accuracy for use in making peak runoff rate estimates. The lack of information on the origins of the method and the lack of scientific testing of the results raises some doubts when very accurate results are needed, but the method is used everywhere in the world when a simple way to estimate some discharge values is needed.

The SCS Curve Number method was implemented in GIS by different authors (Crăciun et al., Mary Halley et al., Xiaoyong Zhan and Min-Lang) using the ArcView or ArcGIS products. The usage of GIS systems for the SCS Curve Number method allows for automatic calculation of the CN parameter based on spatially distributed data obtained from measurements in the field or through remote sensing.

Water infiltration capacity of the soil was classified by the USDA-SCS into four classes called hydrologic soil groups (Mihalik et al, 2008, Matziaris et al 2005) Every type of soil has a Hydrologic Soil Group (HSG) that indicates an infiltration capacity and a rate of water transmission through the soil. The four types of HSGs are presented in table 1.1. The HSG values are based on the intake and transmission of water under the conditions of maximum yearly wetness (thoroughly wet) and are valid for unfrozen soil. When assigning a HSG to a soil, bare soil surface is considered. The land cover and land use are used in conjunction with these HSGs in order to obtain the final value of the Curve Number (CN) parameter.

Table 1.1. - Classification of hydrologic soil groups (USDA-NRCS 2007)

Hydrologic soil group	Characteristics of the hydrologic soil group	Infiltration rate
A	Soils in this group have low runoff potential when thoroughly wet. Water is transmitted freely through the soil. Group A soils typically have less than 10 percent clay and more than 90 percent sand or gravel and have gravel or sand textures	> 7.62 mm / h
B	Soils in this group have moderately low runoff potential when thoroughly wet. Water transmission through the soil is unimpeded. Group B soils typically have between 10 percent and 20 percent clay and 50 percent to 90 percent sand and have loamy sand or sandy loam textures	3.81 - 7.62 mm / h
C	Soils in this group have moderately high runoff potential when thoroughly wet. Water transmission through the soil is somewhat restricted. Group C soils typically have between 20 percent and 40 percent clay and less than 50 percent sand and have loam, silt loam, sandy clay loam, clay loam, and silty clay loam textures	1.27 - 3.81 mm / h
D	Soils in this group have high runoff potential when thoroughly wet. Water movement through the soil is restricted or very restricted. Group D soils typically have clay textures. Soils in this group may have high shrink-swell potential, a water table or a water impermeable layer close to the surface. , infiltration rate from	0 - 1.27 mm / h

The CN has a value between 0 and 100; lower numbers indicate low runoff potential while larger numbers indicate increased runoff potential. The lower the curve number, the more permeable the soil is. A CN of 0 is specific only to surfaces where no runoff exists like inland marshes or coastal lagoons.The values of the CN can be found in chapter 9 of the NEH – National Engineering Handbook for different land uses and soil groups. These values were obtained from data on floods, annual rainfall and runoff values taken from scientific literature for a large variety

of catchments, usually small catchments under 1 km^2 (USDA-SCS 1985).

<p style="text-align:center">Table 1.2 – Landuse categories and Curve Numbers (from Chow et al. 1988)</p>

Description	Mean % Impervious	CN for each hydrologic soil group				Typical landuse
		A	B	C	D	
Residential (High Density)	65	77	85	90	92	Apartment buildings, multiple family houses, trailer campings
Residential (Medium density)	30	57	72	81	86	Single family houses with land up to 1 acre
Residential (Low density)	15	48	66	78	83	Single family houses with land over 1 acre
Commercial	85	89	92	94	95	Shops, commercial areas
Industrial	72	81	88	91	93	Light industry, schools, treatment centers
Constructions	5	76	85	89	91	Parkings with gravel, construction areas, quarries
Agriculture	5	67	77	83	87	Cultivated areas, cereals and other cultures
Pasture or Range Land, good condition	5	39	61	74	80	Parks, golf fields, pastures on hills
Meadows	5	30	58	71	78	Grass , no grazing, mowed for hay
Forests (good condition)	5	30	55	70	77	Dense forests and brush that covers ground

Forests (rare)	5	43	65	76	82	Open forests with meadows or orchards
Impervious	95	98	98	98	98	Paved parking lots, roofs, driveways
Water	100	100	100	100	100	Areas covered with water, lakes, swamps

Maps with landuse distribution can be obtained from local administrations or from aerial photography. The categories of landuse can be determined according to the required level of detail in the results. Table 1.2 presents an example of standard CN values associated to the most frequent landuse categories for each of the four hydrologic soil groups. The landuse categories are the main landuse categories used in hydrologic analysis with the SCS method (USDA-SCS 1986)

The Curve Numbers presented in table 1.2 correspond to Antecedent Moisture Condition II (AMC II). The antecedent moisture condition can be defined as being the moisture state of the soil before the studied rainfall starts.

Due to spatial and temporal variability of rainfall, quality of measured rainfall-runoff data, the variability of antecedent rainfall and the associated soil moisture amount, the SCS-CN method has sufficient room for variability (Ponce and Hawkins 1996). A source of variability is also the antecedent moisture condition (AMC). Though the term antecedent is taken to vary from previous 5 days to 30 days (USDA-SCS 1986), there is no explicit guideline for varying the soil moisture with the antecedent rainfall of certain duration.

The SCS methodology represents this parameter based on the cumulated precipitation over the previous five days in the following way (McCuen 1982):

1. AMC I represents dry soil, with cumulated precipitation < 12.7 mm in the dormant season and < 35.6 mm in the growing season.

2. AMC II represents medium soil moisture, with cumulated precipitation of 12.7 - 28 mm in the dormant season and 35.6 – 53.4 mm in the growing season.

3. AMC III represents moist or saturated soil, with cumulated precipitation > 28 mm in the dormant season and > 53.4 mm in the growing season.

These values of AMC correspond, respectively, to 90, 10, and 50% cumulative probability of exceedance of runoff depth for a given rainfall (Hjelmfelt et al. 1982).

Table 1.3 presents the values corresponding to each AMC depending on the season and rainfall

Table 1.3. Antecedent Moisture Conditions

AMC	Five-day precipitation	
	Dormant season	*Growing season*
I	< 12,7 mm	< 35,6 mm
II	12,7 – 28 mm	35,6 – 53,4 mm
III	> 28 mm	> 53,4 mm

The Curve Numbers are calculated for AMC II and then adjusted by addition to simulate AMC III or subtraction to simulate AMC I. Different formulas can be used to adapt the AMC II curve number values to another AMC.

A GIS can be used to determine these parameters and to apply the presented equations and determine the runoff depth. This method will be used later in the study to obtain runoff depth according to a single torrential rainfall. The SCS Curve Number method was implemented as a Python script that can be used in GIS to obtain the runoff depth using freely available databases and spatially distributed rainfall.

Complete spatially distributed hydrologic models.

The blueprint for the complete spatially distributed hydrologic model built by Freeze and Harlan in 1969 is a completely tri-dimensional scheme for saturated and unsaturated groundwater runoff coupled with a bi-dimensional scheme for surface runoff and a uni-dimensional scheme for channel flow.

The interaction of different processes can be made by information transfer in the areas where the model components connect. For example water stored in micro-depressions can be used for infiltration and then added to the groundwater flow component. In this way, all the processes can be represented and solved as a single system of equations. In practice, finding the solutions to these equations for an entire basin requires a large memory and high computational power, so different models have simplified the description of certain processes.

The solution to this problem can be obtained using a lower resolution in the representation of the catchment or by splitting the problem in a number of subproblems. The approach of the SHE model was the representation of runoff in unsaturated areas, with predominant vertical processes,

as a one-dimensional problem and the runoff in saturated areas where lateral flow is present as a bi-dimensional problem.

The **SHE (Système Hydrologique Européen)** model was developed and applied starting in 1977. A description of this model was published by Beven et al. (1980) and an explanation of the model's concepts was published by Abbott et al (1986).

The SHE model is a hydrodynamic model with distributed parameters that simulates seamless integration of all the important processes of the hydrologic cycle at catchment scale (Abbott et al. 1986). The model works based on a discretisation of the catchment in a matrix of square cells that cover the catchment connected to stream channels at sides of the hillslopes.

Interception, evapotranspiration, snowmelt and runoff are determined in every cell and the cells are connected by the bi-dimensional processes in the model (surface runoff and groundwater flow).

The SHE model is developed by the DHI (Danish Hydraulic Institute) under the name MIKE SHE. A MIKE SHE model can include any or all of the processes in the terrestrial phase of the hydrologic cycle using the modules corresponding to each process (Fig. 1.14)
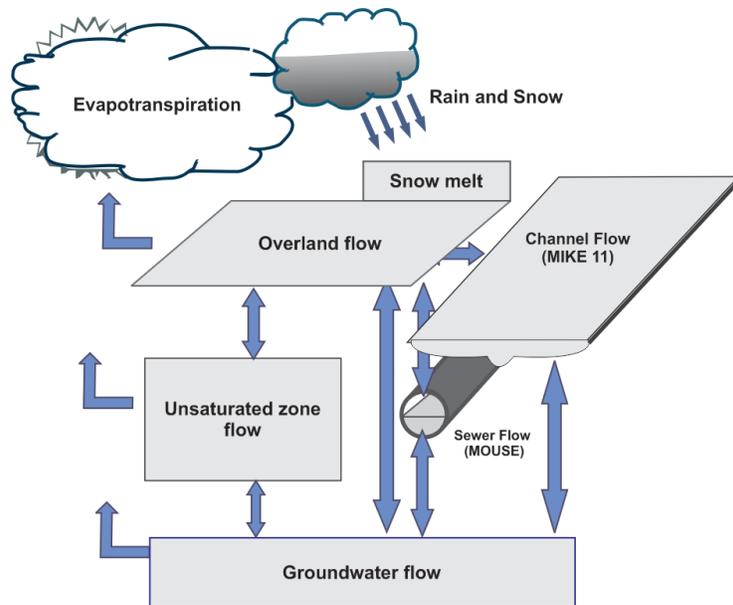
Fig.1.14 Processes modeled by MIKE SHE (DHI 2011)

MIKE SHE has different components for the simulation of each process (DHI 2007): precipitation, interception and exapotranspiration, surface runoff, groundwater flow in saturated areas, channel flow, sediment transport, freeze-defreeze dynamics, sewer flow in urban areas

There are other spatially distributed models based on the representation of the terrain as a grid of cells. Some of them are fully three-dimensional (Binley *et al.* 1989*)* (Paniconi and Wood 1993). The **ANSWERS** model (Beasley *et al.,* 1980: Silburn and Connolly 1995) is based on one of the first spatially distributed hydrologic model created by Huggins and Monke (1968). This model considers only infiltration excess overland flow and uses the Green-Ampt infiltration equation to determine the excess rainfall in every cell. The runoff is then routed towards the steepest slope from every cell.

The GIS is an excellent tool for the manipulation of input data and presentation of the results obtained using these models. Also, lots of data is available in spatial databases belonging to different agencies or authorities. More and more hydrologic models are adapted to use spatially distributed data and GIS interfaces are created for initial data processing for these models.

Spatially distributed models based on hillslope elements

These models are still spatially distributed but use an alternative discretization of the catchment. They are based on splitting each hillslope in planes. The splitting is usually done along the flowlines so there is little lateral exchange between adjacent hillslope planes.

These models can be used in places where the hydrologically active layer is mainly close to the surface and the underground processes are not very significant. The discretization of the hillslopes can then be made using the terrain topography obtained from different maps or from Digital Elevation Models. When modeling areas with significant underground processes the plane two-dimensional models (SHE) or fully tri-dimensional models for the subsurface flow domain are more appropriate. (Beven 2003)

Some models use elements that represent the areas split from the hillslopes, with different dimensions and shapes, as 'equivalent' planes with a regular form (uniform width, height and depth) and an uniform slope. Some of these models did not account for underground processes and treated infiltration as a loss. A well-known model using such a discretization was created by Smith and Woolhiser în 1971 and later evolved in the **KINEROS** model.

In Australia two such models were also developed: **THALES** and **TOPOG**. Both are based on a topographic analysis package named TAPES-C that can automatically split a hillslope into the hillslope elements needed for modeling based on the flow paths on the hillslopes.

One of the most important models based on this approach is KINEROS (KINematic runoff

and EROSion model). KINEROS is an event-oriented model based on physical laws. The model can describe the processes of interception, infiltration, surface runoff and erosion in small basins from urban or agricultural areas. KINEROS started in the '60s at USDA (United States Department of Agriculture) at the SWRC (South West Research Center).

The model represents the basin as a one-dimensional series of hillslope planes that contribute laterally to the feeding of streams (Woolhiser, et al. 1970). Rovey (1974) coupled an infiltration module to this model and called it KINGEN (Rovey et al. 1977). After important modifications based on field validation, KINGEN was modified to include soil erosion and sediment transport modules and called KINEROS. KINEROS was first used in 1990 (Woolhiser et al. 1990) and described by Smith in 1995. Later, KINEROS received a series of transformations and KINEROS2 was launched as an open-source program available to the public along with the necessary documentation.
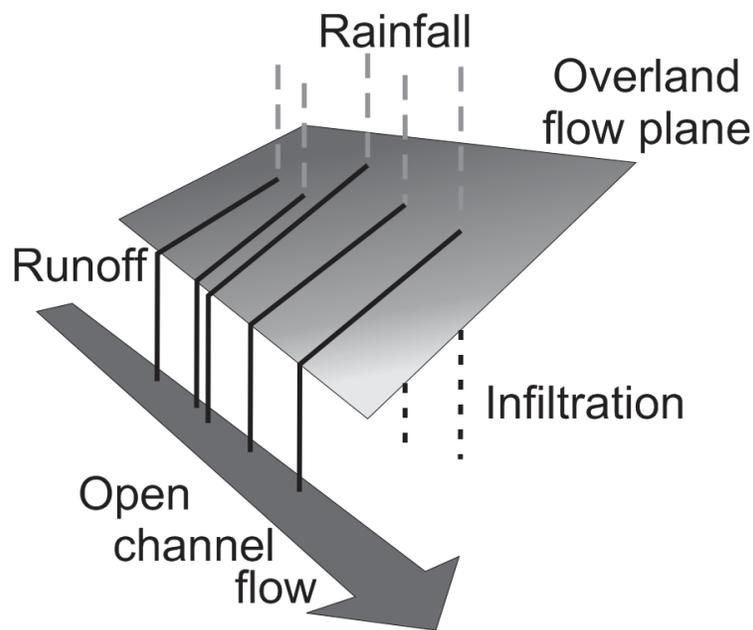
Fig. 1.15 *KINEROS model scheme (adaptation after Smith 1995)*

The partial differential equations that describe the processes of surface runoff, channel flow, erosion and sediment transport are solved by finite differences. The variation of precipitation, infiltration, runoff and erosion can be considered.

The application of this model needs spatially distributed data on the parameters, and the discretization of the catchment in hillslope elements with assignment of parameters for each element is a complicated and time consuming process. The development of GIS and the spatial

databases created a need for a GIS interface to the model. Such a GIS interface named **Automated Geospatial Watershed Assessment (AGWA)** was developed in 2002 (Miller et al. 2002) by the research department at the USDA, the Environment Protection Agency of the USA (EPA) and the University of Arizona.

AGWA was initially an extension to ArcView versions 3.x, a GIS created by ESRI (Environmental Systems Research Institute) starting with 2001. The interface and possibilities offered by AGWA are good for catchments where the topography is important. AGWA allows for an automatic extraction of parameters for KINEROS from GIS besides a way to visually present the model results. AGWA is also available as open source software on the internet. The newer versions of AGWA also support the newer versions of ArcGIS (9.x).

Models based on distribution functions

One of the most popular models based on distribution functions is **TOPMODEL (a TOPography based hydrological MODEL)**, a model that started more than 20 years ago with a simple structure. The model has a distribution function calculated based on the high resolution. DEM.

TOPMODEL appeared from the necessity to avoid the calculations made in every cell of a usual spatially distributed model. The idea of TOPMODEL is a discretization based on hydrologic similarity obtained from soils and elevation data through the catchment. In catchments with moderate slopes and thin soils over an impervious layer, the terrain characteristics have an important role in runoff generation, especially when the soil moisture is high. This was the basis of a hydrological similarity index created by Kirkby (Kirkby 1975) called **topographic index** that was later used as a basis for the rainfall-runoff model named TOPMODEL.
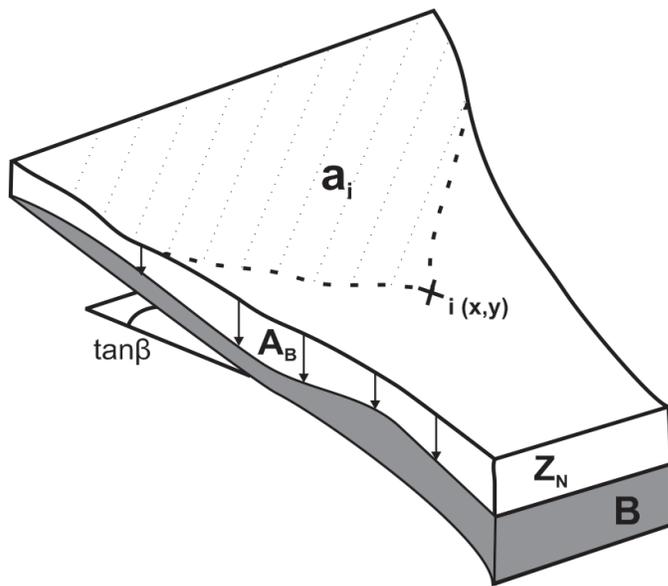
Where:
i – point of calculation
a$_i$ – upslope area from point i
tan β – slope in point i
B – Saturated (base) zone
Z$_N$ – Unsaturated zone
A$_B$ - percolation

Fig 1.16 *The scheme of the topographic index (adaptation from Beven 2003)*

TOPMODEL was created by Beven and Kirkby in 1979 on the basis presented in the previous paragraph, that the points in the catchment with the same topographic wetness index have the similar response to the rainfall input. This approach is useful because the calculations are only made for the values of the topographic index present in the catchment rather than for every cell regardless of the similarity to the other cells in the catchment representation. After the equations are solved, a distribution function will create the response at catchment level.

These details offer the possibility of obtaining the water deficit before saturation based on Kirkby's topographic index (4.10) (Fig. 1.16):

$$G = \ln (a / \tan \beta) \qquad (1.11)$$

Where:

a – upslope area from point i

tan β - slope in point i

The topographic index practically represents the capacity of every point in a catchment of reaching saturation. New improvements given to the topographic index calculation include the possibility of inserting the soil characteristic from the basin as an additional parameter.

The topographic index is used in the following way:

The soil profile can be vertically split into several zones (Fig. 1.17) The zone towards the surface represents the water stocked in the superior unsaturated part, where rainfall infiltrates until reaching saturation. If the ground is covered by trees or bushes an additional layer for interception

is needed. Another zone represents the saturated part of the soil that the water reaches after infiltration. The water reaches this zone with a delay and the accumulation of water in this zone lowers the distance between the saturated area and the soil surface.

TOPMODEL calculates the moisture deficit in every point at time intervals. The topographic index represents the water deficit and can be obtained from a DEM using functions from the last TOPMODEL version, The topographic index extraction from a DEM is based on a computer program written by QUINN in 1991 (Quinn et al. 1995).

The development of TOPMODEL was initiated by Professor Mike Kirkby at the School of Geography, University of Leeds under funding from the UK Natural Environment Research Council in 1974. The first versions were programmed by Keith Beven in Fortran IV on an ICL 1904S mainframe computer. The punched cards that were the program hard storage medium at the time have sadly (thankfully?) long since disappeared. Since 1974 there have been many variants of TOPMODEL developed at Leeds, Lancaster and elsewhere but never a "definitive" version. (Beven et al. 1995)

Processes:
P - Precipitation
E - Evapotranspiration
I - Infiltration
A_B - Input to saturated zone
Q_S - Surface runoff
Q_B - Base flow
Q_T - Total flow
Zones:
Z_I - Interception zone
R_Z - Root zone
U_Z - Unsaturated zone
S_Z - Saturated zone
D_S - Water deficit
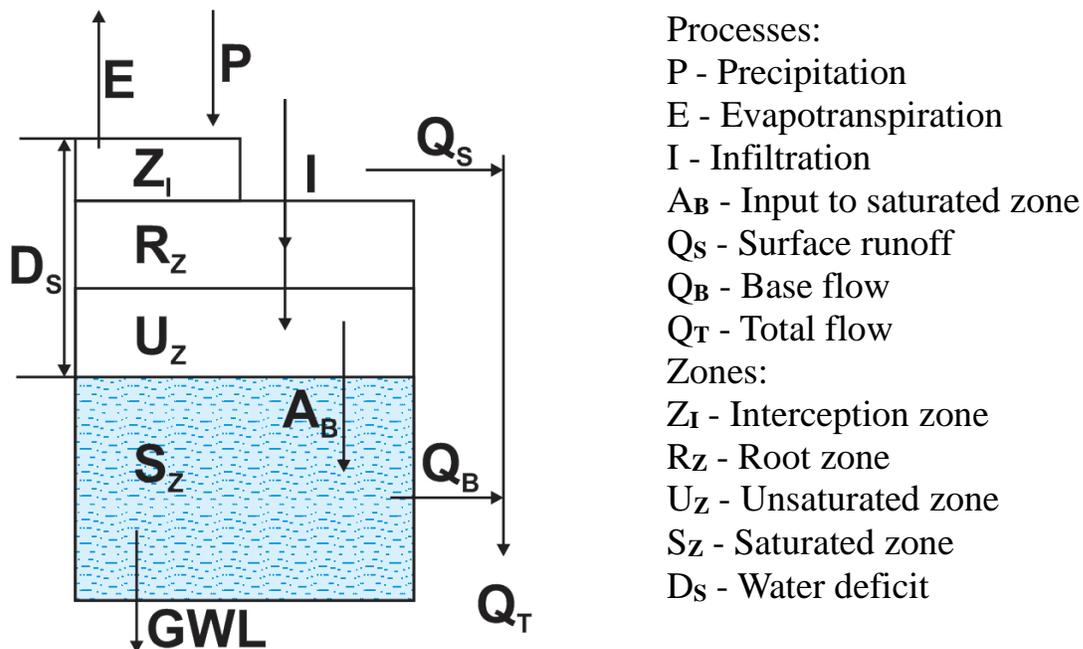
Fig. 1.17 The vertical structure of the TOPMODEL model

The TOPMODEL model was implemented in different GIS products like the open source GRASS, where it can be found as the r.topmodel module. SAGA GIS also allows for calculating the topographic index according to slope and drainage area of each cell in a raster.

The TOPMODEL model has been validated using data on rainfall and runoff (ex. Beven et al 1984, Hornberger et al., 1985, Robson et al., 1993, Obled et al., 1994, Wolock 1995) and different studies have been made regarding the application of TOPMODEL in water quality problems (Wolock et al 1990, Robson et al 1992).

The calibration of TOPMODEL was initially a difficult task, but the recent versions made this process simpler because they use a small number of parameters for the model. The model is influenced by hydraulic conductivity of the soil, water transmission of the soil when thoroughly wet, storage capacity of the unsaturated zone and flow speed of water through the streams when the catchments are bigger. There are also parameters that do not need calibration, like the resolution of the DEM and the time step used to run the model.

VIC type models: Xinanjiang/ARNO/VIC

The model **Xinanjiang/ARNO/VIC** is a model developed since year 1973 and published in 1980 (Zhao et al. 1980). Different publications use the three different names for very similar models, so all the three names can be used: Xinanjiang (Zhao and Liu 1995), Arno (Todini 1996) or Variable Infiltration Capacity (Wood et al., 1992; Liang et al., 1994; Lohmann et al., 1998a).

The Xinanjiang model is a conceptual model (similar to the ESMA models presented) with good results in flood predictions in humid and semi-humid areas in China and through the world for flood forecasting since its initial development in the 1970s. The main advantage of the model is the possibility of using the spatial variation of soil humidity. This possibility makes the Xinanjiang model perform better than other conceptual models. Like most conceptual hydrological models with lumped or semi-distributed structure, the spatial variation of hydrological variables is generally difficult to be considered. (Chen *et al*. 2007).

The VIC (Variable Infiltration Capacity) type models are developed as lumped models but have a function that allows for spatial variation or runoff in the catchment. This is the reason why this type of models have more things in common with the ESMA type models.

The details for the implementation of the VIC-2L (2 layers) (X, Liang *et al.* 1994*)* are presented as follows. The procedure is applied in subcatchments of the main catchment and the obtainded hydrographs are integrated and routed towards the outlet of the main catchment.

The runoff generation is based on the application of the saturation excess overland flow method. As the rainfall input continues, the storage capacity of the soil is filled and any, once filled,

excess rainfall on that part of the catchment becomes surface runoff. The spatial distribution of soil moisture and storage capacity is modeled through a parabolic curve representing it (Fig. 1.18).

The variable infiltration capacity allows for a non-uniform distribution of infiltration according to the equation(1.12):

$$i = i_m [(1-A_i)^{1/b}]$$ (1.12)

Where:

○   i is the storage capacity

○ i_m is a maximum storage capacity for the area

○ A_i is the percent of the surface with infiltration rate lower than i

○   b is a parameter that controls the shape of the distribution

For any level of soil storage the deficit towards saturation i_0 and the equivalent saturated surface A_s can be calculated The distribution of the storage deficit towards saturation (d_i) can be calculated using the formula(1.13):

$$d_i = i - i_0 = i_m[1 - (1 - A_i)^{1/b}] - i_0 \quad , \quad i > i_0$$ (1.13)
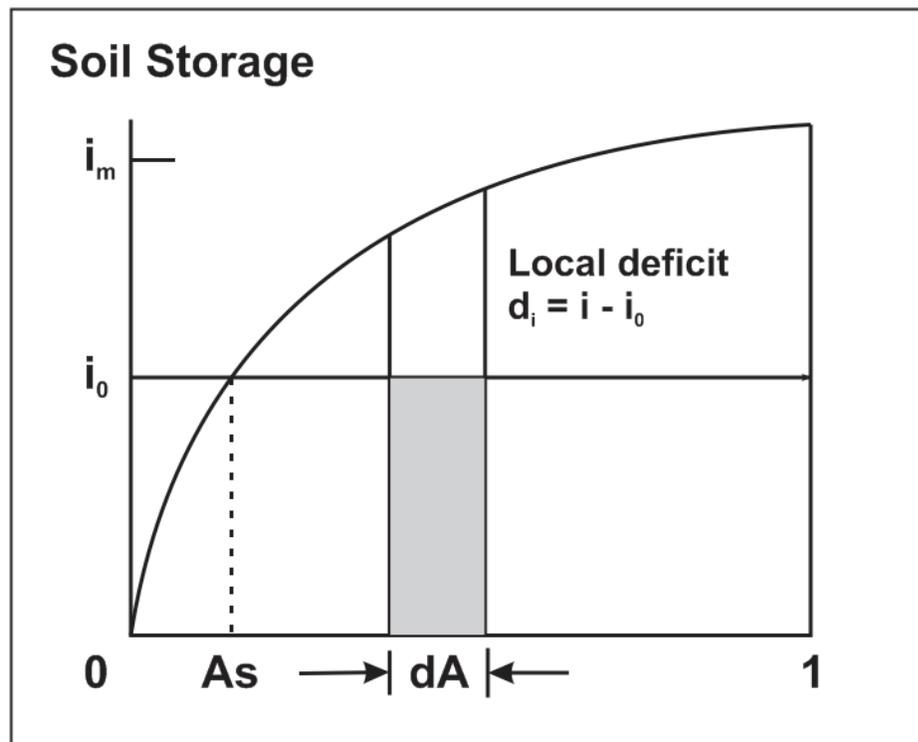
Fig. 1.18 *Storage deficit curve*

The model runs at fixed time intervals and the average quantity on rainfall over the area is

added to the water quantity in the superior soil layer in each timestep. The saturated zone is calculated as the zone where the storage capacity is overrun by the effective rainfall quantity. Any quantity of excess rainfall is then considered runoff that flows towards the stream.

In dry periods, the layers representing the soil dry gradually and the antecedent conditions for the next storm are created.

Evapotranspiration is considered to be at full potential in the saturated zone and at a reduced rate in the unsaturated zone, according to the moisture deficit. Three types of evapotranspiration are considered: evaporation from wet vegetation, transpiration from vegetation and direct evaporation from the soil.

At the end of each time interval, the storage in the superior soil layer is freed by the flow towards the deeper layer caused by gravity. The groundwater flow is represented by a function that can be applied towards the accummulated water from the deep soil layers. This function is linear for small quantities of water and becomes non-linear for large quantities, when a significant groundwater flow occurs.

An additional component for snowmelt can be added to the model, but requires calibration for a number of extra parameters. Liang et al. (1994) shows a way to apply the model separately for each landuse type.

The final part of the model, the flow routing towards the outlet, uses two linear reservoirs with different storage times: one for the groundwater with a longer storage time and one for surface water with a shorter storage time. The parameters representing these times can be obtained by calibration and are quite difficult to be measured.

The VIC model was updated and the current version is developed by Washington University (version 4.1.1) (VIC model overview). This version contains a number of improvements compared to the previous version and a larger number of processes can be represented in a spatially distributed way.

The new version allows for splitting the field in square cells, but these have dimensions larger than 1 km on one side. every cell can be split into zones with different land cover, but only the percent of each land cover category is considered, not the actual location. Therefore, the land cover surface is correct, but its distribution is not very accurate.

The soil can have any number of layers but it is usually modeled with three. The last model versions also allow for simulating the freeze/defreeze processes and the behavior of frozen soil.

The snow can be represented in three ways: snow on lakes, snow on the soil and snow in the canopy of the trees. The snow on the soil is split in two layers and the superior one is used for surface energy transfer.

The model can use daily or more frequent meteorological data (temperature, rainfall, wind) and can disaggregate daily data to sub-daily via algorithms of Kimball et al. (1997) and Thornton and Running (1999) and others.

The rainfall can be spatially distributed but, like in the case of land use, a cell can only have a time-varying wet fraction (where precipitation falls, dependent on on the intensity of the precipitation) and dry fraction (where no precipitation falls). The result at cell level is averaged using a weighted average function when it is written to output files.

The routing of runoff is modeled using the algorithm of Lohmann, et al. (1996; 1998) where each cell is a node in a flow network. The hydrograph from each cell is routed using linear approximations of the Saint Venant equations.

Even if the model is not completely spatially distributed, the spatially distributed models do not have a significant advantage over this model because of the lack of data. A comparison between models made by Reed (Reed at al. 2004) shows that spatially distributed models do not have better performance than lumped or semi-distributed models because of the data available (quantity, resolution, accuracy) used when applying the model.

The performance of each model depends on a large number of factors besides the structure of the model (lumped, semidistributed or spatially distributed), like the physical characteristics of the catchment, the available data, the accuracy of the data, the sources of error, so a perfect model for a certain application does not exist.

Models based on Hydrological Response units

The SLURP model, developed at the NHRI (National Hydrologic Research Institute) in Canada is a spatially distributed hydrologic model that continuously simulates runoff. The main model parameters are the interception coefficient, the storage in microdepressions, the surface roughness, the infiltration coefficient and the ground conductivity. The model can account for changes in the distribution of land cover and is very good for studying the effects of climate change.

The model can use local data for simulation or free remote-sensing data available on the

internet. Data like vegetation indices, land cover, cloud cover, snow cover can be extracted automatically from satellite images by the model.

The model splits the catchment in hydrological response units (HRU) (Fig.1.19). a HRU is not a homogenous area but a group of smaller areas with well-known properties. For example, the land cover can be obtained from images with a resolution of 10m, but it would not be practical to use such a high resolution on a very large catchment. Therefore, a number of cells are aggregated in convenient zones for modeling. The number of HRUs used in the modeling of the catchment depends on the size of the catchment and scale of input data.
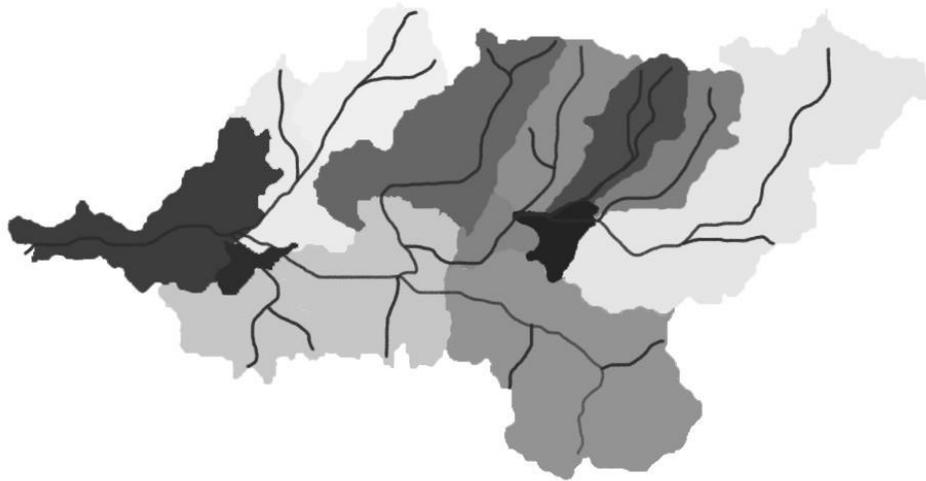
Fig. 1.19 *Discretizing a basin in HRU*

The split in subcatchments can be easily made using GIS because GIS can couple data from different sources and convert them in a form suitable for the application of the model. If the model needs calibration, at least some of the HRUs have to be in areas with measurements. For accurate results, the number of gauges must be larger than the number of land cover types. (Kite 1995).

At each time step, the model is applied in sequence in each HRU. Each HRU is represented by four reservoirs: interception, snow cover, fast runoff (subsurface flow) and slow runoff (groundwater flow). After routing the rainfall through the corresponding processes, the model offers results for evaporation, transpiration and runoff. The runoff is then accumulated from each HRU and routed according to a time-area relation depending on the land use. The accumulated flow is routed through each unit. The model takes into account the limits of each reservoir, infiltration, groundwater flow and other losses in flow routing (Kite 1995).

Where:
Processes:
P - Precipitation
I_C - Interception
E_T - Evapotranspiration
S - Sublimation
I - Infiltration
A_G - Groundwater recharge
T - Transpiration
A - Other (losses)

Output:
Qs - Surface Runoff
Q_H - Hypodermic runoff
Q_G - Groundwater flow

Reservoirs:
Z_I - Interception
Z_S - Snow cover
S_M - Soil Moisture
G - Groundwater
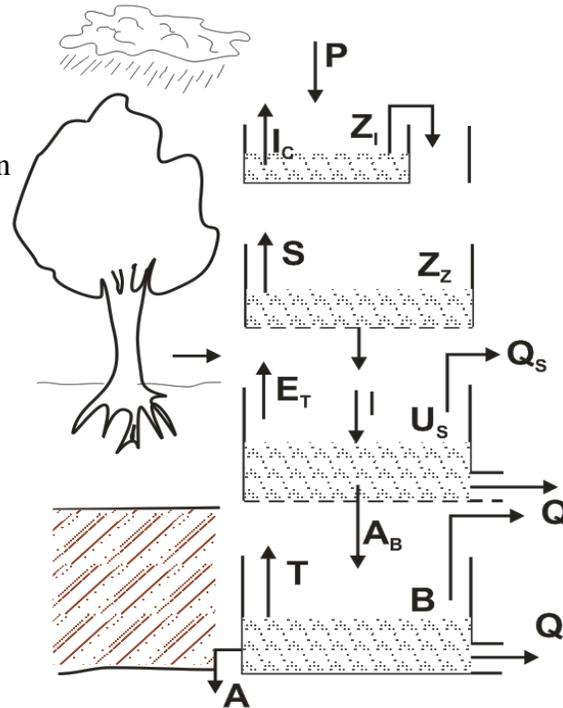
Fig 1.20 *The scheme of water transfers in the SLURP model*

The SLURP model was built to efficiently use the data obtained from remote sensing. The usage of data gathered by satellites is very good, especially in large catchments where data obtained in the field is not enough. A GIS is used in a land cover analysis and in the determination of distances from each land cover to the closest stream. The GIS is also used to determine the distance on the stream towards the HRU outlet. The runoff generated in each land use zone are routed to the closest stream, in the first phase, then routed on the stream towards the outlet of the HRU. After the runoff in all the zones from the HRU is routed, they are integrated to obtain the corresponding discharge for the HRU (in m^3/s), which is then sent to the next HRU downstream up to the outlet of the initial catchment. The routing can be done using non-linear reservoirs and the Muskingum-Kunge procedure or algorithms written by the user.

The parameters of the SLURP model can be derived from land use categories (ex> infiltration rate, Manning's n, hydraulic conductivity, soil depth etc.). A calibration is necessary to apply the model on large areas but once the parameters have been calibrated for a specific land use type, the catchments that use the same land use type can use the already calibrated parameters.

2. GIS Features for hydrologic modeling

2.1 Models in GIS – an introduction

Models are simplified conceptual representations of essential aspects of reality (process, phenomenon, object, element, system, etc.). In the case of GIS a model is a computer representation of a geographic feature or process.

According to Longley at al. (2005), the term of a model in GIS can have two meanings; therefore, a first classification of models can be made into data models and spatial models.

A **data model** is, in essence, a statement about form or about how the world looks like, a representation of different aspects of the earth surface in a digital data structure. Data modeling limits the options open to the data model's user to those allowed by the model's template.

The landscape is continuous and very complex compared to the possibilities of representation offered by a computer. A machine representation of a landscape as a digital stream of binary zeros and ones on a hard disk or diskette necessitates a considerable amount of abstraction, to say the least. (Brimicombe 2009)

The process of abstraction and translation into zeros and ones, known as data modeling, needs to be a formally controlled process to obtain usable results. The representation of data in the computer memory is based on a specific structure depending on the data and the model usage. Depending on the final data usage, only a certain part of the landscape characteristics are represented by each data model. The data represented using such a data model is later used in GIS for spatial analysis, presented as maps or used in modeling some natural processes (process models).

Spatial or process models are expressions of how the world is believed to work, expressions of process. The process of creating spatial models in GIS can be called GIS modeling. Spatial models may include dynamic simulation models of natural processes such as erosion or flooding. They may also include models of social processes (residential segregation or the movements of cars on a congested highway) and processes designed by humans to search for optimum alternatives (finding locations for a new retail store).

From the practical point of view, a GIS spatial model can be a workflow that uses existing GIS functions according to different rules and conditions. A spatial model can also include new algorithms that use functions from outside the GIS or functions implemented especially for the

model. An external software that only uses data derived from the GIS and does an independent processing of this data can also be a form of a spatial model.

The types of analyses that can be carried out and the spatial analysis functions that can be used depend on the memory representation of the earth surface characteristics. Therefore, a process model that needs to realize a certain analysis on some datasets needs an appropriate data model that allows the analysis.

2.2 Spatial modeling in GIS

This chapter presents the stages in creating a model for any kind of process based on a process description. The stages of the modeling process presented here correspond to the approach made by Beven (2003).

The first stage that needs to be considered in the choice or creation of a hydrologic model suitable for the needed application is a synthesis of the hydrologist's ideas on the modeled phenomenon. The result of this stage can be called a **perceptual model** because it represents the hydrologist's perceptions on the way the modeled hydrological system works. The perceptual model for the same phenomenon is not necessary the same for every hydrologist because it depends on the experience of each person, the datasets used, and the knowledge of different methods.

Even if this stage in modeling cannot have results or conclusions presented as scientific publications it is very important because every other stage in modeling will be based on simplifications and interpretations of this perceptual model. These simplifications and interpretations appear because the perceptual model does not necessary have a mathematical base. The perceptual model is formed of the hydrologist's ideas and can consist only of quantitative observations without any idea on its representation in mathematical language. Models that make quantitative predictions will need a mathematical description of processes and this is usually taken into account in the next stage of modeling.

The mathematical description of the processes involved in the perceptual model is the second stage of the modeling process. This stage can be called a **conceptual model** because it presents the concepts that lie at the base of the final model. At this point the ideas and assumptions being made to simplify the description of the processes from the perceptual model need to be made clear.

The result of this stage consists of equations representing the studied or simulated processes and the model's assumptions. The complexity of the conceptual model may vary. Depending on the expected results, some models may use simple mathematical equations while others use nonlinear partial differential equations representing complex processes. The choice of the simplification and abstraction level assumed in deciding on these equations must depend on different factors. Some equations can be directly translated in code that can be run in a programming language. In other cases, the equations do not accept analytical solving and an additional stage of approximation is necessary to solve the equations using numerical methods, só that the resulting equations can be implemented and run on a computer.

The translation of the resulting equations in procedures written in a programming language is called a **procedural model.** The results of this stage are the procedures that can be run on the computer representing the equations obtained in the conceptual model.

During the transformation of the equations from the conceptual model in code important errors relative to the true solution of the initial equations can be introduced in the results. Therefore, the knowledge and assessment of these errors needs to be taken into account and controlled at this stage as much as possible.

Before the practical application of the code obtained in the previous stage, a **calibration of the model** needs to be made.

All hydrological models depend on different kinds of data, input as parameters or defined as internal variables. Some of these data can be considered constant (landuse or catchment geometry characteristics). Some of these data can vary in time (rainfall or other meteorological parameters). Some of the data define the state of the catchment at a certain point in time and vary through time. The state of the catchment is known at the start of the modeling process and is modified according to the results of the equations. These state variables can include characteristics like the depth of groundwater or soil water quantity.

Other parameters define different characteristics of the catchment or surface. They can be constant (soil properties) or variable, but their values are not always known from the start. The adjustment of the parameter values at the start of the model is needed to achieve a good correspondence between the modeling results and the real results of the modeled process. This is

called model calibration and can be made based on previous observations from the field or by comparison with other models known to work well.

After the calibration of the model is finished, the last stage of modeling follows. This stage presumes a comparison between the modeled results and field observations and calculation of some indices of performance for the model. This stage is called **model validation** and is usually carried out by applying the model in different scenarios and comparing the results of the model with the results observed. The result of this stage is a quantitative evaluation on the performance of the model and a knowledge of the accuracy obtained in the results.

Another possibility of model evaluation is a comparison between the final model and the initial perceptual model. This is a much more difficult task because the initial parameters can be modified and the results of the final model can even lead to a change in the initial perceptual model in the case of new knowledge obtained after applying the model.

Types of spatial models

Different authors choose different methods of classifying the spatial models. Some of these are valid for any kind of models and do not consider the specifics of spatial models. The first classifications did not consider the IT component either.

For example, Ackoff (1964) classifies the models in the following way:

- **Iconic** - Uses the same materials as the studied phenomenon at a smaller scale
- **Analogue –** Uses different materials and a smaller scale
- **Symbolic** – Uses a symbolic system, like mathematics

Thomas and Huggett (1980) classify the models in a similar way according to the degree of abstraction required in modeling (Fig. 2.1)
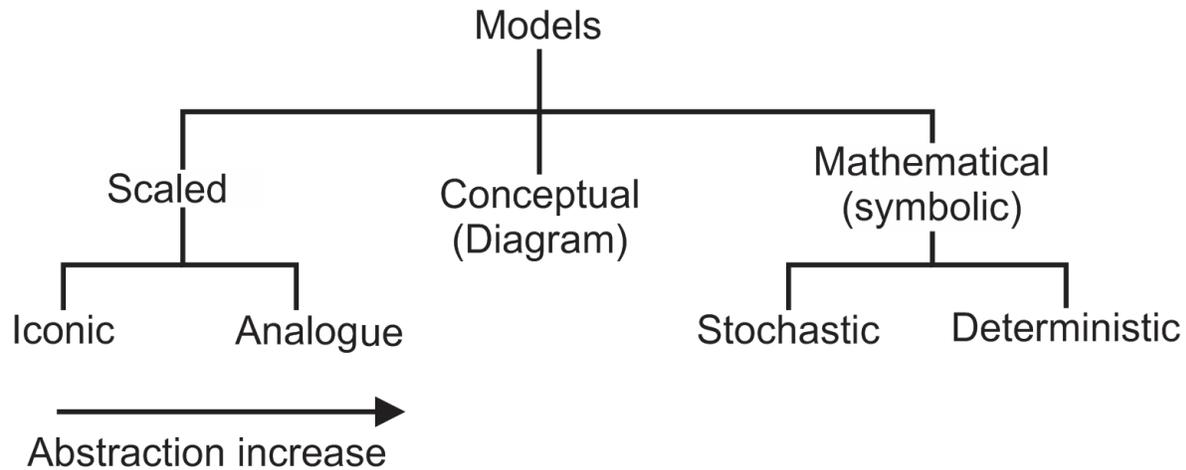
Fig. 2.1 *– Classification of the models according to abstraction (Huggett 1980)*

Another classification was made by Chorley and Huggett (1967) with regard to the modeling process:

Models based on analogy – These models use an event from the past to explain the present event (historic analogue) or a similar event from another location (spatial analogue)

Physical models – These models are a physical setup created from similar materials to study the forming, evolution of state or influence of different factors on a process

Mathematical models – A phenomenon and its evolution are represented by a series of mathematical relationships (equations, functions or statistical laws). A mathematical model can be deterministic if it returns a unique solution or stochastic if the solution is a probabilistic answer with random behavior taken into account.

Because of the recent technological advance, Brimicombe (2005) considers that the addition of a fourth category of models is mandatory:

Computational models – A computer is used to manipulate the data and code written in a programming language is used to express phenomena and their workings. These models cannot be easily classified according to the criteria presented in the last paragraph because they can include deterministic and stochastic elements together with rules, logical (and, or, not) or conditional (if) operators, iterations and even artificial intelligence components. Computational models tend to become very complex and offer a high degree of flexibility and realism in the way the result can be presented, sometimes at the cost of precision. Another advantage of the computational models

is the possibility of modeling complex phenomena, which cannot have a scaled equivalent created and models that require very high costs or difficulty to model using the other methods.

There are also other factors that can be used to classify models according to the role of time, the degree of specification of the model and the way in which the model is being used. Therefore, a model can be:

- Static – The elements of the model are fixed through time or a single state of a system is modeled

- Dynamic – The model variables vary through time

- White box – The model internal workings are fully specified

- Gray box – The model internal workings are partially specified

- Black box – The model internal workings are unspecified or little details are known about the process and the inputs are transformed into outputs based on correlations

Hydrological processes are variable both in space and time so they can also be classified according to other more specific characteristics.

The most important criterion used to classify a hydrological model is the spatial component: does the model take spatial distribution into account or not? Two main types of models exist according to this criterion (Beven 2003):

- **Spatially aggregated models** – These models treat the whole catchment or study area as a whole and use the mean values throughout the catchment in calculations. Using this type of model quantitative results can be obtained but their spatial distribution is unknown.

- **Spatialy distributed models** – These models offer results that are spatially distributed throughout the catchment by discretising the catchment in a large number of elements (subcatchments, hillslopes) or cells which have characteristic state variables. The values for the parameters in this type of models must be known for each element from the discretisation and the characteristics of the phenomenon throughout the study area must not be constant. The results obtained for each element can then be integrated according to different rules and assumptions.

A classification of the models according to the concept of causality and chance (Şerban

1995) can split the models into two categories: deterministic models and stochastic models.

- **Stochastic models** – These are black box models that use mathematical and statistical concepts to link an input (for example precipitation), represented by random variables, to an output (for example discharge). The modeling process implies the determination of a relation between the input data and the results without accounting for the physical mechanisms involved. A model of this kind does not try to explain the processes involved in the modeled phenomenon. The mainly used techniques are regression functions, transfer functions and neural networks.

- **Deterministic models** – These models try to represent the physical processes observed in the real world. Usually, these models are based on the knowledge of physical processes in the catchment and contain simplified representations of processes like : surface runoff, groundwater flow, evapotranspiration os overland flow but can be more or lass complex. A deterministic model is a model where the result is determined by mathematical equations representing the processes with a single value for each input parameter. Deterministic hydrologic models can be subsequently split into single-event models and continuous models. The deterministic models ensure the conservation of water volume in the system without regard to its location.

Apart from these classifications, there is a newer method based on fuzzy systems that is very appreciated and may be a very used method in the future. The number of fuzzy models is still small (Bárdossy et al. 1995) but their applications can be numerous. Fuzzy models offer the possibility of passing from a complex perceptual model to a procedural model in a simpler way.

The nature of the surface runoff processes in small catchments and the complexity of these processes require a deterministic spatially distributed model. Rainfall is not uniform in space and time and the runoff and flow processes depend on the rainfall. In small basins where surface runoff is the main hydrologic process generating the discharge, the spatial characteristics of the terrain (land use, soils, infiltration rate, snow cover, slope etc) are very important in runoff generation and need to be taken into account.

Initially, the hydrologic models were not created to use data obtained by remote sensing or Geographic Information Systems (GIS) because the availability of these systems and data was very limited. Along with the development of GIS, the increase in computing power and the availability of new datasets the hydrologic models are being based on more and more spatially distributed data,. With time, remotely sensed data and geographic information system data have become almost a necessity for spatially distributed models and such databases are also being made available to the public more conveniently. (Adhikari 2003).

Geographic Information Systems became a very useful instrument in hydrology. GIS is useful both in scientific studies and in water resources management. The recent climate changes lead to the need of higher interest in water resources and more knowledge related to this field. As every hydrologist knows, the water is in a constant movement and variation through space and time, so its study using computer systems (in this case GIS) makes the tasks of water management easier. Although GIS systems were initially built to be static in data representation, they became more and more dynamic with time, helping reduce the gap between historical data and actual reality in hydrology. (Maidment 2002)

Reasons for spatial modeling

The A to Z GIS dictionary (Tasha Wade 2006) defines spatial modeling as a methodology or set of analytical procedures used to derive information about spatial relationships between geographic phenomena. Therefore, unlike spatial analysis which presumes the application of different GIS functions to obtain new information from existing attributes or spatial relationships between existing data, a process model can be more complex and can do a large set of operations on spatial data. A process model is, therefore, an automation of a GIS workflow created to execute a specific analysis on some datasets.

There are many reasons why a GIS analyst can benefit from process model. Some of these reasons are presented in the following paragraph (Verbyla 2005):

- Simplify a tool for beginner user – even if this action does not presume the creation of a model but just the usage of modeling instruments, a GIS analyst can create a simplified version of a tool for an inexperienced user. This can be realized by giving default values to some parameters or creating a friendlier user interface for the tool.

- Chain together many geoprocessing tools – Some geoprocessing workflows need a number of geoprocessing functions to be executed together, in the same order. All this functions, represented as geoprocessing tools, can be chained in a single model that holds the order and conditions for calling these functions. An example would be the terrain preprocessing that needs to be carried out before any hydrological analysis.

- Parameters for flexible models – The parameters of the model can be organized in many ways and some models can use datasets of multiple kinds as parameters. A model can be used to create an appropriate interface for a certain usage. For example a layer can be located in a directory, a spatial database or a spatial database. A model parameter can allow the usage of one or more types of layers according to their location.

- Sharing of models among users – Probably this is the main reason why spatial modeling is useful, because there are a lot of cases when a number of users need to apply the same workflow for different datasets or in different locations. Models may allow an entire company to carry out the same operations on different datasets and reduce the time needed to obtain usable results. The possibility of sharing models among users is mandatory to make this situation possible.

- Using a model diagram to explain spatial analysis procedures – Whether the model is created using a graphical modeling tool or a program or a script written using a programming language, there is a possibility to create a diagram that presents the way the model works. UML (Unified Modeling Language) or a similar language is frequently used for this purpose. UML is a language used in Object Oriented Programming for description of classes and the connections between objects from the classes. Some graphical modeling tools have a specific way of presenting the internal works of a model, usually based on the UML representation used in programming.

There are some major benefits in GIS automation using one of the methods for creating process models (see chapter 6) presented in this work:

- Automation makes work easier. Once a process is automated, the analyst does not have to put so much effort in remembering the exact tools to use with the exact parameters and the proper sequence in which they should run.

- Automation makes the work faster. A sequence of tools executed by an automatic process model is executed much faster than anyone can accomplish by running each of the tools manually

- Automation reduces the possibilities of errors and makes work more accurate. As an analyst performs a manual task on the computer there is a chance of error. The chance multiplies with the number and complexity of the steps in the analysis. In contrast, a model can be tested extensively during its creation and then trusted to perform the same sequence of steps every time.

There are many reasons that lead to the necessity of creating models of hydrologic processes. One of the most important reasons is the limitation or lack of measurement techniques in hydrologic systems. Not all the components of a hydrologic system can be measured in detail and measurement techniques only allow for measurements limited in space or in time. Therefore, interpolating these measurements in space, in ungauged catchments or in time for simulating future event predictions or future hydrological changes is needed. Different models offer quantitative extrapolation methods that help in taking administrative decisions.

Hydrological models are simplified conceptual representations of a part of the hydrologic cycle. A lot of rainfall-runoff models exist and are still used in practice today. The models were created for different uses like understanding the processes occurring in hydrological basins or taking the right administrative decisions in different applications. The application of hydrological models can help in landscape planning, flood protection, water resource planning, management of chemicals and pollutants transported by water etc.

The increasing need for water resources in the world makes hydrological models very important for taking decisions in the context of actual climate change.

2.3 Perceptual models for catchment runoff

There is a lot of scientific literature that describe the processes involved in the response of a catchment to rainfall. Most of these texts present the processes involved in the response of a catchment to rainfall in more or less detail. The works of Kirkby (1978), Anderson and Burt (1990) present the experiences of different hydrologists in some chapters. The hydrologic systems are complex enough for every hydrologist to have its own opinion on the processes involved in the system, and thus its own perceptual model. However, there are some general phenomena common to more perceptual models along with phenomena specific to a certain climate, relief, soil, vegetation or other characteristics.

A model does not need to represent all of these processes in a catchment. One of the widely used methods for understanding certain processes in more detail is the study of a small part of the hydrologic cycle in detail. Because of this, there are studies realized only on certain hillslopes, experimental basins or in the lab. If the study is more detailed, the model representing the natural processes (in this case surface runoff contributing to flood discharge) is more complex.

This complexity can lead to more accurate results but often creates difficulty in choosing the equations and hypotheses used to define the conceptual model.

In the following part, a conceptual model suggested by Beven in 1991 will be presented in some detail. A schematic view of this model can be seen in Fig. 2.2
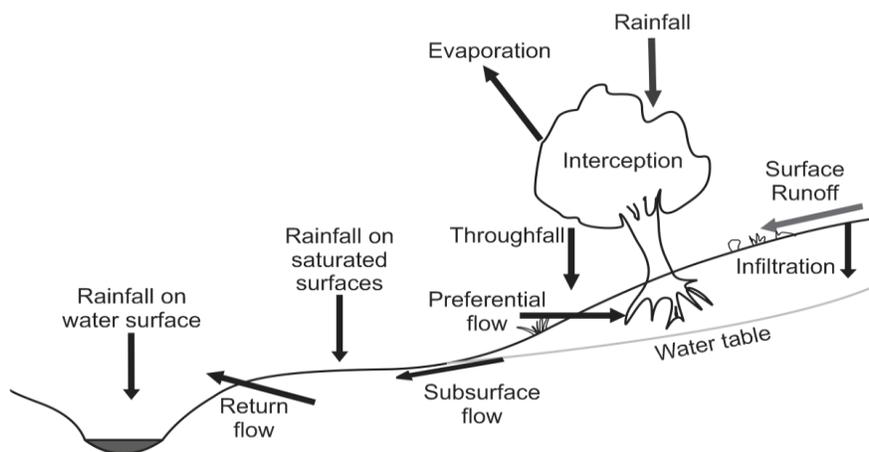
Fig 2.2 Perceptual model of runoff processes in a catchment (Beven 1991)

In dry periods, the storage in the soil and rock declines and the water table (if existing)

level gradually decreases. Storage may be higher and the water table may be closer to the surface towards the valley bottom because of downslope flow and the return flow from deeper layers. The vegetation has a greater or lower effect on runoff generation depending on the season, temperature and type. The interception of water by vegetation increases the time needed for the water to reach the soil. Evapotranspiration causes water loss that can be more or less significant depending on the climate, season and vegetation type. The plant roots can extract water from different depths but can also be pathways for water infiltration.

In case of a storm, the first answer to rainfall is given by rainfall falling directly on the stream. Although the surface of the stream can be quite small related to the surface of the catchment, the water falling directly over the stream can influence the final hydrograph in catchments with low runoff coefficients.

Rainfall not falling directly on the stream is affected by vegetation and soils in different ways. A part of rainfall falls directly on the ground and is subject to infiltration. Another part of rainfall falls over the vegetation canopy and is subject to interception. From the water intercepted in the canopy a part will reach the ground later on as throughfall and a part may run down the branches to the ground contributing to a local concentration of water in the area near the tree trunk.

Once water from any of these processes reaches the ground, it starts to infiltrate in soil layers. Exceptions to this are impervious surfaces (rock or frozen soil). Each type of soil has a specific infiltration capacity, and if the water intake to the soil is past this infiltration capacity the infiltration excess overland flow (Hortonian flow) will appear.

The soil infiltration capacity varies with soil type or soil moisture conditions. Therefore, the infiltration excess overland flow appears only in remote areas with soil permeability is the lowest or on areas where the soil becomes saturated. Vegetated areas have soils with a higher infiltration capacity and tend to delay the water input to the soil, so these areas are less vulnerable to infiltration excess overland flow.

Another type of runoff that may appear is saturation excess overland flow. The saturated soil areas occur first in places where soil moisture deficit is the smallest and then gradually increase with rainfall duration or intensity. Saturated soils may be subject to subsurface flow which may later come back to the surface as return flow. The return flow can cause the runoff to continue even

after the rainfall ends.

Snowmelt is also a factor in the generation of runoff and rise of water discharge. The main problem that arises in snowmelt modeling is the snowmelt rate related to different temperature and sun exposure conditions. Another problem that appears is the snow water content that varies according to the conditions under which the snow cover accumulated and can be influenced by temperature variations or wind amongst other factors.

Groundwater flow processes are more complex and the knowledge about these processes is limited by the possibilities of measurements in the areas. The groundwater flow can also contribute to the discharge rise in the streams and should be taken into account in flood modeling.

Regardless to the runoff mechanism used, the surface depressions need to be filled before continuous runoff can appear and runoff will sometimes concentrate and follow paths depending on characteristics of the terrain instead of flowing uniformly as a thin sheet of water over the land.

Although a large number of processes were presented, these do not appear in every case. Just a small part of these processes may be involved in the hydrologic response of a certain area and the hydrologic response can consist of different processes in different parts of the catchment or at different points in time, Some models are able to model just a part of these processes and other models can consider a process to be dominant over the others (for example the Hortoniam model only considers the infiltration excess surface runoff). Different authors presented the dominant hydrologic processes according to the climate or area, but there is no widely accepted classification for these.

Different perceptual models similar to the one presented contributed to a large number of rainfall-runoff models being developed worldwide. These models differ because of the processes modeled, the way the spatial distribution of the processes is considered, the expected results or other characteristics.

2.4 Flash flood conceptual model proposed for this study

Due to the characteristics of flash floods and after taking into account the data available, a perceptual model was created for the scope of this work. The model will be implemented in GIS

using a data model designed for hydrology presented in chapter 2.5 and the spatial modeling techniques presented in chapter 2.7. A schematic representation of the methods used can be seen in Fig. 2.3
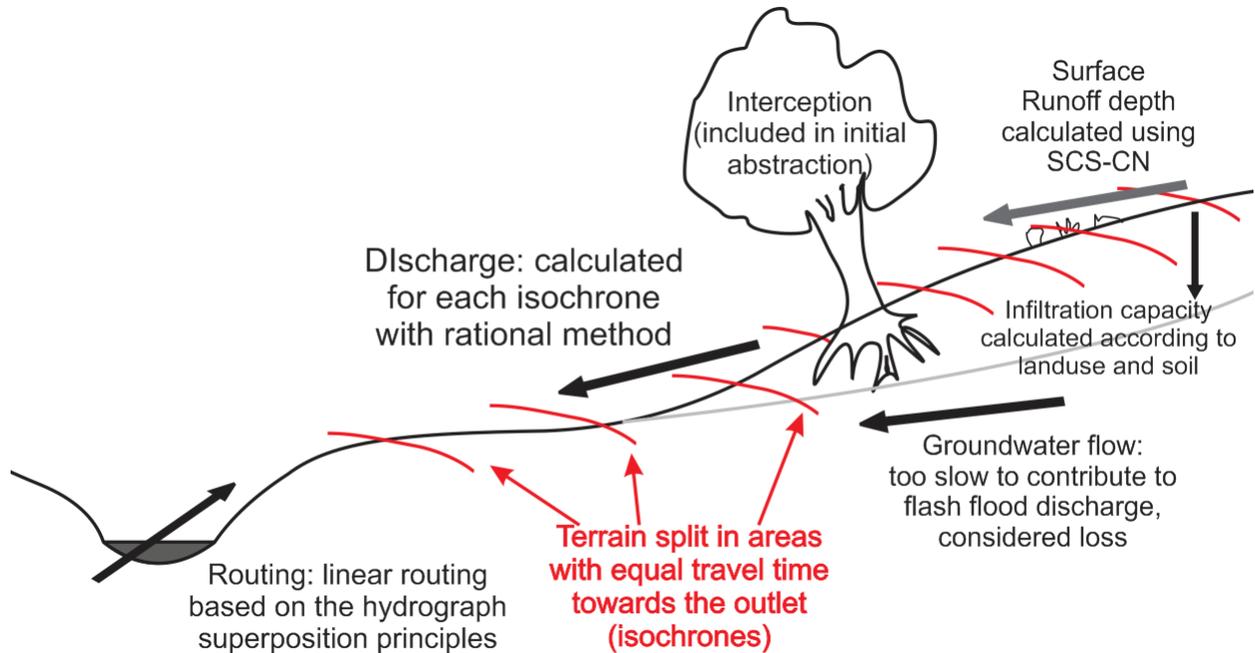
Fig 2.3 *Perceptual model of runoff processes used for this study*

The model will be based on the following assumptions (derived from the concepts presented earlier in this chapter at the SCS method, the unit hydrograph and the time-area methods):

- The surface runoff depth can be calculated using data on landuse, soil types and characteristics of the rainfall using the next workflow:

- Interception, surface depression storage and infiltration during the early parts of a storm are included in a single parameter: the initial abstraction. This parameter can be estimated from the surface and soil cover conditions at the start of the rainfall.

- Baseflow occurs when there is a fairly steady flow from natural storage. The flow comes from an aquifer that is replenished by infiltrated rainfall or surface runoff. Changes in this type of runoff seldom appear soon enough after a storm to have an influence on the hydrograph for that storm (USDA-SCS 1985) so the baseflow will be considered a loss and ignored

- The SCS Curve number method used calculates the direct runoff that contains channel runoff, surface runoff and subsurface flow in unknown proportions. All these types of runoff contribute to the hydrograph during a storm.

- The soil moisture conditions that affect the infiltration during the storm are included in the calculation of the empiric parameters used in the SCS-CN method: the Curve Number and the Antecedent Moisture Condition

The surface runoff depth is calculated using this method in each subarea of the basin represented by a raster cell.

Although the NEH suggests that an infiltration approach should be used to determine the variation of runoff during a storm, the limitations given by the data available call for a simpler method. The curve number runoff equation is not an infiltration equation (Smith 1976, Chen 1982) but it can, however, be used as a surrogate. The approach is quite similar to an approach suggested by Linsley, Kohler, and Paulhus (1982).

The runoff depth from the previous step and the rainfall depth are used to calculate a runoff coefficient according to the formula (2.1):

$$C = P / Q \tag{2.1}$$

The runoff coefficient is then used to calculate the discharge in every subarea using the rational method. The rational method can be used when the intensity lasts longer than the time of concentration of each catchment subarea. In this case, the catchment is split into a grid of cells with a small dimension compared to the limits of the rational method (max. 8 ha) and therefore the maximum discharge can be considered to appear from the start of the rainfall. The result of the application of this method is considered to be an equilibrum flow rate and the flow velocity is considered constant through time at this flow rate.

The time-area method is used to do the runoff routing. The flow velocity can be calculated using different methods and the model will use one of the methods available as a GIS module for flow velocity calculation from available data. The flow velocity at maximum flow rate using the DEM, the Curve Number, Manning's n and the rainfall intensity as parameters is calculated.

After the travel time to the outlet from each cell is determined, the areas with the same travel times are cummulated to obtain the total discharge corresponding to that time step. These

70

values correspond to an instantaneous hydrograph for the basin, but the rainfall duration also has to be included in the final hydrograph calculation.

The rainfall intensity is considered constant through the storm and a series of discharge values through time is obtained for each time step during the rainfall duration. The hydrographs obtained are then superpositioned and cummulated to obtain the final hydrograph.

2.5 Data modeling in GIS

Data representation for hydrologic modeling

The modeling of runoff in drainage basins needs datasets associated to each factor that influences the runoff created in a way that allows for integrated spatial analysis on all these datasets. A data model for hydrology has to facilitate the spatial analysis specific to this domain, not only to represent the hydrologic data. Maidment (2002) created a hydrologic model named ArcHydro at the CRWR (Center for Research in Water Resources) in Texas. He considers that the main characteristics of such a model must be the definition of a structure that is simultaneously a layer for hydrology in GIS and a base for the application in water resources.

An accurate runoff model must consider all data used by the mechanisms explaining the runoff generation and behavior. These data include, for the surface: soil type, land use, land cover, roads, cities and built areas. If the groundwater is also taken into account, the data can include: hydrogeological characteristics of the ground and aquifer information.

Different agencies that handle hydrologic and climate data in the world have different formats to represent the data, but not all these formats are directed towards GIS and the data is not available worldwide.

Data on land use, locations of the meteorological stations and different climatic variables already exist for free on the internet. Some areas also have databases on the soil properties or locations of gauges on the rivers. The algorithms for preprocessing GIS data were created to prepare the data for runoff or water quality models. GIS is now a very important and useful instrument for assembling the information on water resources, and the community of hydrologists working with GIS is in a continuous grow (Maidment 2002).

The representation of these data can obviously be done in different ways using the available base data models in GIS (raster and vector) but a system that allows for integration of these data is needed for an efficient runoff model.

Data models are very useful in modeling the hydrological processes because they allow for the usage of standard data structures in more spatial analysis operations or even in more models. Therefore, a data model should not limit the user to a single GIS product or methodology but offer a representation of data that can be used in different GIS products to create a process model.

The ArcHydro data model - description, usage, capabilities

One of the most popular models used in modeling hydrological processes is ArcHydro, the model created at the CRWR Texas and implemented in ArcGIS. ArcHydro is a geospatial and temporal model that runs inside ArcGIS. ArcHydro has a set of associated functions, designed and implemented by ESRI and CRWR that help in creating a spatial database, populating it and connecting the features in the database. The ArcHydro extension for ArcGIS also includes a number of tools that can be used in manipulating and analyzing the hydrological data or creating spatial models.

ArcHydro is an extension for ArcGIS and a data structure usable in hydrologic simulation models, but it does not contain any functions to simulate hydrologic processes. The modeling of hydrologic processes can be realized through the exchange of data between ArcHydro and another independent hydrologic model, through the creation of the model attached to ArcHydro as a dynamic link library (dll) or through extending the ArcHydro objects. Therefore, the functions and data available in ArcHydro can be used directly in the internal models of ArcGIS or in external programs calling the library methods.

Another scope of the ArcHydro data model is the creation of a possibility to couple geospatial data, representing the environment where hydrological processes occur, with time series, representing hydrologic measurements, to form a complete water resources management information system.

The data structure in ArcHydro is not definitive and can be extended by the user by adding new information useful for certain studies. ArcHydro was built to be useful in modeling water

resource problems at any scale and to adapt to specific modeling tasks.

The representation of hydrologic and hydrographic data in ArcHydro is a set of layers (Fig. 2.5) including different terrain characteristics and structures that create relations between these characteristics that will be presented.

The physical approach in modeling presumes the representation of all terrain characteristics specific to hydrology. The other approach, the "behavioral" one (Maidment 2002), defines the way the interaction between different features from different layers can be modeled.

The layers from the ArcHydro data model are stored in a geodatabase, a spatial database structure specific to the ESRI products. The geodatabase is a collection of geographic datasets of various types used in ArcGIS and managed in either a file folder or a relational database. It is the native data source for ArcGIS and is used for editing and data automation in ArcGIS.

Geodatabases have a comprehensive information model for representing and managing geographic information. This comprehensive information model is implemented as a series of simple data tables holding feature classes, raster datasets, and attributes. In addition, advanced GIS data objects add GIS behavior, rules for managing spatial integrity, and tools for working with numerous spatial relationships of the core features, rasters, and attributes.(Zeiler 2001). The geodatabase is a suitable environment for storing the data needed for hydrologic modeling in an efficient way along with the spatial relationships between datasets and assuring the spatial integrity of the database layers.

Water resources can be modeled for different purposes (floods, water quality, water supplies, infrastructure design, case studies over the landscape in a certain region) and each of these purposes needs another aspect of the water behavior, along with the laws governing it, to be modeled. The approach on modeling this behavior in ArcHydro is the creation of certain object attributes representing the connectivity between different elements from the same layer or from different layers.

ArcObjects uses ObjectId to uniquely identify a feature in a feature class, which is created along with the object and does not change its value. In ArcHydro each feature has two attributes:

HydroId - An integer attribute identifying the feature inside the database

HydroCode - A text attribute representing a public identifier of the feature

The connectivity between objects using this attribute is further presented.

The structure of an ArcHydro database regarding the connections inside and outside the database is presented in Fig. 2.4:

HydroId is an integer identifier for every feature used for internal management of the ArcHydro database

HydroCode is the public identifier of an ArcHydro feature, a text attribute that links ArcHydro and external sources of information

Using these two fields, ArcHydro can communicate with other information systems and automatically gather data necessary for hydrologic studies.

Fig. 2.4 *The ArcHydro identifiers and connections within the geodatabase*

HydroId is an internal identifier in the ArcHydro database that uniquely identifies each feature in an arcHydro database. The ObjectId used in each feature class and controlled by ArcGIS is not suitable for the unique identification because it can be modified by the system in different cases like copying the feature from one feature class to another. The ObjectId field identifies the feature within a single feature class and not within the entire database, so it cannot be used to create connections between different feature classes. Using the HydroId identifier relations between different feature classes can be made and different types of hydrologic structures can be used in

74

the same analysis.

HydroCode is an external identifier for each element in a feature class and can be used by other information systems outside ArcHydro. For example, a gauging station can have a unique identifier from the water resources administration and a database server can serve the data from this station using this identifier.

The unique identifier HydroId also identifies the feature within a single feature dataset. The standard format of a HydroId is the number that identifies the dataset followed by the number that identifies the feature. The management of the HydroId identifiers at database level is made through a pair of tables inside the database, namely **LayerKeyTable** and **LayerIdTable** that are created automatically. W hen a new feature is created in the database, a new unique HydroId is assigned to this feature and inserted in these tables.
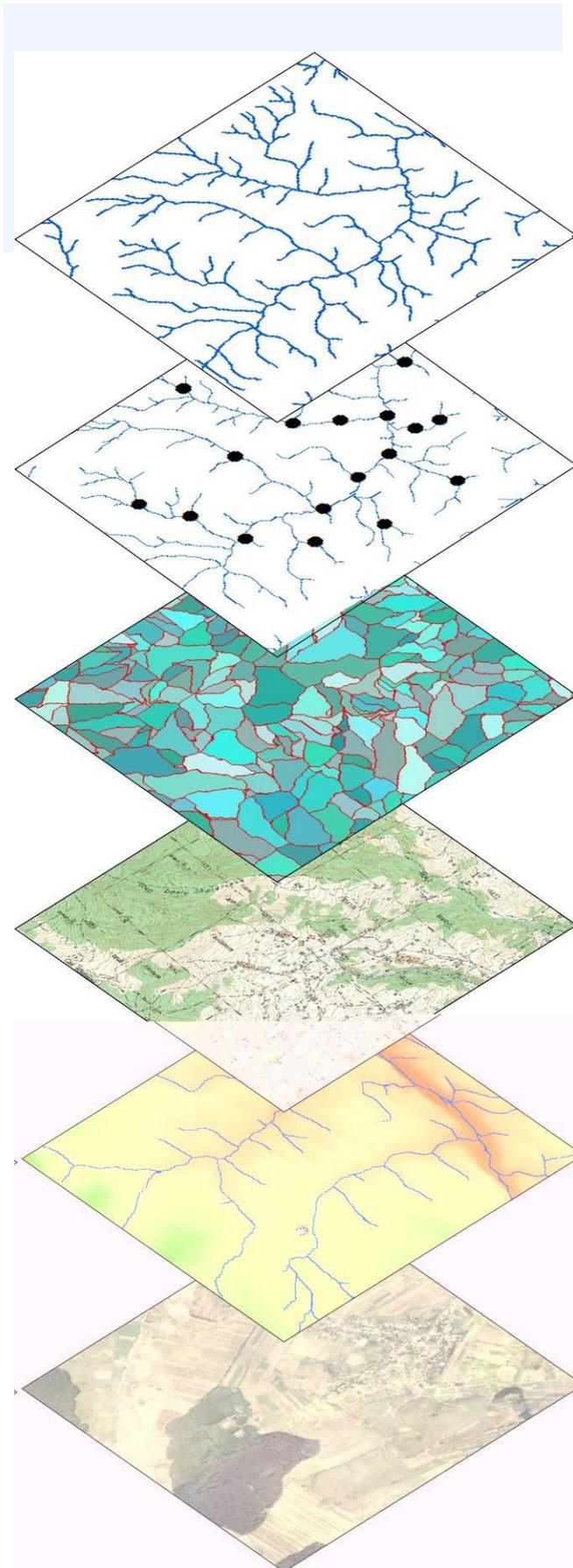
Representation: Edges and nodes for the flow network, polygons for lakes

Spatial relations: Each edge has a flow direction and flows into another edge or depression

Hydrographic points

Usage: Gaging stations or dams

Representation: Junctions and points in the flow network

Spatial relations: Can be junctions in the flow network

Hydrographic basins

Usage: estimation of flow into rivers

Representation: Polygons and points at outlets

Spatial relations: Fiecare bazin acoperă o secțiune de curs de apă

Topographic maps

Usage: Obtaining hydrographic data

Representation: data as points lines or polygons can be digitized from topographic maps

Spatial relations: the topology depends on the digitized information

Digital elevation model

Usage: automatic extraction of streams or catchments

Representation: TIN or raster DEM

Spatial relations: each cell in a raster has an altitude and the surface in a TIN is continuous

Satellite images

Usage: cartographic base

Representation: Raster

Spatial relations: cells in the raster fill the study area

Structure of the ArcHydro database

The main data structure used for automatic extraction and processing of hydrologic characteristics is the Digital Elevation Model (DEM). The DEM is a structure representing the altitude of the topographic surface in every point of the study area. The creation and data structures used for the DEM are presented in chapter 3. The ArcHydro data model uses a raster DEM representation where the terrain is represented as a grid of cells and every cell has a certain altitude value. A more extensive discussion on the raster DEM representation can be found in section 3.2.

The manual processing of some database features can be made using topographic maps. The topographic maps are usually available for every country in paper format at a certain scale, they can be scanned and georeferenced and then used for manual input of different parameters (permanent streams, lakes, flow control structures, bridges etc).

The satellite images can be used as a cartographic base for representation of the maps created and stored in the data model. The satellite images can also be useful in manual or automated determination of some characteristics (lakes, landuse etc).

The representation of the layers presented in Fig. 2.5 is made by organization of layers in five main categories: hydrography, flow elements, network elements, channels, time series. Features in a category can create features in another according to the techniques used in the analysis (Zieler 2009).

Flow elements

These features are derived from the Digital Elevation Model or from the surface topography. The lines represent streams or rivers, points represent outlets and polygons represent a catchment delineation organized on several levels.

Hidrography

Hydrography is obtained from the topographic maps through digitizing and can contain natural and anthropic details. The network can contain points (bridges, dams, gauges, pumping stations and other point features), lines (streams and rivers) and polygons (lakes).

Flow networks

Flow networks are represented as graph data structures. They contain points (nodes) connected by lines (edges) that show the path of flow through the drainage network. **HydroEdges** are edges in the network corresponding to streams. **HydroJunctions** are nodes in the network corresponding to point features (junctions, outlets). **HydroNetwork_Junctions** are additional nodes created by the GIS where the ones created by the user are insufficient. **SchematicLinks** and **SchematicNodes** are an alternative representation of flow where drainage areas are represanted by SchematicNodes and the connections between these areas are represented by SchematicLinks. Using these classes complex flow networks can be represented and the flow through these networks can be calculated with algorithms frequently used in graph theory.

Flow channels

These feature classes do not represent the exact location of features in the field but are a three-dimensional representation of channel bottoms (Fig. 2.6) with cross-sections, banks and flooding areas. The cross sections are made through depth measurements in the field across the flow channel. The representation of these profiles is useful in creating maps with flooding potential.
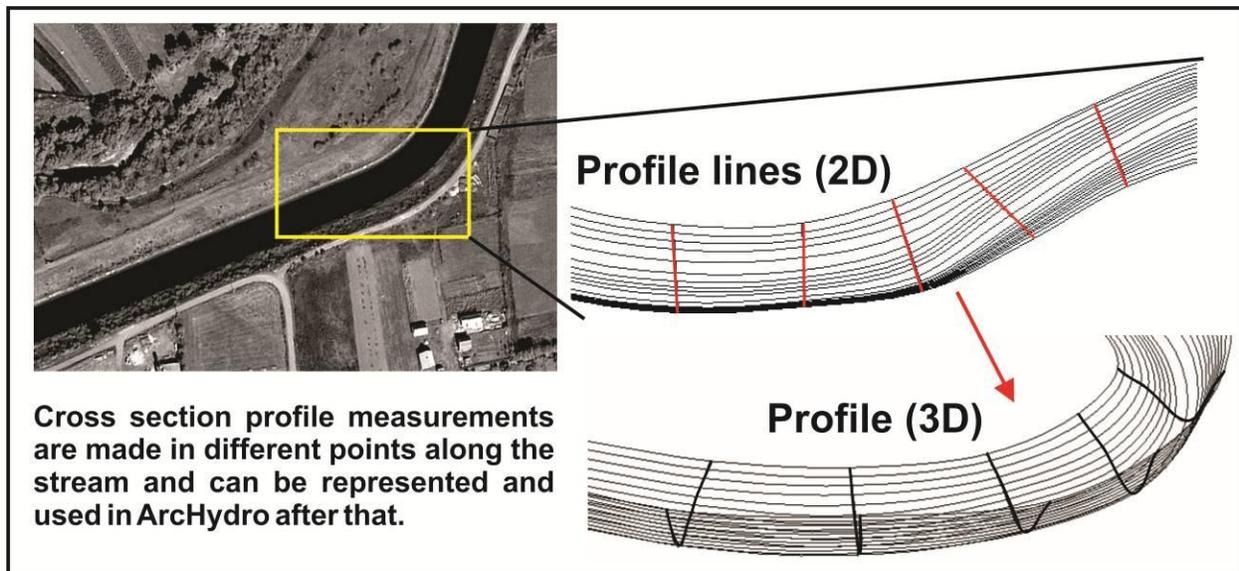
Fig. 2.6 Cross-section profiles of flow channels

Time series

Time series are used for periodical measurements from gauging stations and other facilities. Each time series has one or more features in the ArcHydro database associated and each feature can have one or more time series. The **TSType** table contains a set of metadata for each time series.

Processing functions available with the ArcHydro model

Although ArcHydro was created as a data model, an associated set of functions designed by ESRI and CRWR are available to create the database, fill it and do the basic operations on the data from the database. Some of the tools are present in any GIS product but some are specific to ArcHydro (especially those working with identifiers or with the database structure). A short overview of the functions available to manipulate hydrologic data will be presented in the next paragraphs.

The functions available with the ArcHydro model use the raster representation of the terrain

as a basis of extracting data on streams, catchments and the other flow characteristics. A more detailed view of the raster representation and processing can be found in section 3.2.

The first thing that should be done before any kind of hydrologic modeling is a review on the quality of data used in modeling and a correction of these data according to the needs in modeling. Almost all GIS products contain tools for **DEM reconditioning**. The DEM reconditioning is applied to make the application of the D8 (eight direction) flow model possible. The D8 method (Douglas, 1986; Fairfield and Leymarie 1991) defines the flow properties in any raster cell through the evaluation of the cell along with its eight neighboring cells (Fig. 2.7).

Regardless of the method used to create a DEM it is possible for it to contain sinks and flat areas, artifacts that appear because of the horizontal and vertical resolution, the methods of DEM generation and the noise present in some altitudinal data. (Garbrecht, Martz 1998) Sinks are cells with their value lower than all the surrounding cells. The sinks that appear in a DEM are not inexistent all the time, there are times when they actually exist in the field. The real sinks can be anthropic (open mine shafts, dams, storage tanks of different material) or natural (sinkholes, potholes and other karst landforms). The flat surfaces can also be anthropic or other types of surfaces like lakes recorded as flat by the measuring techniques. The flow ends in sinks and the flat areas have an undetermined flow direction. Therefore, a hydrologic model usually must correct a DEM to prevent problems that appear because of non-existing sinks or flat areas that cause the impossibility of routing flow towards the outlet of a catchment.

The deletion of the erroneous sinks is a step made by the hydrology applications where an uninterrupted and continuous flow network is needed to model runoff through the catchment (Tarboton et al., 1991; Wise 2000). Before modeling runoff using any DEM a sink removal procedure needs to be applied to dispose of erroneous sinks. Usually this operation is called **Fill Sinks** and takes the cell with the lower level to the same level with the lowest of the surrounding cells so that water flows past this cell when modeling. Some GIS tools allow for pointing the real sinks and ignoring them when filling the DEM. Other tools allow for manually specifying the lakes so that the altitude inside these is not misinterpreted.

After the preprocessing of a DEM, all the other base methods for hydrologic data

manipulation can be used.

The delineation of a catchment using GIS is usually based on the eight direction flow model (D8). In this model, each cell is connected to the eight adjacent (Fig. 2.7) and the flow direction is considered to be towards the steepest slope (largest difference of altitude between the cells)

1	2	3
4		5
6	7	8

Fig. 2.7 *A raster cell and the eight neighbours*

The result is the catchment corresponding to a certain point (Fig. 2.8) Different important hydrology tools are used in the catchment delineation.
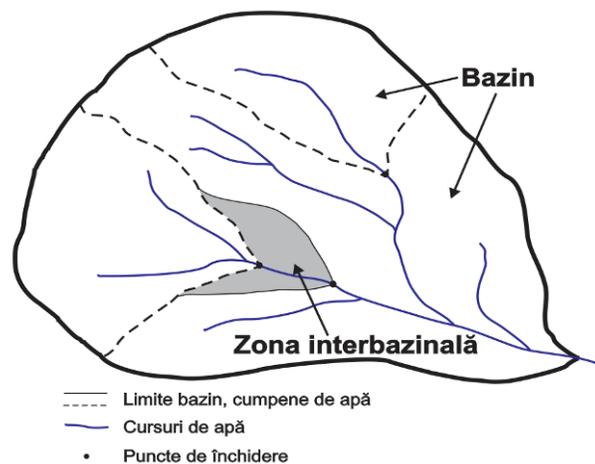
Fig. 2.8 *Hydrographic basin (Merry 2003)*

The **Flow Direction** tool creates a raster with the value corresponding to the flow direction from each cell. The most common method to represent flow direction is the D8 method, which presumes finding the direction of the steepest slope between the cell and the eight adjacent cells. The D8 method was introduced by O'Callaghan and Mark. (1984), and widely used since then (ex: Tarboton 1991 1997)

The resulting raster will have the value of every cell corresponding to the flow direction of water through that cell. The result of this function is a value represented on 8 bits (between 1 and 255). The majority of GIS products use the powers of 2 for the values to make the storage of this values in the computer easier (Fig. 2.9).

32	64	128
16		1
8	4	2

Fig. 2.9 *Flow direction*

The powers of 2 are used because the numbering scheme is derived from a series of binary numbers represented in the memory as 00000001 (=1), 00000010 (=2), 00000100 (=4) and so on up to 10000000 (=128). As the figure shows, a cell from which water flows to the right will have the value 1 for flow direction and the value doubles with each step clockwise.

There are other more complex methods of extracting and representing flow direction using hillslope orientation, multiple flow directions in a single cell or additional data, but each of these methods has its advantages and disadvantages.

The **Flow Accumulation** presents an indirect way to determine the drainage paths based on a DEM. The accumulation in a cell is defined as the number of cells that drain through that cell. The result of running the flow accumulation tool is a raster with the accumulation value calculated as specified stored in each cell. The calculation of the discharge accumulated on different segments of the drainage paths and the determination of streams can be done using the flow accumulation function. The stream network and the potential drainage lines are extracted from the accumulation of the flow on the terrain. If some models need the catchment area contributing to each cell it can be obtained easily by multiplying the value of the flow accumulation with the area of a single cell. Some GIS products have a separate Catchment Area tool specific for this purpose.

The **Catchment Area** tool determines the area that drained through each cell in the raster representation of the catchment. Unlike flow accumulation, the area is represented in area unit rather than the number of cells covering the area.

The definition of the streams can be done using the **Stream Definition** tool. This tool considers every cell with a flow accumulation larger than a value, given by the user, as being a stream. The value has a great effect on the density of the streams obtained from the operation and different applications can use different values for this purpose. There are also algorithms that determine this value in order to obtain a network as dense as possible without changing the geomorphologic properties of the flow network like the one implemented in the TauDEM program (Terrain Analysis Using Digital Elevation Models in Hydrology).

Another function that splits the streams according to their structure is the **Stream Segmentation.** Using this function, each stream is split in segments. The first segment represents the first level in the graph assigned to the stream network (from the source until the first confluence) and each edge of this graph (between two confluences) becomes a segment. Each segment has a grid code for unique identification of the segment.

The determination of subcatchments is then done using the **Catchment Grid Delineation** tool. This function assigns each cell a value (Grid Code) corresponding to the closest drainage segment that drains the cell. The result is a raster that contains the subcatchments corresponding to each segment from the Stream Segmentation result. (Fig. 2.10).
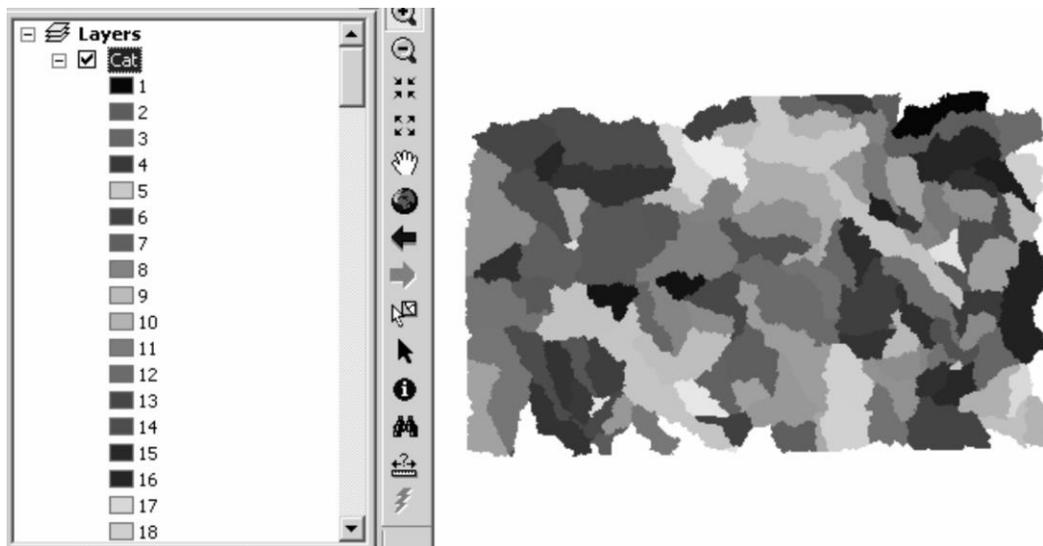
Fig. 2.10 *The determination of subcatchments using Catchment Grid Delineation*

The result of applying these tools in the order given creates the layers representing the catchment for each segment in the drainage network. If the user is interested in the watershed draining through a specified outlet, the **Watershed** tool can be used to obtain this result after the flow direction and the flow accumulation are obtained.

The ArcHydro model has another two tools that allow for the delineation of multiple watersheds if a layer with multiple points is given as a parameter. The tools are called **Batch Watershed Delineation** and **Batch Watershed Delineation for Polygons**.

The catchments obtained as a raster with the Catchment Grid Delineation tool can be converted to polygons using the **Catchment Polygon Processing** tool which converts from the

raster format to a polygon format and corrects the frequent errors (catchments represented by just a few cells, artifacts generated by the errors in delineation). After the conversion is done, the tool assigns a HydroId to each generated catchment for later use in the model. After the determination of each catchment as a polygon a tool named **Adjoint Catchment Processing**. can be used to obtain polygons representing the whole drainage area associated to each catchment determined earlier.(the surface of the catchment merged with the surface of al catchments upstream.)

Another tool that can be useful in some types of models is the **Drainage Point Processing,** that determines the drainage points for each catchment from the catchment delineation and stores them as point feature classes.

The suggested value for the parameters of these functions can be obtained using the "a thousand – a million" rule. The surface of the whole region divided by a million is the suitable cell dimension for the catchment representation. The multiplication of this value with 1000 is the minimum value for the drainage area that defines a stream (Maidment 1996). An example of using the functions presented can be seen in Fig. 2.11

Flow direction values

Digital Elevation Model (DEM)   Flow Direction

Flow Accumulation   Stream Network

Fig. 2.11 *Raster based functions for terrain analysis in hydrology*

These functions combined with other functions present in a GIS can be used to create different types of models that solve a large number of hydrology problems.

2.6 Database structure for the model proposed in this study

The ArcHydro data model is used to create a database of the layers needed in the model.

The database structure needed for this model will contain the main characteristics that contribute to surface runoff generation in the catchment. The main inputs are the DEM and the land use layers and the other data needed for the model is derived using the functions presented.

The model proposed in this study will only generate a hydrograph at selected points within a catchment. Therefore, the user must be able to select key points on the main streams within the catchments to apply the model. The database will be constructed using the functions available with the ArcHydro data model and the important characteristics of an area will be extracted from the DEM.

The terrain database construction will follow the following workflow (Fig. 2.12):
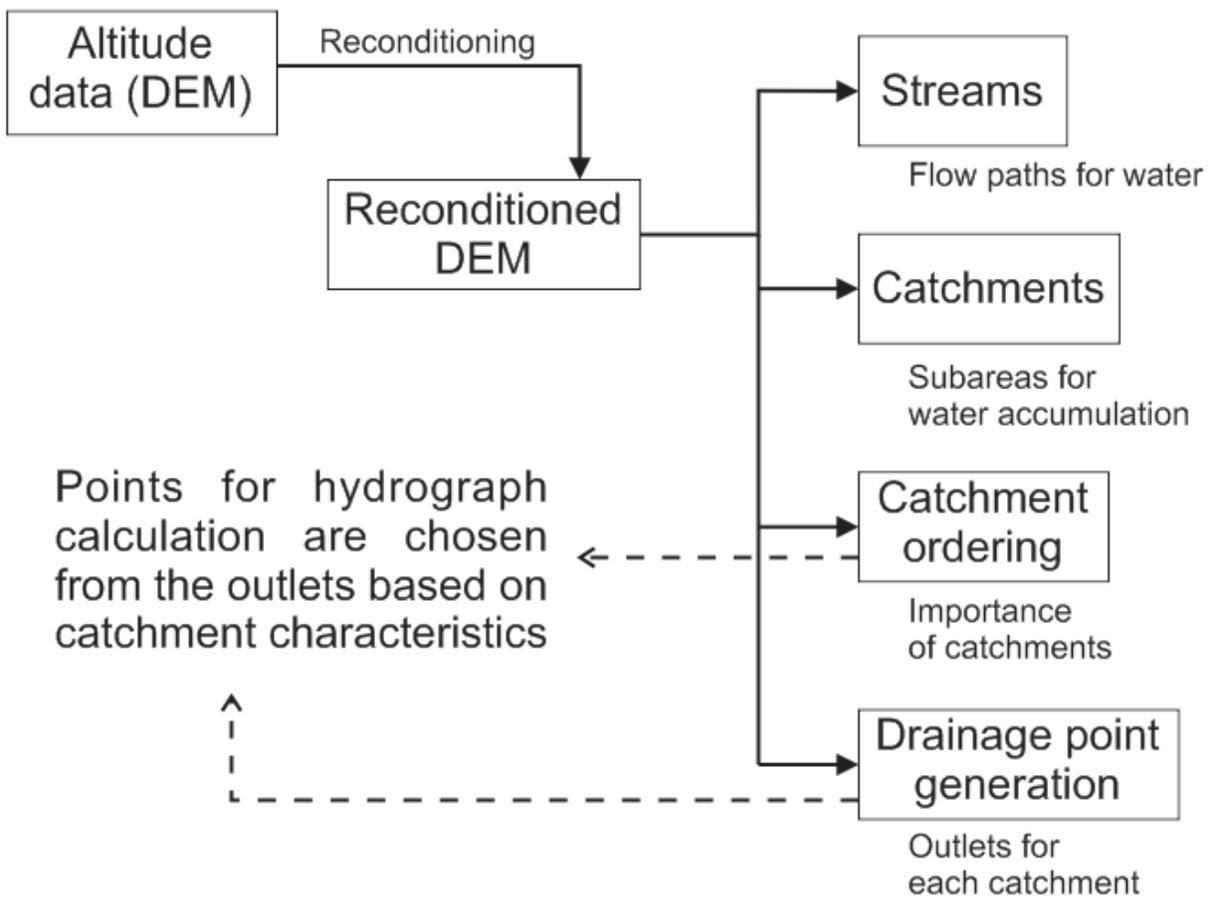
The input data for the database creation is a raster DEM obtained using the methods presented in chapter 3.

This DEM is reconditioned using the DEM reconditioning workflow. The main tool used in this process is the Fill Sinks tool, but other tools may be used if needed for correct determination of lakes or marking real sinks in the field.

The **reconditioned DEM** can then be used to apply all the other functions presented. The most important results of the processing sre the ones seen in Fig. 2.12. They will be used in the following way:

The **streams** (flow paths for water determined automatically) can be used along with topographic maps, satellite images or other information from the field (ex GPS recorded points) to determine the key points for model application.

The **catchments** are needed for different purposes. One of the most important purposes is the calculation of the contributing area for each stream. Another purpose is the determination of the order of water flow to the catchments which is needed in choosing the main points for the model application.

The catchments can be **ordered** using a standard stream ordering function available in the GIS. The order of the catchments can be used to obtain their importance in flood creation, according to the scale at which the model is applied.

The **outlets** of each catchment are determined using the drainage point processing tool. These points are the possible points for application of the model, but the user must choose between these points according to the needs of the application. Besides the terrain data other data may be used as an aid in this choice: satellite images and topographic maps to determine inhabited areas or GPS points to determine areas with specific needs for modeling.

Chapter .3 is dedicated to the obtaining of the altimetric data from different data sources and the creation of the DEM. Chapter 4 will present the processing needed for the terrain database creation and an automated model for all the steps of the workflow from Fig. 2.12.

2.7 Implementation of spatial models in GIS

Depending on the type of model chosen, there are different ways of implementing the model. The intrinsic models can use interfaces for graphical modeling provided by the GIS or can be created as scripts, while the models created as dynamic link libraries have to be written using code in a specific programming language, depending on the APIs (Application Programming Interfaces) offered by the GIS product, for the tightly coupled models or the possibilities of data manipulation offered, for independent models

Graphical modeling is a way of modeling assisted by a user-friendly graphical user interface that is integrated for modeling purposes in some GIS products. These models can use only internal functions from the GIS product and share the GUI with this GIS.

Programming based on the available libraries can also be a way of creating a spatial model. These programs that run outside the GIS are usually created using an object oriented programming language that has a specific API for the GIS product used. These programs can not be integrated in the GIS and need to be provided separately. In this case the modeling possibilities vary from using functions available in the GIS (an analyst can also use the graphical modeling method presented earlier for this purpose and it can save a lot of time) to implementing new operations based on data specifics that solve new problems and provide extra functionality. The models created in this way need to have enough complexity to justify the harder effort involved in their creation. If the modeling problem can be solved by a graphical model or a script, programming a new module is not justified. Another possibility offered through the development of programs outside the GIS is the creation of complex graphical user interfaces that do not impose running a graphical GIS product at the same time.

The creation of these programs requires a good knowledge of the API provided by the GIS used or the source code of this GIS. Open source products offer the entire source code for free and allow an experienced user to solve complex problems using this method.

A large number of models created as independent applications exist that depend on the data provided by the GIS. Some examples used in hydrology and coupled with ESRI GIS products are: HEC GeoHMS, NFF, StreamStats, Mike 11, MIKE FLOOD, MIKE 21, WISE and SMS (Kopp

and Noman 2008).

Automating GIS functionality by **scripting** is a newer concept for the GIS products. Even if certain products had the possibilities of creating scripts in different scripting languages in the past languages (ex: Avenue for ArcView, BASH scripting for GRASS GIS, Python or Batch scripting for SAGA GIS), there was no possibility of coupling functions from more GIS products in a single model. This fact changed in the recent developments of software, as the programmers created **Python** libraries that can be used for automation of modeling tasks. The Python language provides many opportunities for integration within GIS computing systems. Cross-platform capabilities and ease of integration with other languages (C++, FORTRAN and Java) mean that Python is most successful in gluing systems together (Butler 2005). Python scripting facilitates the communication between different GIS products and allows an analyst to use functions from more of these products in a single model.

From the methods presented, there are some possibilities for modeling created for persons with only the basic programming knowledge, without the need for extensive experience in programming. Graphical modeling and scripting are presented in the following part of the work.

Graphical modeling

Graphical modeling includes a new interface element in the GIS, an environment that allows the expression of the connections between different GIS functions for the creation of an automated workflow. This workflow can be a process model and usually it can be shared between users of the same GIS. The Graphical modeling interface offers an easy to use but powerful environment for model development, that allows for testing the steps in the model execution without modifying the data, modifications in the model while step-by-step testing and other operations similar to the debugging in software application development.

Some of the graphical modeling environments present in current GIS products are presented as follows:

ArcGIS Model Builder

ArcGIS includes an extensive set of geoprocessing functions organized as tools in the ArcToolbox window. The ArcToolbox window contains tools organized by category according to the licence and extensions available to the ArcGIS user. Every tool can ba called from the ArcToolbox window and has a GUI for setting its parameters before running.

Model Builder is the graphical ArcGIS environment for creating models and offers different possibilities in creation, administration and debugging of different models. A model created with Model Builder is available in ArcToolbox as any other tool and has a similar interface. The elements of Model Builder that can be combined to create more complex models are (Fig. 2.13):

- Input Data

- Tools (functions) applied on the input data

- Intermediate data sets

- Output data

The input and output of each tool are linked to the tool with arrows pointing in the direction of the processing flow: input data is processed by the tool and the result is returned as output data.

The figure presents a model with a single tool, one input dataset and one output dataset. The models created with Model Builder can be very complex. Besides creating linear models, Model Builder offers the possibility of implementing loops and iterations, creating parameters of different types and saving the model for future use or sharing.
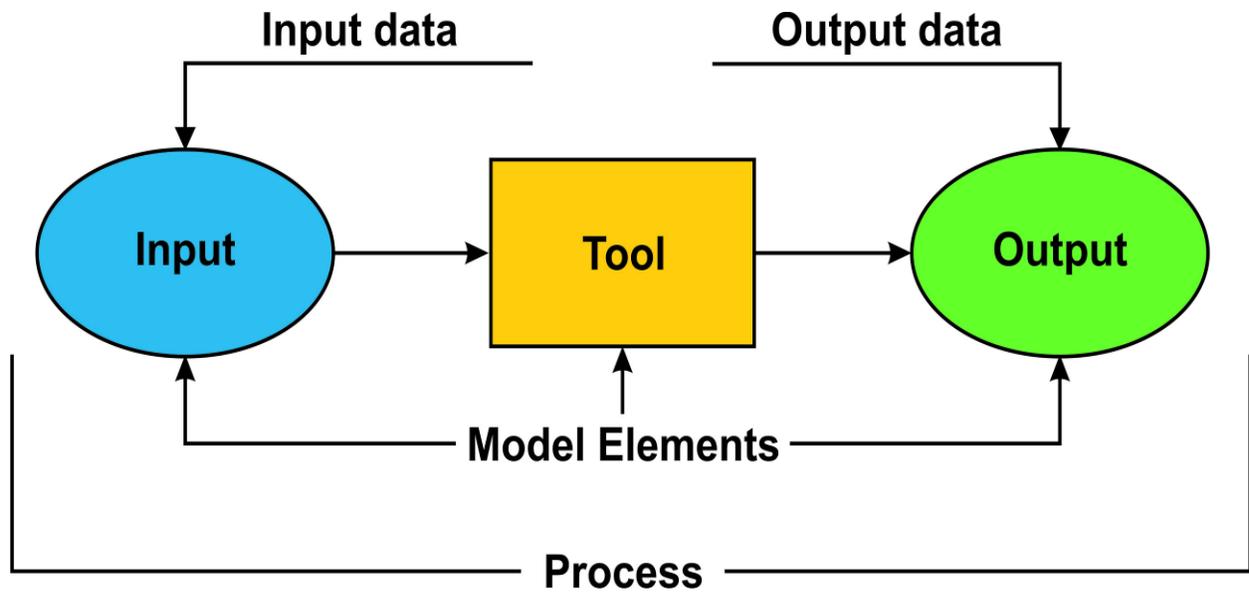
Fig. 2.13 *Elements of a ModelBuilder model*

IDRISI - Macro Modeler

IDRISI Macro Modeler is another graphical environment for the creation of models includes in the IDRISI GIS. A model created in Macro Modeler is similar to a model created in ArcGIS Model Builder in terms of elements: input data, processing modules and output data.

Macro Modeler includes functionality for group processing of multiple datasets and for dynamic modeling (iterations with data transfer from one iteration to another) (Eastman 2003). The Graphical User Interface of Macro Modeler is similar to the one of Model Builder, with different colors used to represent different data types (Fig. 2.14).
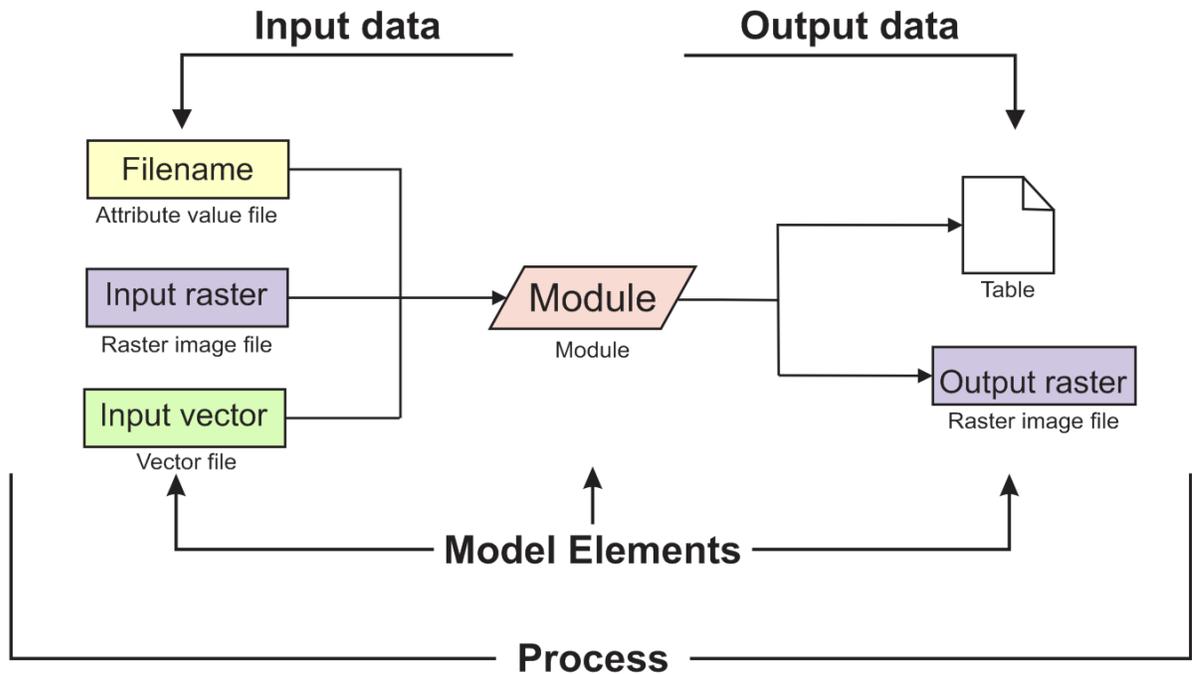
Fig. 2.14 *Modeling in IDRISI Macro Modeler*

SEXTANTE

SEXTANTE is an example for an Open Source graphical modeling interface. The Sextante library is written in Java and contains an extended set of geoprocessing modules. The library was developed under the auspices of the government from the Extremadura region in Spain. The main purpose of the Sextante library is the creation of a platform for the implementing, sharing and using a rich set of geoprocessing functions. The library leverages the access to a comprehensive (300+) set of algorithms and spatial operations, along with additional components to maximize its productivity and ease of use.

The main purpose of the Sextante library is the creation of a platform for implementing, sharing and using a rich set of geoprocessing algorithms. Similar to other Open Source projects, Sextante was initially developed as a product for a specific project, in this case the management of forests (Olaya, Gimenez 2011). Sextante has gradually become a solution applied in most types of geospatial applications. Sextante is an independent library written in Java that integrates tools from multiple open source GIS products written in Java (like gvSIG, uDIG or OpenJUMP). The

library can also be linked to geoprocessing modules from GRASS GIS to obtain more complex functionality. (Olaya 2010).

The graphical modeling interface in Sextante is also similar to the previous two solutions (Fig. 2.15):
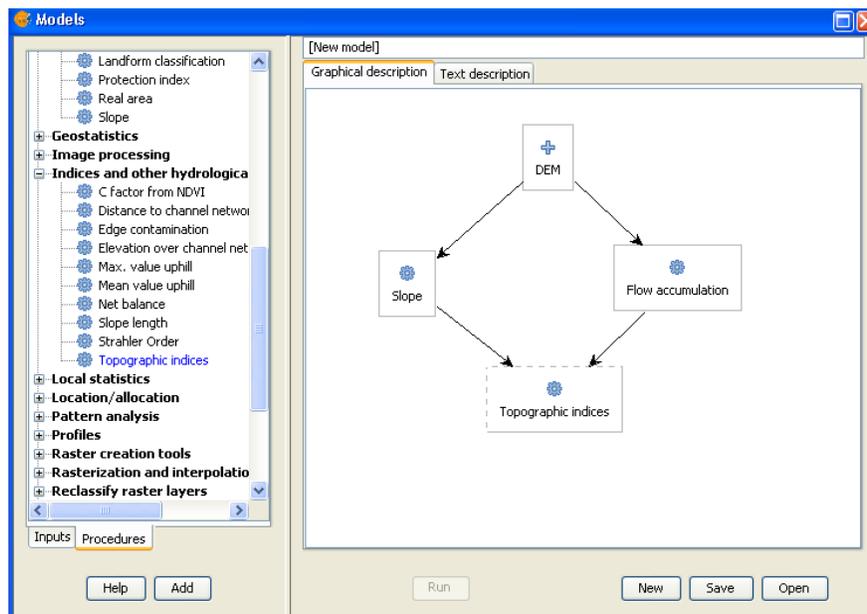
Fig. 2.15 *SEXTANTE graphical modeling interface*

Scripting

Scripting languages are programming languages that can be associated with applications and execute the code immediately without compiling. Scripting languages can be extended with modules that enable different operations (accessing different data types, graphics, complex mathematical calculations, network administration procedures etc.). Scripting languages are very useful in gluing systems together, allowing a user to create solutions for manipulating geospatial data and create process models. The rapid and interactive development of scripting solutions makes this procedure very attractive for the "user-developers" (Rosson, M.B 2005). Libraries from the geospatial domain that can be coupled using scripting languages are more and more common and a lot of GIS products implement their specific library for geoprocessing (Jolma et al. 2008)

The majority of programming languages need to compile (transform the source code input by the user in machine instructions interpreted by the computer) a program before using it. The source code written in scripting languages is interpreted line by line by an interpreter that comes

with the language. The interpreter executes the commands in the script line by line and every command is transformed in machine code and executed on-the-fly. The advantages that are obtained by running each command line without compiling the program are multiple, the main advantages being: a step-by-step debugging of the program is very simple and the wrong instructions can be rerun if needed; the commands can be input from a command line interface without writing an entire script beforehand.

A number of scripting languages existed during the computer evolution and some of them are still used today. Different scripting languages were focused on certain types of applications ar a specific programming task. The next paragraph includes a short list representing the evolution of scripting languages (TCL/Tk History of Scripting):

- JCL (Job Control Language) was used in the '60s to order the data introduced in the computer through punched cards.

- UNIX scripting languages (sh, csh, ksh, bash etc.) appeared in the '70s and are still widely used in Linux and UNIX platforms.

- Perl is a scripting language created at the end of the '80s and is well known for creating dynamic web applications and for server-side web application programming

- Tcl, along with the TK set of instruments became popular for being the fastest method of creating Graphical User Interfaces at the end of the '80s

- Visual Basic was created by Microsoft at the start of the '90s and combines the scripting language characteristics with characteristics of other programming languages for the Windows operating system. Visual Basic became well-known for creating GUIs in Windows and was used for a long time by applications running under this operating system.

- Python is an object-oriented programming language created by Guido van Rossum in the early 1990s to bridge the gap between shell and C programming. Its elegant, easy to learn syntax, high level data types, elaborate library, portability, and ease of extending and embedding in C/C++ all contribute to its popularity. Originally designed as an advanced scripting language, it found new uses as a rapid application development language for web, database and GUI applications, as well as for distributed systems and mobile code.

- Created in the mid 1990s by Netscape Corporation to perform scripting functions in Web browsers, such as simple form validation. JavaScript has become the de facto standard for client-side Web scripting, although it doesn't have much to do with Java.

There are many other scripting languages that are not described here or are derived from the ones presented. From these languages, Python was widely adapted by the GIS developers because it allows for easy communication between the native libraries existing in GIS and the scripting language.

The Python scripting language

The Python scripting language was developed in 1991 by Guido van Rossum at the Wiskunde en Informatica centre (CWI) in Netherlands. The characteristics of the programming language are inspired from other languages like ABC, Modula, LISP or Haskel. Initially, Python was developed as a scripting language but it developed in a language suitable for user interface programming, web programming, database applications, mobile and distributed systems (Butler 2005). Python includes advanced functionality for advanced users but the majority of GIS analysts writing Python scripts will only need the basic functions and data types. Python was designed to be a dynamic and easy to use language. Being an interpreted scripting language, Python scripts do not need compiling and the code is interpreted when running the script. The language is also interactive and the user can use a command line to write the commands of a script and run them one by one, which makes the language learning easier by trial and error.

As shown, the Python language tends to become the standard in GIS automation and most of the GIS developers created their own versions of libraries with the specific functions available in a GIS.

The development of GIS products did not take into account the interoperability between different products at the start and different languages were used for automation (either proprietary or open source). None of the libraries developed for these languages allowed the usage of functions from other libraries in a single program.

The ease of learning Python for non-developers is an obvious advantage, and a good reason to integrate a Python library in a GIS product. On the other hand, scripting in GIS can also face some problems and risks affecting the development time.

The first such problem is error management. Due to the fact that the commands in a script are executed line by line, the errors will not be detected until the program reaches the line with the error (Downey 2008) A compiler would detect errors when compiling and will not create a situation similar to the one in which an error will only appear after running a significant part of a long script.

Another problem is the speed of a script. A program compiled in an efficient programming language like C++ will offer much better performance than a script running the same tasks. This fact is no longer such a big problem because of the evolution in efficiency of GIS processing tasks and in computer performance nowadays. The main part of a script is usually processed by the internal functions of a GIS product which are compiled with the GIS and only called by the script. This is also the reason why writing a script can be faster than writing a program working on the internals of each GIS function, so it is not necessarily a problem, especially when well trained programmers are not available.

Communication and coupling between models and GIS products

Hydrologic analysis and modeling involves solving different mathematical equations that describe the runoff, flow or quality of the water. These equations are not usually available as internal commands of a GIS product, so they have to be coupled with the GIS. This can be done in three ways (Maidment 2002):

- Using the internal commands available in a GIS (**intrinsic modeling**). These models use the internal functions available in a GIS application for the calculations necessary in modeling. This type of coupling does not require advanced programming or UI design knowledge, but it requires experience and knowledge about the functions available in the GIS application used. An example for this type of model coupling is the calculation of runoff maps from rainfall maps using the raster calculation functions available in any GIS software.

- External commands from dynamic link libraries (DLL). The equations describing the hydrologic processes are implemented in a separate code module, like a DLL file, that can be written in different programming languages. The dynamic link libraries are coupled with the GIS when the functions implemented in these libraries are called, rather than when they are compiled. Functions implemented in dynamic link libraries can be called from the user interface of a GIS product and parameters are used in transmitting the data. The communication between

the GIS product and the library is possible using these parameters, and this method is faster than creating an independent application for the model. The models coupled with the GIS in this way are called **tightly coupled** models.

- Independent modeling. The model runs as a standalone application and the only link between the GIS and the model is at the data level. The data is passed between the GIS and the model as text or xml files (or any other file format that can be interpreted without GIS functionality). This is the most common way that GIS interfaces have been created for hydrologic models, and the model does not depend directly on the GIS product used. Still, the GIS has an important role in creating the input datasets and displaying the results. Independent models offer the programmer a wide range of possibilities because they are not limited in any way by the functions available in a GIS product, they can do a lot of processing on the data and can solve different requirements not directly related to GIS or spatial data. The independent model can have a dedicated GUI (Graphical User Interface) that does not use components from the GIS user interface. These kind of models are called **loosely coupled** models because they are only coupled at data level with the GIS product.

In each case, the hydrologic data needed for modeling is stored in a specific GIS data model. Even if the same type of data can be represented in different ways, there is no link between a specific model coupling way and the usage of a specific data structure. The coupling of models with GIS differs just in the type of model construction, not in the data storage technology. (Maidment 2002)

The graphical modeling solutions presented allow for creating user interfaces with components from the main user interface of the GIS used. The main problem of the models created using this technique is sharing. Even if the software offers the possibility of saving and sending the created models through email or other services, the models are limited to the specific product for which they were created. SEXTANTE makes an exception from this rule because any model created in SEXTANTE can be used in a number of Open Source GIS products written in Java.

The problems involved with sharing models between analysts using different software products calls for a solution allowing easier transfer of models between different products. This can be achieved by dropping the graphical modeling tools in favor of scripting in Python.

The interoperability between GIS products using scripts

Besides being easy to use, the Python programming language is open source and was created for easy integration with other programming languages like C, C++ or Java. This is why any GIS product written in one of these languages can create a Python library with wrappers for the internal geoprocessing functions of the product. Many libraries and tools have already been developed for working with GIS data in Python. (Butler 2005). Some GIS products also offer a Python programming interface that can be used when a more performant development environment is not available.

The Python programming language allows for using libraries coming from multiple sources, so tools from different GIS products can be used in a single Python script if some analyses require functionality from more than one GIS.

Conversion between data storage formats can be achieved using the open source **OGR library**. The OGR library can convert between a large number of vector formats, including: shapefile, ArcInfo coverage, spatial databases like PostGIS or Oracle Spatial and others. The OGR library is developed along with the **GDAL (Geospatial Data Abstraction Library)** and can be downloaded along with this library.

For raster formats, the GDAL library also has a Python programming interface. GDAL is the equivalent of OGR for raster operations and is able to convert and work with a large number of raster formats like: JPEG, TIFF/GeoTIFF, ERDAS IMAGINE, Enhanced Compressed Wavelet (ECW), Geographic Analysis Resource and Support System (GRASS) and others.

GDAL and OGR are actually two different libraries, but they are found under the same compiling system for historical reasons and because they are maintained by the same programmer. (Ramsey 2007).

Raster calculations (Map Algebra) can be obtained with GDAL along with the **NumPy** (Numeric Python) library. The NumPy library allows for complex raster operations on any of the GDAL supported raster formats. (Oliphant 2006) NumPy is a Python library allowing the manipulation of multi-dimensional matrices with a large set of mathematical functions. NumPy is

open source and was developed from the Numeric library created by Jim Hugunin in 1995 (Dubois et el. 1996). The project in which the NumPy library is developed also contains the **SciPy** library with functions specific to mathematics, engineering and other sciences.

Another open source library that is very important and widely used is the **PROJ.4** library (Urbanek 2007), that also offers a Python interface. This library allows for projecting datasets between any of the over a thousand spatial coordinate systems available in the EPSG standard (European Petroleum Survey Group)(D'Hont 2007) and also defining new coordinate systems.

The R programming language, used for applications in statistics, also offers different Python interfaces. The user can choose between **RSPython** and **RPy** to create Python scripts. These libraries can work with all the types of objects available in the R programming language and the majority of functions (even the graphical ones). The errors in the R language are converted to Python exceptions and any module installed for the R programming language can be used within Python (Churches 2011).

These are some of the main open source libraries working with GIS data that have a Python interface. Besides these, there are a lot of other libraries and most of the GIS products nowadays offer a Python programming interface (ex: GRASS GIS, QGIS, SAGA GIS, ArcGIS, MapInfo, IDRISI, PCRaster and others).

The great number of Python libraries for working with geospatial data raises the question: "Can I use functionality from multiple libraries in a single application?"

Like in any programming language this is possible and also quite easy. The "import" command is used to make all the methods present in a specific library available for use in a script. (Lutz 2010) The user can use methods from different GIS libraries together by importing multiple libraries in the same script.

The storage format for spatial data is different in different GIS products so the conversion problem appears. There is a variety of formats for representing the same data in different GIS products. When using tools from multiple products in a single model, conversions may need to be made between the specific formats for each of the products used. The conversion libraries

presented above (OGR, GDAL) along with NumPy allow for conversion between different raster and vector formats without the need to know the implementation details of each format.

There are some formats that cannot be directly converted using these libraries, but these are proprietary formats specific for a single product (ex: the ESRI ArcInfo GRID format or any kind of data stored in a geodatabase spatial database). In most of these cases, the conversions can be done by using intermediate formats or tools specific to a certain GIS product.

The possibility of using tools from different GIS products in a single script makes Python scripting a very important instrument for the future development of geospatial applications. New technologies like web services are also supported by Python and the online documentation is quite rich and usually available online for new users.

Implementation of the rainfall-runoff model from this study

The rainfall-runoff model that makes the subject of this study can be implemented using any of the methods presented above. After an evaluation of the possibilities along with the advantages and disadvantages of each method I arrived at the conclusions presented as follows.

The coupling between different GIS products using scripts can be done in order to use the functions available in different GIS products if needed. However, this coupling has some significant disadvantages:

- The needs for data conversion between formats used by different GIS products create unnecessary significant processing. This can slow down the model and make the task of cleaning temporary data without losing any important information quite difficult.
- The Python scripting allows for customization of GIS functionality at almost any level starting from raster cell level up to batch dataset processing. Even if some tasks can be more difficult to implement directly as source code, due to the open source character of Python and different geospatial online communities a lot of functionality was already implemented and the source code is available with open access.
- The storage of all data in a single database structure greatly enhances the processing speed and reduces the time needed for model application. Each different GIS product has its own data structure and the possibilities of using a single structure and more

products are not many. Therefore, the usage of a data structure specific for hydrological modeling like the ArcHydro data model has several benefits.

- The scripting libraries from different products update regularily, especially for open source products. Even if the updates offer increased functionality the scripts need to be kept up to date with these libraries and the development process for every library used should be taken into account.

Due to these reasons I chose to implement the model using the ArcHydro data model and the ArcPy library provided with ArcGIS. The latest version of the ArcHydro data model (version 2.0 in 10 Dec. 2011) offers a large number of new geoprocessing tools that can be used in modeling.

The graphical modeling is a good option for simpler models and some of the tasks are implemented as graphical models. These tasks are mainly the tasks that need to be run once and do a clear workflow without iterations, mainly the initial database processing. The part of the processing that was only implemented as a graphical model is the spatial discretization of the catchments presented in chapter 4 and used in initial database processing.

The rest of the model is implemented as Python scripts. The scripts use the arcpy geoprocessing library available with the ArcGIS Desktop 10 software. ArcGIS 10 Desktop introduces ArcPy, which is often referred to as the ArcPy site package. ArcPy provides Python access for all geoprocessing tools, including extensions, as well as a wide variety of useful functions and classes for working with verifying the integrity of GIS data. ArcGIS Desktop is needed to run the scripts, but the scripts can be run in two ways: from the desktop environment or from a Python shell (the command line).

The only tool that is not available in most of the GIS products due to its complexity and the uncertainties involved is the calculation of the runoff speed (flow velocity). Different methods exist for this calculation but most available GIS tools do not offer a continuous result over the terrain or raise other problems in speed calculation. A module available in the open source SAGA GIS and presented in the next section will be used for runoff speed calculation. The module needs user intervention in the choosing of the outlet from the map so calling the method from a script

was not possible.

Complex data processing, interpolation methods for discharge values and graphical representation of continuous discharge as a plot were not possible using the arcpy library so the MATLAB matrix processing software was used for this purpose. The scripts writtern for these tasks use the MATLAB scripting language and are available as .m files usable in MATLAB. The Matrix Laboratory package referred to as MATLAB was originally designed to serve as the interactive link to the numerical computation libraries used by engineers and scientists when they were dealing with sets of equations.Today, MATLAB is a computer language designed for technical computing, mathematical analysis, and system simulation. (Kalechman 2009)

Over the years, MATLAB evolved creating an extended library of specialized built-in functions and some of these can be used to generate among other things two-dimensional (2-D) and 3-D graphics.

The coupling between the different products used in the implementation of the model can be seen in Fig. 2.16
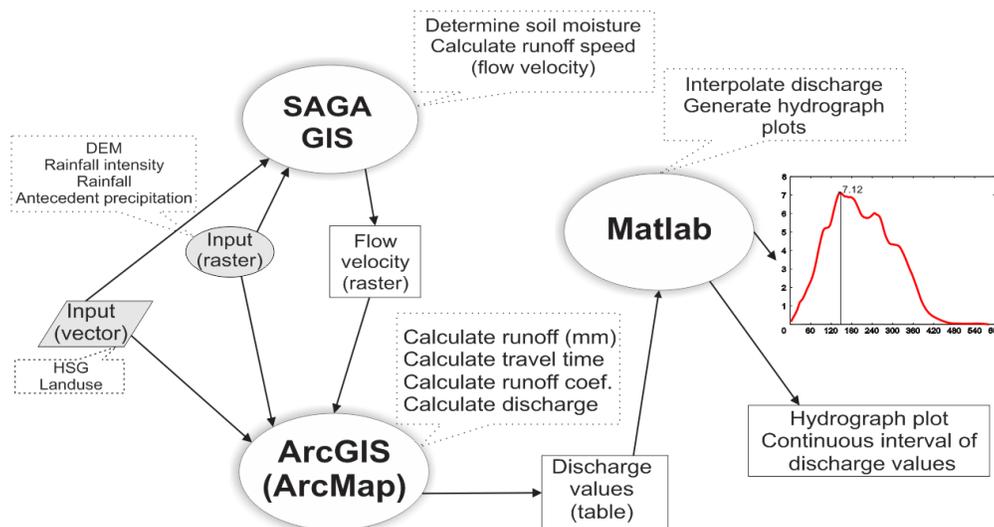
Fig. 2.16 *Workflow and programs used in hydrograph generation*

101

2.8 GIS functions created for surface runoff modeling

The tools presented in chapter 5 are the main functions used for primary manipulation of hydrologic data that are used in any hydrologic model and are not specific for runoff models.

Besides these standard functions, there are others created especially for modeling the surface runoff. Two of the most important problems that can appear and cannot be solved by using the standard instruments are the runoff speed (and associated time needed for water to flow through the entire basin until its outlet) and the progressive flow accumulation.

The **runoff speed** is a very important parameter used to get information about the time distribution of runoff from a basin. (Olaya 2004).

V. Olaya implemented two algorithms for runoff speed calculation in SAGA GIS. The first one, which is easier to implement, is called **Isochrones - Constant Speed** and considers the runoff speed constant throughout the watershed. A speed is given as a parameter and this speed is used everywhere in the basin as the runoff speed. This algorithm is not so complex and the same result can be used by multiplying the result of the flow length function with the time needed to traverse each cell. This algorithm is useful for the calculation of the unit hydrograph according to the real distribution of runoff throughout the basin, but is not suitable for obtaining discharge or other results needing more exact data.

The second algorithm, more complex and very useful, is called **Isochrones - Variable Speed** (Olaya 2004). The algorithm is based on an adaptation of the well known Manning formula presented in the work of Martinez et al. (2000). The result of applying this algorithm is the runoff speed in each cell from the watershed raster representation and the total time to the outlet from each cell (Fig. 2.17).
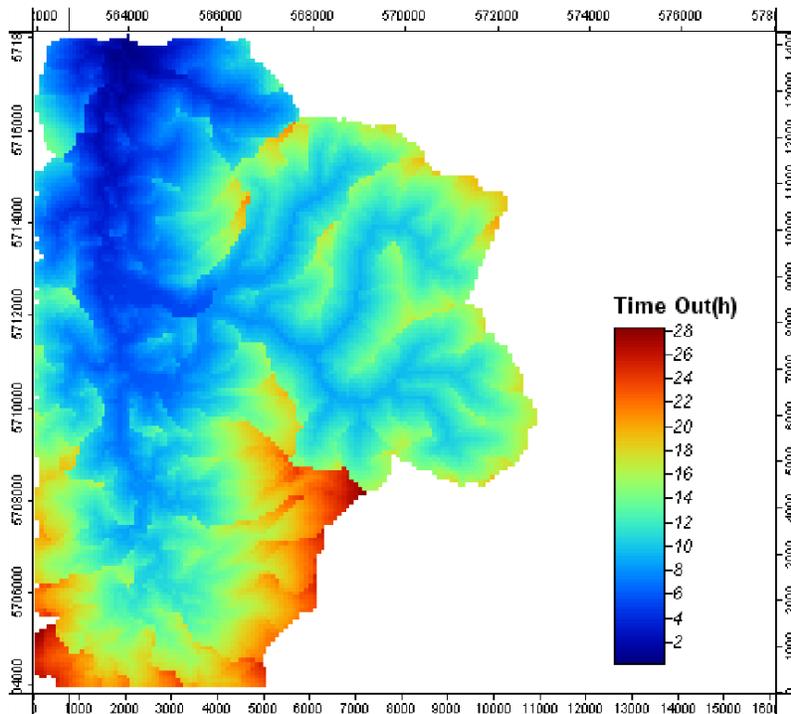
Fig. 2.17 *Flow time calculated in SAGA GIS with Isochrones - Variable Speed*

The algorithm calculates the runoff speed in every cell of the watershed according to the characteristics of the terrain. The constant speed from the previous algorithm is replaced by the speed obtained through these calculations. The runoff speed in each cell depends on different characteristics of the terrain represented by parameters (slope, estimated quantity of water in the cell, terrain roughness, the Curve Number coefficient (section 1.5). The Curve Number coefficient (a non-dimensional coefficient used for determination of runoff depth) and the terrain roughness coefficient (Manning's n, an non-dimensional coefficient used in speed calculation) are calculated through a spatially distributed algorithm that I implemented and also used in the runoff model presented in this work. The mean rainfall intensity (mm/min) also influences the runoff speed and is given as a numeric parameter. This intensity can be calculated for past rainfalls from the duration and the runoff depth measured in the catchment.

Other parameters that the user has to specify are the thresholds where the runoff becomes mixed flow and the threshold where a channel is defined. These parameters have standard values calculated from the characteristics of the area and the user does only need to change this in special cases when more information on the area creates the need for other parameter values.

The last two parameters used in the calculations are the slope of the channel sides and the minimum runoff speed. The channel is considered triangular and the user can choose the slope of the channel sides that will be used in the calculations. The minimum speed is used when the surface slope is very small or there is insufficient runoff and the calculations return a value of zero for speed.

Another phenomenon that can be described by mathematical equations if the **flow variation through time.** The runoff is a dynamic phenomenon that can be described by differential equations (the Navier-Stokes and the Saint-Venant or shallow water equations) that can be solved through approximate methods.

The Navier-Stokes Equations are a general model which can be used to mod el water flows in man y applications. However, when considering a specific problem such as shal low-water flows in which the horizontal scale is much larger than the vertical one, the Shallow Water Equations will suffice.

The basic assumptions for the analytical derivation of the Saint Venant Equations are the following (Elisa Aldrighetti 2007):

• the flow is one-dimensional, i.e. the velocity is uniform over the cross section and the water level across the section is represented by a horizontal line

• the streamline curvature is small and the vertical accelerations are negligible, so that the pressure can be taken as hydrostatic

• the effects of boundary friction and turbulence can be accounted for through resistance laws analogous to those used for steady state flow

• the average channel bed slope is small so that the cosine of the angle it makes with the horizontal may be replaced by unity.

The solving of the saint-Venant equations for runoff on the irregular land surface is implemented in the **GRASS r.simwater module**. The result of this module is a time series of maps with the runoff variation in the field. Optionally, the method can return time series of discharge in certain points represented in a point layer.

The module calculates the depth of surface runoff or the discharge according to data on the terrain characteristics and rainfall. The terrain characteristics can be obtained from the DEM and

used later as parameters for the module. The main parameters for the module are: rainfall excess, considered spatially distributed but time invariant, flow gradient vector given by first-order partial derivatives of elevation field (dx, dy parameters), a surface roughness coefficient given by Manning's n and the infiltration rate in mm (Neteler, Mitasova 2008). The flow gradient Partial derivatives raster files can be computed along with the interpolation of a DEM using the -d option in v.surf.rst module. If elevation raster is already provided, partial derivatives can be computed using r.slope.aspect module. Partial derivatives determine the direction and magnitude of water flow and can be modified to include predefined water flow in channels.

An example of maps returned from this functions after 200 (a) and 2400 (b) seconds can be seen in Fig. 2.18
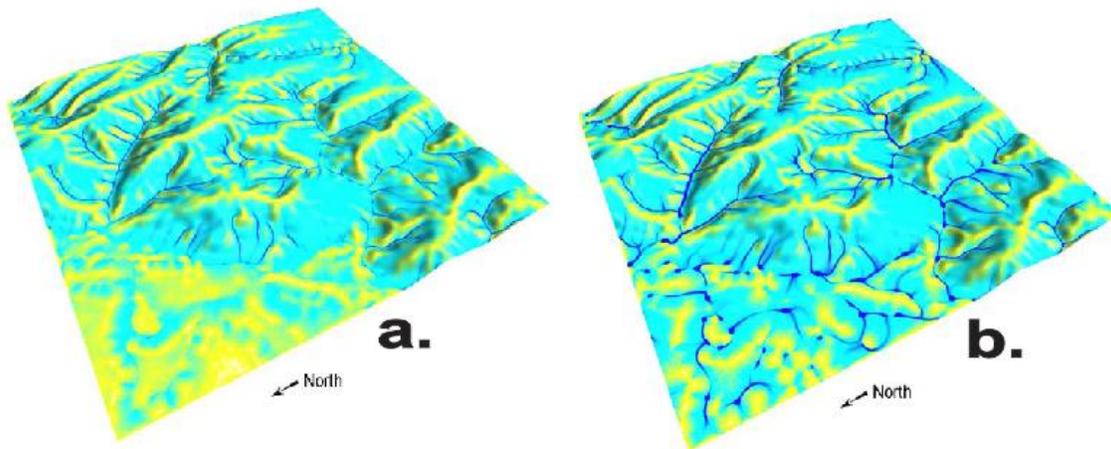
Fig. 2.18 r.sim.water simulation after 200 sec(a) and 2400 sec (b) (from Hofierka et al.. 2009)

Other software products that allow for hydrologic and hydraulic calculations are the HEC-HMS and **HEC-RAS**, developed by the US Army Corps of Engineers Hydrologic Engineering Center (HEC). HEC-HMS (Hydrologic Modeling System) is a model used to simulate the rainfall-runoff processes in large catchments and urban areas. HEC-RAS (River Analysis System) is a model developed for hydraulic calculations in a complete network of natural or man-made channels.

HEC-RAS includes components for one-dimensional analysis for: water surface profile in the case of uniform flow; turbulent flow simulation; sediment transport calculation; water quality analysis. The water surface profile is very useful in determining the water level in case of floods

and its calculation using a tested model leads to better results in a shorter time. The calculation of the water surface profile using a tested model is good for obtaining faster results with less effort in case of a flood.

HEC-RAS is an independent program with no GIS coupling. It is the implementation o f well-known model from the '60s named HEC-2 using object oriented programming. Along with the development of GIS technology in the '90s, GIS interfaces were built for conversion and transfer of data between HEC-RAS and the ArcGIS product (Maidment 2002). The necessary data can be stored in an ArcHydro database as presented in section 2.2.

The last version of the GIS interface for ArcGIS is called HEC-GeoRAS. Using the HEC-GeoRAS interface, the HEC-RAS model is loosely coupled to ArcGIS. All the operations besides the creation of the database and the presentation of the results are made outside the GIS.

The HEC-GeoRAS interface allows for creation of a specific spatial database and for conversion of necessary data from the ArcGIS formats to the format needed for HEC-RAS. After the HEC-RAS processing, the result can be converted to GIS datasets and presented in ArcGIS.

A very important result that can be obtained using HEC-RAS is the map of **flooded areas**. Also, the water surface profiles and data on the runoff speed can be used to estimate the damage caused by flooding and preparing the procedure of intervention in case of extreme events (Merwade 2010).

HEC-GeoRAS uses the DEM to extract important information about the application area. The user creates a series of line layers used to obtain the geometric data required by HEC-GeoRAS. These layers include: the central line of the streams, stream segments and the connections between these segments, stream banks, cross-sections, hydraulic structures and other physical attributes of river channels. The pre-processing using HEC-GeoRAS involves creating these attributes in GIS and exporting them to a HEC-RAS geometry file.

The accuracy of the data is very important, so high resolution satellite or aerial images are recommended when creating the input datasets. HEC-GeoRAS offers an easy to use interface for creating all the input datasets and offers some useful functions that help the user obtain these data (for example, creating equidistant cross sections along a river) (Fig. 2.19)
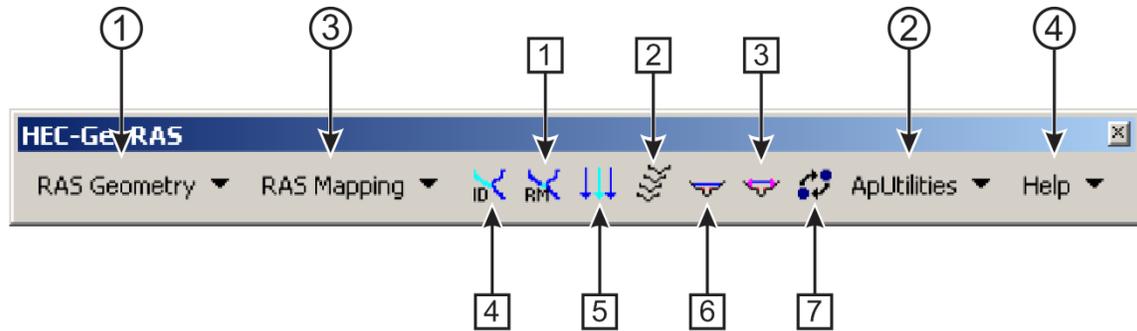
Fig. 2.19 *– HEC-GeoRAS toolbar*

The HEC-GeoRAS toolbar contains four menus (RAS Geometry, RAS Mapping, ApUtilities, Help) and buttons for seven tools (Assign RiverCode/ReachCode, Assign FromStation/ToStation, Assign LineType, Construct XS Cutlines, Plot Cross Section and Assign Levee Elevation).

The RAS Geometry menu contains functions for preprocessing data sent to HEC-RAS. The RAS Mapping menu contains functions for postprocessing the results obtained from HEC-RAS and creation of flooded area maps. The apUtilities are tools for organizing input and output datasets, and the Help menu is used for assistance.

Besides the basic operations used for creation of cross-sections, streamlines and floodplain, HEC-GeoRAS has other tools used for creating bridges or areas without flow (buildings in the water), areas with water but zero flow speed (behind the bridge piers) or automatic Manning's n extraction from standard terrain datasets.

After the preprocessing, the user verifies the input data in the HEC-RAS interface and can do the last corrections and adjustments on this data (for example, filtering redundant points from cross-sections or entering the dimension and number of bridge piers). The maximum discharge calculated for each subsection of a river is input and the HEC-RAS model executes the simulation.

The HEC-GeoRAS interface is again used to import the results back to ArcGIS and present them in different ways like flooded areas, flow speeds or maximum water levels. An example representing a flooded area polygon can be seen in Fig. 2.20
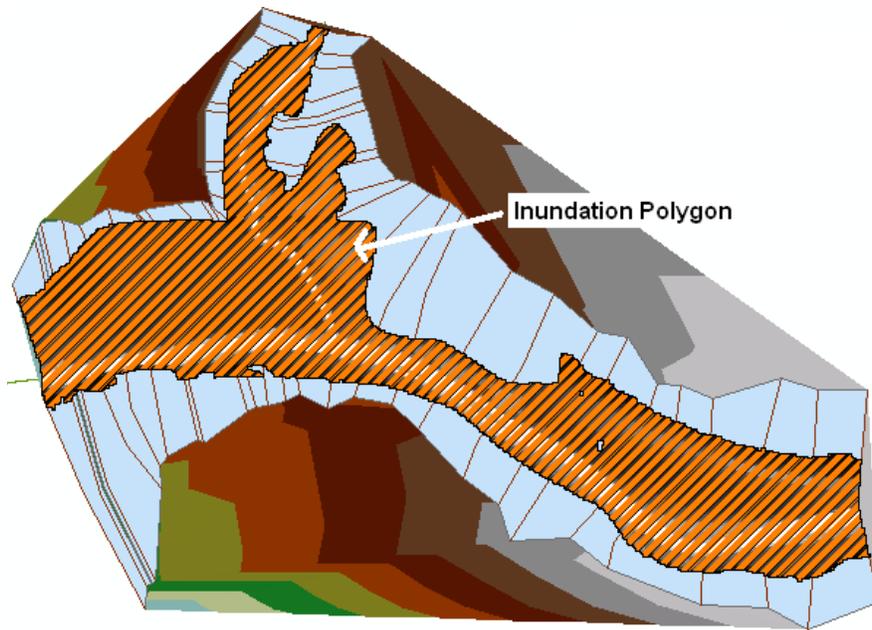
Fig. 2.20 *Flood area calculated with HEC-RAS*

Usage of the functions in the modeling process from this study

The runoff speed calculation and the travel time to the outlet of the catchment is a very important parameter that needs to be calculated . These functions aid in the calculation of this parameter and they will be used in the final implementation of the model.

Depending on the data available some of these functions can be easier to apply in a simulation or analysis process than others. The model I create has to be realistic in terms of data availability so some of these functions will not be used now.

3. Database construction

3.1 Data necessary for surface runoff modeling using GIS

The surface runoff modeling can be done using different datasets, but some of these datasets are needed regardless of the runoff modeling mechanism used. A good runoff model has to answer some very important questions to be useful to local authorities in building flood protection infrastructure or to evaluate damage from extreme events.

The first and most important question is: "Where does the water flow?" To answer this question using a GIS, the user needs a **digital elevation model** of the area. These datasets will be used to calculate the **flow direction** and **flow paths** using the functions presented in chapter 2.5 The result of this stage is represented by the maps of flow direction (the direction of water flow from every cell representing the study area), **catchment area** (the surface of the basin from which water flows through each cell) and **streams** (lines representing the paths of water flow).

The second question, more complex and requiring more complicated operations, is "How much water flows?". To answer this question, one needs to consider the used model and obtain the data needed to apply the model. Usually these data include what is on the surface of the ground and data about the soils: **soil description, land use/cover, roads, areas for water accumulation**. If the groundwater flow is considered, data can also include **hydrogeology characteristics** of the terrain, **soil humidity** and description of **the aquifer**. The result of this stage is the quantity of water available after rainfall in each cell (or subarea of the study area if working with vector datasets).

The third question, which gives the final result and is the most important part of a rainfall-runoff model is "What is the discharge generated by the runoff in every point of interest on the flow path?" Usually, this is solved by routing the water available for runoff obtained earlier using a known routing method. A linear translation according to runoff speed can be performed using the flow direction raster calculated earlier and a **flow velocity** raster calculated using a tested method like the one available in SAGA GIS presented in chapter 8. The time required for water to cross each cell is obtained by dividing the length of flow through the cell with the flow velocity.

The time of flow to every interest point can be then obtained by summing the time of flow through each cell on the flow path.

After the determination of the flow path and travel time from every cell of the watershed to the points of interest, the time distribution of the runoff generated in the catchment can be obtained by cumulating the quantities of water available for runoff according to the distribution in time. A last step is the interpolation of runoff values to obtain a hydrograph and the estimation of the moment when peak flow will be expected in each interest point.

The data needed to model the runoff according to the structure presented earlier are available in different forms and some of these forms are presented in the next section.

3.2 Altimetric data (DEM)

The relief has a major impact on the evolution of runoff processes and digital elevation models are very important in any spatially distributed hydrologic analysis.

The main factor determining the runoff of eater and its accumulation in channels causing floods, the altimetry is a essential dataset in runoff modeling. All the runoff processes depend on the movement of water due to gravitation, movement that can be modeled when knowing the topographic structure of the terrain.

The digital elevation model is also important in determining the slope of the terrain and flow direction, which are later used for delineating drainage basins corresponding to a measuring gauge or needed for an analysis.

The shape of the terrain is a surface that varies continuously into space and that can be symbolized using contours in a plan. Any digital representation of the continuous variation of altitude in an area is called a **digital elevation model (DEM).** Other characteristics with a continuous variation in space can be modeled with the methods applied to altitude: pressure, temperature, pollution, depth of soil layers etc. In this case, the representation of any thematic variable Z on a continuous area is studied (Niţu, 1992).

DEM data structures

The altitude data is usually organized in one of three data structures: raster (or GRID), TIN or isolines (contours). Each of these structures has a usage specific and is associated to a certain way of processing the data.

Raster modeling of the DEM

Raster models discretizing the surface in square cells disposed as a grid is a very popular structure because of its simple memory representation and the ease of implementation for analysis tools on the data.

A raster is a rectangular grid of equally sized cells representing thematic or spectral data. Raster data can represent anything starting from the qualities of a terrain like altitude or vegetation to satellite images, scanned maps or aerial photos (Zieler 2001). In the case of the DEM the modeled characteristic is altitude and the value in each cell contains the elevation of the terrain in the area represented by the cell.

The raster digital elevation models are the most used type of model because their representation is extremely simple (a simple matrix where the topological relations between the cells is implicit) and for the ease of representing them in memory (Moore et al 1991).

The advantages of raster data representation include **(Satheesh 2007):**

- The geographic location of each area is given by iots position in the matrix; therefore, except for an origin point and a cell dimension, no geographic coordinates are stored in the elements of the raster
- Due to the storage specific, the programming of spatial analyses functions can be implemented with ease and some analyses run very fast on the raster model
- The nature of raster data is suitable for mathematical models and quantitative analysis because every cell in the raster contains a single numeric value for use in calculations
- Zonal data like forests can be represented as easy as continuous data like altitude and the two types of data can be integrated in analysis. More datasetc covering the same area can easily take part in a spatial analysis operating on all datasets

The main disadvantages of raster data structures are:

- The dimension of the cell determines the resolution at which the data is represented. Any information with a apatial dimension smaller than this resolution cannot be represented in a raster.

- Linear structures are difficult to represent due to the rectangular shape of the cells, so networks of linear features are very difficult to represent. This can affect hydrologic analysis because of the linear nature of streams.

- Processing of associated attributes can be difficult at a large wuantity of data. The raster data structure can only represent one characteristic of the terrain and any other data must be stored separately and linked to a raster velue or cell position. This does not affect the analysis only needing a digital elevation model (altitude values)

- Any vector data must be converted into a raster for compound analyses and this fact can affect data integrity because of generalization or wrong choice in cell dimension

- Even if there the characteristic represented is not present in a certain location, the memory of the computer is occupied by a specific value called a NoData value. In the case of altitude, any point on the surface of the Earth has an altitude so this does not apply.

Practically, a raster DEM is composed of points distributed at equal distance represented as numbers in a two-dimensional matrix. The rectangular matrix where the data is stored allows for the storage of any number of rows (equal to the number of cells in the model on the Y axis) and columns (equal to the number of cels in the model on the X axis). Each cell in the matrix has the X and Y coordinates known from the line and column of the raster and contains a single value representing altitude. There are other representations that store the geographic coordinates or the distance between cells on the horizontal and the vertical (to be able to create rectangular rather than square cells), but these structures are rarely used in practice.

The resolution of a raster is the smallest surface on the terrain that can be represented with an independent value in a raster.

Fig. 3.1 presents an example of a DEM at different resolutions.
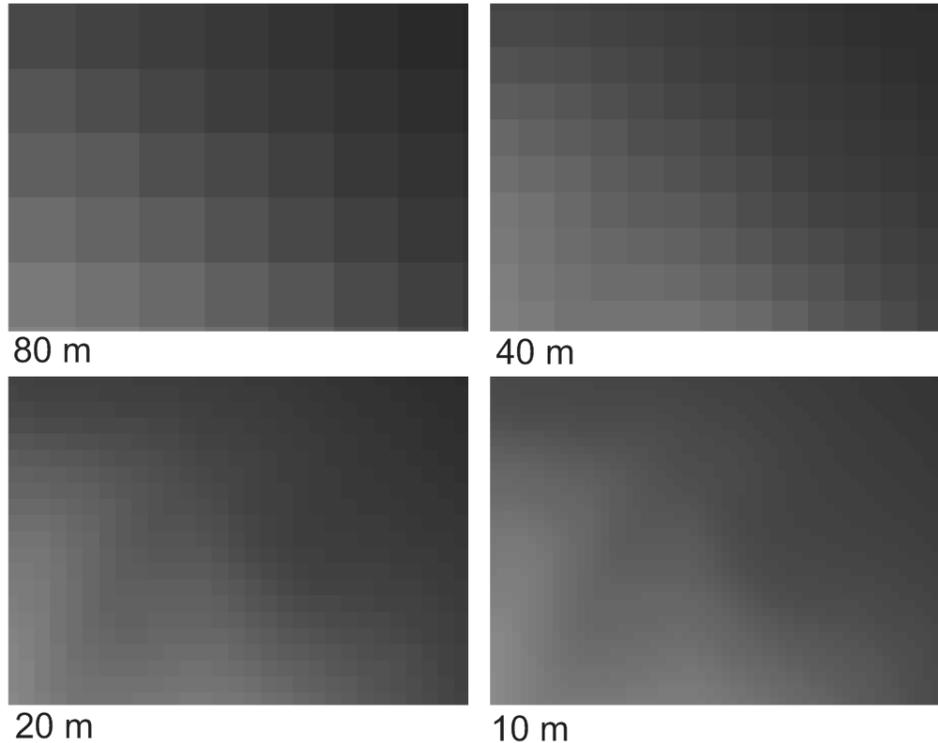
Fig. 3.1 *The different in detail due to the storage of a DEM at different resolutions (from the DEM of the same area from the Râşca catchment)*

The majority of raster file formats include information related to the cell dimension, the coordinates of a corner X_0, Y_0 (usually the lower left corner), information on the coordinate system, number of lines and columns, etc stored in the header of the raster file or sometimes in a separate file.

After these characteristics of the raster the list of values for points follows in the file.

Even if a raster DEM has the disadvantage of storage space occupied in memory, the spatial analysis possibilities make this format suitable for hydrologic analysis.

TIN - Triangulated Irregular Network modeling of the DEM

The triangulated irregular network is a structure a little more complex than a grid composed from a network of triangles connected on the edges. The slope of the ground is considered constant in every triangle and the triangles have different dimensions according to the altitude variation. The areas with little variation are represented by larger triangles and the areas with more variation are represented by smaller ones in order to have a better representation of the

ground shape. (Fig 3.2)

The Delaunay triangulation for a set of points presumes the calculation of triangles that respect the criteria invented by Boris Delaunay and have vertices in every point of the set. In mathematics and computational geometry the Delaunay triangulation for a set of points in a plane is the determination of triangles with vertices in the points so that no point P is in the circumcircle of any triangle from the triangulation. The Delaunay triangulation was discovered in 1934 by Boris Delaunay (Delaunay 1934) with the scope of maximizing the minimum angle of all the angles of the triangles in the triangulation in order to avoid skinny triangles. The triangulation can be extended in three or more dimensions by considering circumscribed spheres.

If every vertex of a triangle obtained by Delaunay triangulation has an altitude value a TIN digital elevation model is obtained. The Delaunay triangulation is presented in Fig. 3.2 for a set of random points. The Thiessen polygons in which any points in the polygon are closer to one of the points in the initial set than any other point in the set, can be obtained from the perpendicular bisectors of these triangles.
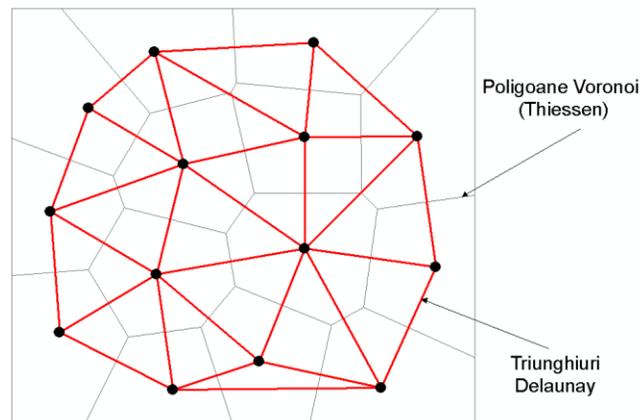
Fig. 3.2 *Voronoi Polygons and Delaunay triangulation.*

The data structure of a TIN is based on an algorithm that contains vertexes, sides and their geometry information (X,Y,Z coordinates of the points) and topology information (neighboring triangles and edges connecting them). All the triangles of a TIN respect the criteria of Delaunay triangulation.

The vertices from the original data used to generate the TIN are connected by edges to form a structure of triangles respecting the criteria presented above. An example of the TIN generated from contours can be seen in Fig. 3.3

114
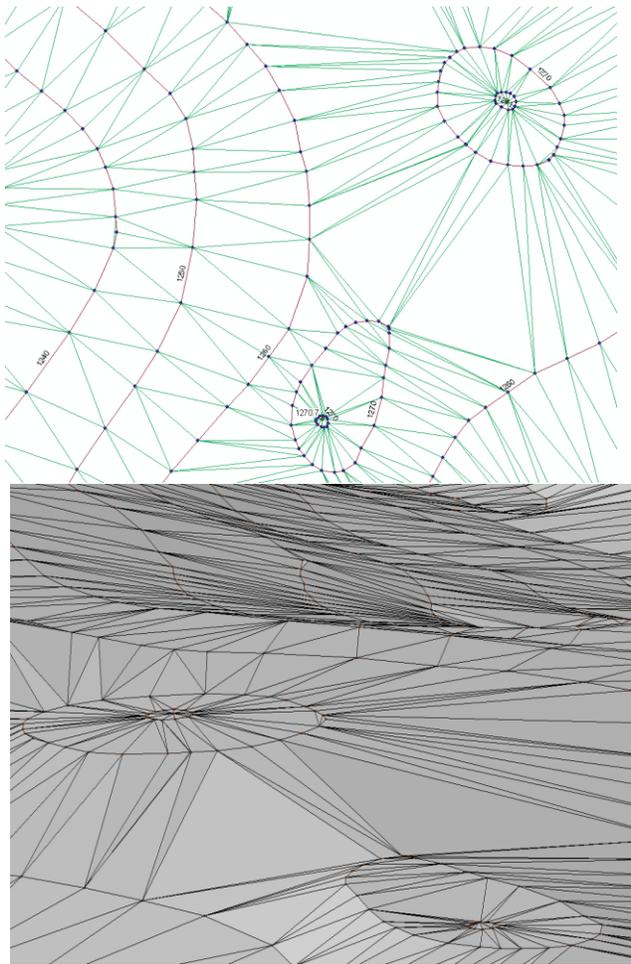
Fig. 3.3 Representation of the surface by a TIN în 2D (left) and 3D (right)

Due to the structure, only the coordinates of the points used to generate the TIN are stored along with the connections between the neighboring points. The TIN model can easily represent discontinuities in the shape of the terrain like peaks or sinks and has an efficient storage algorithm because the dimension of the triangles can vary according to the irregularities in the terrain shape (Moore et al. 1991). This can be a disadvantage in analysis when the terrain is flat or uniform. To avoid the problems due to representation, some software products include the possibility of setting certain structures that are definitive (hardlines) in order to have a correct representation of features like roads, rivers or banks.

The main advantage of a TIN is the efficiency in the storage because only the information for minimum number of vertices to correctly represent the result of the interpolation is stored in

memory.

Other advantages include the possibility of adding features to the TIN from more sources, either points or lines (ex: contours, points with altitude measurements), without the need to make multiple interpolations like in the case of a raster DEM. Due to the structure of a TIN new points can be added of the data can be updated in an area without the need to recalculate all the structures in the model.

The main disadvantage of this structure is the difficulty in addressing the independent primary components of the network from the computer memory (points, edges, triangles).

Another disadvantage that limits the use of a TIN in hydrologic modeling is the impossibility of making an analysis on multiple layers. Any analysis requiring other data besides altimetry becomes difficult due to the different representation of data in memory.

Even if the debate regarding the use of raster or vector models in geomorphometric analysis is still unsolved, the majority of application use the raster models (Hengl et al 2009).

Modeling through contours

This type of modeling presumes the representation of altitude through isolines connecting all points with the same altitude (contours) (Fig 3.4) at constant altitude intervals.

The main advantage of this structure is the possibility of obtaining the data directly through digitizing existing topographic maps. The data obtained like this can contain more details than a free DEM available from the internet, but the accuracy of these data depends on the scale of the map being digitized.
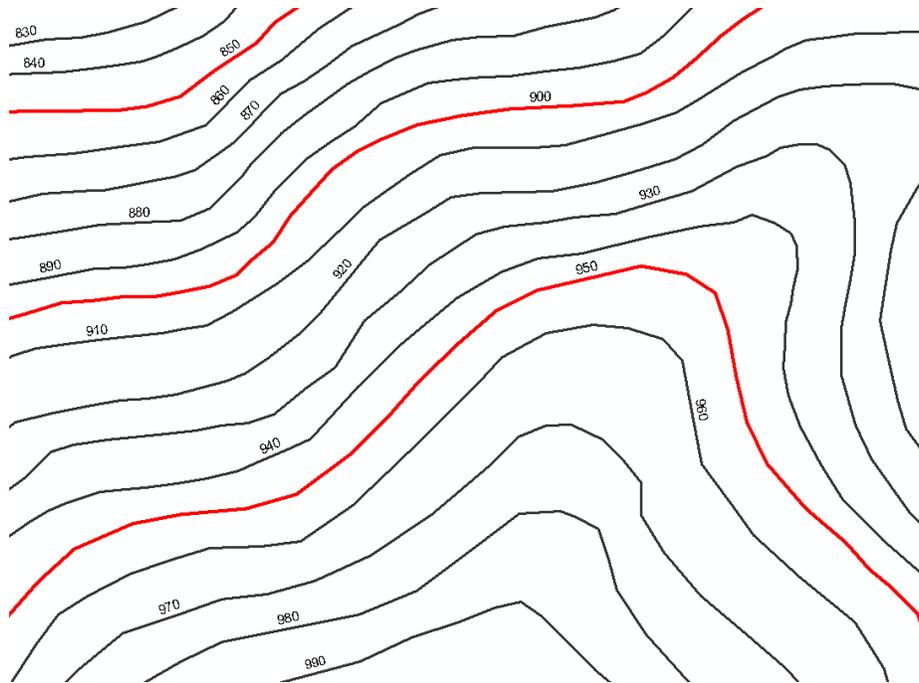
Fig3.4 Altitude representation through contours

This structure cannot be considered a DEM because it does not represent the continuous variation of terrain altitude but only discrete data on the contours. This structure is very good in cartographic representation but it is not useful for hydrologic or other spatial analysis.

The contour representation of the terrain surface is important because it helps in obtaining the other types of DEMs through interpolation of contours. Besides these contours, the maps may also contain points with altitude measurements which can be used in interpolation. If the maps being digitized have a large scale or topographic surveys are used, this model allows for obtaining DEMs of higher resolution. Also, this model enables obtaining more accurate data for high scale digitized maps.

The possibility of obtaining the altitude data as points with altitude disposed at regular intervals also exists. Some data providers can offer altitude data in text files in this format with the equidistance of the points specified. This format is quite similar to the raster format and can be included in the first category of raster DEMs if every point is considered to be the center of a square cell. Some GIS products have a special raster format called an ASCII raster which is this exact representation of equidistant points disposed in lines and columns in a text file.

This format will not be detailed because most of the characteristics of a raster are also true for this format and the two formats are similar.

Obtaining the altitude datasets in a digital format

The majority of altitude datasets available are derived from data obtained through remote sensing (Moore et al 1991). These sources of data are based on obtaining altitude variation either by interpreting stereo aerial photography using automated or manual instruments, either by the direct recording of altitude using radar or other technologies presented as follows. Besides data obtained through remote sensing there are other data sources of altitude data used frequently.

The five main sources of altitude data used in the present are (Hengl et al. 2003):

- topographic surveys;

- aerial images;

- maps or existing topographic plans;

- aerial laser scanning;

- stereo satellite imaging or radar.

Altitude data derived from these well known datasets have the resolution and characteristics as follows (Table 3.1):

Table 3.1 – Altitude data resolution according to acquisition method

Acquisition data	Characteristics	Technologies used	Typical resolution
Direct field surveying	- Greatest accuracy - Low density of measuring points - High costs	Differential GPS Total station	<= 1 m 1 mm - 1 m
Stereo imaging	- High density of measurement points - Can be semi-automated or completely automated - Problems when vegetation exists	Aerial photography Satellite images (SPOT, ASTER)	0.1 m - 1m 10 m (20 m)
Laser scanning	- Laser scanner installed on GPS guided plane - Raw data needs filtering and scaling before use - Laser can pass through vegetation and record altitude on ground level	Airborne laser scanning (LIDAR)	~ 0.2 m - 1.0 m
RADAR scanning	- Lowest cost - Needs calibration data from the ground - Complex data processing	Airborne Radar (SAR - Synthetic Aperture Radar) Spaceship Radar (ERS, SRTM)	~ 0.5 m - 2 m 10 m (30 m)

Measurements in the field

The horizontal and vertical locations of points on the surface of the ground can be obtained with precisions at the order of millimeters. Measurements for obtaining these data can be made

through different methods.

The classical method is the usage of the theodolite (an instrument used to measure angles in a horizontal and vertical plane) and triangulation methods (calculations on the distance and angles between points) to create a network of triangles with measured points in every vertex. These triangles are then transformed into digital format and a high precision surface measurement is obtained.

Even if this method is not so expensive concerning the technology acquisitions and the process requires trained surveyors and a large amount of work for measurements. The development of electronic theodolites and total stations that make the calculations automatically reduced the amount of work needed for these measurements.

By processing the terrain measurements in the computer maps with accuracy in the order of millimeters can be obtained for large surfaces. The main problem in this method of obtaining spatial data is the high cost of workforce that is not justified in small projects.

Another measurement technique is the GPS. The GPS instruments are not that exact, but the development of differential GPS (DGPS) leads to a higher accuracy by using a GPS unit with known location that transmits the measurement error for correction in the position measured by the GPS. GPS manufacturers specify vertical and horizontal error of 4-20 m and 8-40 m for GPS, 1-3 and 2-6 m respectively for DGPS. In good conditions,m(ore than five satellites), these errors vary very little and a horizontal accuracy of less than one meter horizontally and 1-2 m vertically can be easily obtained (Hengl et al 2009). GPS measurements are faster; they just require covering the ground with a GPS receiver and recording points according to a defined interval.

Besides these systems there is also potential to use GPS systems for geodesy / surveying composed of a mobile GPS receiver and a rover with a known and fixed position. These systems have a high purchase cost but can provide horizontal and vertical accuracy of less than 5 cm.

Advantages of field measurement techniques are high accuracy (altitude can be determined with precision of about 1 cm or higher), flexibility (density measurements may vary depending on terrain and needs) and minimal post-processing after taking measurements in the field. Problems that make this method difficult to use for research are equipment price, the amount of work

required and the long duration required to achieve gather data in the field.

In the past national mapping agencies created topographical maps using these measurement techniques, but maps made by this method have been widely replaced by remote sensing methods in the current period. (Smith, 2005 cited by Hengl 2009). For Romania the most famous institution that created topographical maps is the Military Topographic Directorate (Direcţia Topografică Militară).

Because data taken from the plane with laser scanning systems (LIDAR data) are quite expensive and surveying large areas is difficult and raises various issues, a source of data with sufficient accuracy that can be used in any area, is not very expensive and has uniform characteristics throughout the country is needed. Data recorded by remote sensing and available for free on the internet can meet these requirements.

Data recorded by remote sensing

Perhaps the greatest progress in hydrological modeling over the past few years has been the public availability of land surface elevation data over the Internet or in other digital format (CD-ROM) and development of advanced data processing methods.

a. SRTM

The data obtained by radar, the second very accurate source, are available in many forms. One of the most popular examples for Romania is the SRTM (Shuttle Radar Topography Mission) program, which in February 2000 measured altitude for about 80% of worldwide land area, using a radar sensor mounted on board of the space shuttle Endeavour. The SRTM was the first set of global continuous altitude data at good spatial resolution: 1 arc-second (approximately 30 m), freely available to the U.S. and paid the rest of the world and 3 arc-seconds (approximately 60 m in Romania). The commercial data is distributed by NASA and the free data is available on the USGS website for free.

These data are available as raster data where each raster cell surface has a corresponding value of the land elevation at that point and cell size is equal to the accuracy of measurements.

b. ASTER

Another database that was created in a recent mission and is available for free is the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) database.

The **A**dvanced **S**paceborne **T**hermal **E**mission and **R**eflection **R**adiometer (ASTER) is one of many tools located on the Terra spatial platform, which was launched in December 1999. ASTER provides images in 14 spectral bands with a resolution ranging between 15-90 m.

ASTER images are available in Universal Transverse Mercator projection (UTM), but some scenes are stored in other projections. The data is referenced in the WGS 84 system

The ASTER remote sensing system consists of three subsystems covering the visible and near infrared (VNIR), medium infrared (Swire) and thermal infrared (TIR) wavelengths.

Details of spectral bands and their resolution can be seen in the table. Each of the three telescopes can be rotated by + / -24 ° in the lateral direction by tilting the whole system of telescopes.

The VNIR system records stereographic data and includes two telescopes, one at nadir and the other in the opposite direction of travel (backward). Because the images recorded on the same wavelength, they facilitate stereoscopic view of the target area (Jay Gao 2009). As can be seen in the table, the third band VNIR sensor, which is registered stereographical, is the most important for the extraction of altimetric digital model.

On June 29, 2009, the Ministry of Economy, Trade and Industry of Japan (METI) and NASA announced a new product, global ASTER digital elevation model (GDEM). GDEM was created by linking 1.3 million stereographical ASTER scenes and covers the entire Earth's surface located between latitude 83N and 83S. GDEM is produced using 15 m resolution stereographical scenes available as GeoTIFF files covering surfaces of 1x1 degrees on the ground. Each scene is accompanied by a GDEM quality control file that specifies the number of ASTER scenes used to calculate the value of each pixel or the external data source used to cover areas where ASTER data are missing, if any (ASTER GDEM) . The resolution of the GDEM for Romania is approx. 25m.

The accuracy of this freely available dataset is suitable for the representation of urban areas or other smaller areas, but in larger river basins and for other purposes data at a lower resolution may be enough (Table 3.2) (Maidment 2002).

Table 3.2 – Characteristics of the sensors in the ASTER system

Characteristics	VNIR	SWIR	TIR
Sensor and wavelength (μm)	1: 0.52 - 0.70	4: 1.600 - 1.700	10: 8.125 - 8.475
	2: 0.63-0.69	5: 2.145 - 2.185	11: 8.475 - 8.825
	3N: 0.76 - 0.86	6: 2.185 - 2.225	12: 8.925 - 9.275
	3B: 0.76 - 0.86	7: 2.235 - 2.285	13: 10.25 - 10.95
		8: 2.295 - 2.365	14: 10.95 - 11.65
		9: 2.360 - 2.430	
Nadir resolution (m)	15	30	90

Table 3.3 – Cell size in free digital elevation models and their applications

Geographic dimension of cell	Application	Linear dimension of cell	Catchment area (km^2)	Region area (km^2)	Zones of use
1"	Small basins	30m	5	1000	Urban
3"	Small basins	90m	40	8000	Rural
15"	Basins	460m	1000	200000	River
30"		930m	4000	900000	State
3'		5,6 km	150000	30000000	Continent
5'		9,3 km	400000	90000000	Global

Source: Maidment, 1996

To use more detailed elevation models commercially available data or manual creation of a DEM from topographic maps of large scale with elevation contours is necessary.

Data obtained by interpolating information on existing maps
DEM resolution

The raster format will be used in runoff modeling in this study because the majority of altitude data is available in this format and the format is suitable for hydrologic analyses.
The raster digital elevation model representation in a very important characteristic is the **resolution of the model**, which determines the accuracy of the data represented by this model. Depending on the resolution, the model can be used for various operations. The higher the resolution (smaller cell size), the more detailed the DEM is and more visible features of the landscape of the area are represented. The figure (Fig.3.5) shows the connection between resolution of a raster and terrain detail representation. A too low resolution (p = 2.5 m) will be less accurate in land surface

representation, whereas a higher resolution (p = 0.5 m) will better represent all peaks and valleys present in the terrain.
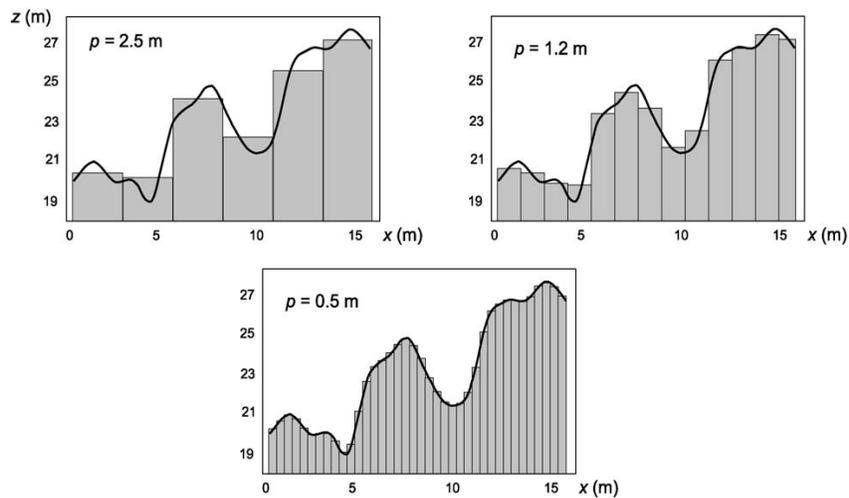
Fig. 3.5 *Characteristics of altitude representation according to raster resolution (Hengl 2003)*

The drawback in this is that the higher the resolution, the more memory space the representation of the same surface will occupy, causing the analysis to be more time consuming. The user interest is the usage of a resolution that offers enough detail for the necessity of the specific application but occupies the least amount of memory. In the case of hydrologic applications the resolution depends on the dimension of the studied catchment and the aspect of the terrain in the study area.

In the case of digital elevation models, appropriate resolution depends on the variability of land. In general, pixel size must be less than the average distance at which a change in the morphology of the land appears. In a region where altitudes have little variation, like in plains or depressions, very high resolution is not so important as resolution increase will cause an unjustified increase in the number of cells and therefore the size of the data set. Digital elevation models in the case, appropriate resolution depends on the variability of land. In general, pixel size must be less than the average distance that a change in the morphology of the land.

Resolution of a DEM depends on the source from which it is obtained and how the data is interpreted. If topographic maps are scanned, the raster size of a cell is closely related to the level of detail of a map. The relationship between cell size and scale on a printed map is very important in cartography, but it is not so important for generation of images from stereographical scenes. As a general rule in the mapping, the minimum size of an object that can be represented on the map

123

printed on paper is 0.5mm. Thus, a 1:25000 scale map corresponds to a raster resolution of 12.5 m.

Basic size of the cell which can be obtained from a particular data source is very important. A smaller size leads to unnecessary use of memory without obtaining a visible improvement in the results of the analysis, while a larger size leads to free memory space at the expense of accuracy in representation. As a raster resolution decreases, the content of information in a map and it will gradually decrease, and vice versa. (Stein et al. 2001)

Sometimes the choice of pixel size is limited by the processing power and scale of the application. For regional and national studies a very high resolution DEM is unnecessary and exaggerated.

The necessity of interpolation for discrete data and measurements

The raster format is a continuous and uniform representation of land surface in which every point on the surface belongs to a cell with known altitude value. Even if the data represented in the model are continuous, usually to achieve a DEM from direct measurement of land or existing maps discrete data are used in the first stage.

Ideally, data sources can be used without interpolation. For example, contour lines represent the altitude of land surface and may be taken directly from topographic maps, but to represent a terrain model continuous data is required.

Interpolation is the transformation of these discrete data in a raster containing continuous data using a function that passes through (or near) the input data points. Because there are an infinite number of functions that can meet these requirements, additional conditions must be imposed, which results in a number of different interpolation methods (Neteler and Mitasova 2004). Interpolation methods can be based on zonal conditions, where the interpolation is done locally on a subset of sampling points (eg Voronoi polygons, Inverse Distance Weighted, Natural Neighbor, Nearest Neighbor). Other methods are based on uniformity, where the result will have as few discontinuites as possible (eg Spline).

To obtain a DEM interpolated data sources that can be multiple, but the most common sources are inexpensively available topographic maps. In Romania there are several series of topographic maps at various scales prepared by the Military Topographic Directorate. One of the latest series of such maps is the topographic map, scale 1:25.000, edition 1982. Topographic maps in this series contain data as elevation contours with equidistance 10m and points with altitude in important locations (eg. peaks).

Altitude database construction for this model

The model presented in this work needs a high resolution altitude database for a determination of the catchment characteristics suitable for the application. Due to the processing capabilities of computers and the small area of the catchments the dimension of the final elevation dataset is not a big problem. If a good resolution can be obtained using an available dataset, the results of applying the model will be better.

However, there are some problems regarding the resolution of the DEM that will be used.

In the case of data interpolation from topographic maps, the first problem is the spatial accuracy of data. Even if the cartographic rule presented in the previous section suggests that a resolution of 12.5 m can be obtained from a 1:25.000 topographic map, it is not a good idea to use that value for interpolation of data extracted from the maps. There are several sources of error in this method and each of these sources will impose lowering the resolution to obtain more spatially accurate results or accepting the error as a compromise.

These sources of error include:

- Accuracy of the original map – there is no way of assessing the accuracy in measurements from the original topographic map. If the advance of the technologies from 1982 until now is taken into account, the digital measuring processes and error correction along with the GPS technologies offer significantly better results in terms of accuracy.

- Errors introduced through georeferencing – the georeferencing presumes taking an image without any spatial significance and positioning the image at the geographic coordinates of the information in the image. This process gives a spatial dimension to a simple raster map needed for usage in the GIS. Different

coordinate systems (methods of representing geographic location) can be used for the map coordinates according to the area and specifics of the application or the other datasets involved in the analysis. As georeferencing is a manual process, the user can introduce additional error in the accuracy.

- Errors in digitizing of the information – the contours on a map are continuous lines obtained through an interpolation of certain measurements when the initial topographic map was created. The GIS uses a polyline or point representation of the information extracted from these maps rather than a continuous line. The dimension of the segments in this polyline and the accuracy of placing the points on the scanned map is another important source of error. If the first two sources could be determined through certain mathematical calculations, this source of error is highly dependent on the quality of the digitizing done by the user.

- Errors in coordinate system transformations – different data obtained from different sources may have different coordinate systems. For example Romania uses a local coordinate system named Stereo 70 and all the data obtained from institutions in other countries will use coordinate sistems specific to the whole world or to the countries where the data is created. The transformations between these coordinate systems can introduce other errors in the result.

Due to these sources of error the accuracy of an elevation model obtained from topographic maps can be lower than the presumed accuracy. An extensive discussion on digital terrain model accuracy based on experiments is presented by Martinoni (2002).

The ASTER GDEM dataset presented in the previous section uses remote sensing to obtain terrain altitude values. As specified, the resolution of the ASTER GDEM is about 25 m at the latitude of Romania. Spatial accuracy of the ASTER dataset is affected by the method of obtaining the DEM from stereographic images. The sources of error in this case include:

- Errors due to characteristics of the terrain – the GDEM is obtained through automatic processing of remotely sensed data extracted from stereographic images. This is the reason why any kind of obstacle on land will be recorded as a difference in terrain altitude. Forests, for example, will be recorded at the treetop level rather than the ground level

- The errors in coordinate system transformation are the same as the ones presented in the previous case

The ASTER GDEM website (2011) states an accuracy of 7-14 m for the ASTER GDEM but Reuter et al. (2009) made an assessment of the ASTER GDEM dataset and concluded that "the Aster GDEM contains a significant number of anomalies that prevents immediate use for a wide range of applications".

To obtain a DEM from topographic map information the following steps were taken:

- Scanning and georeferencing of the topographic map: results in a raster without any significance for modeling terrain; A GIS product does not automatically obtain elevation data from a topographic map, so data on this map should be extracted by the user, usually through a manual process.

- Digitizing the scanned map: results in a spatial database with different elements: contour lines, points with altitude attributes, hydrographic network boundaries. The result is still not a DEM because the elevation values are only on the contours and the points with altitude and a DEM presumes continuous elevation throughout the area under work;

- Interpolation of sampling points: after this operation the DEM is obtained in raster format, as a form of continuous spatial representation of topography. In this case, the ANUDEM (TOPOGRID) method implemented in the ArcINFO GIS (Hutchinson 1996) was used for interpolation. Because ANUDEM is an interpolation method specifically designed to obtain a DEM from elevation data in the form of contours or points with altitude and offers many special features to create a correct DEM (eg. drainage constraint), this algorithm is very suitable to achieve a proper digital elevation model through interpolation.

A comparison between TOPOGRID interpolated DEMs and the ASTER dataset was made as follows and gave the following results (the figures are extracts of a map from the southern area of the Apuseni Mountains):

The TOPOGRID algorithm suggested a default value of 38m for the interpolation so that

127

value was also used for testing besides the 25m value that I considered suitable for the purpose. An example of a topographic map sheet interpolated with this algorithm can be seen in Fig. 3.7 However, if areas larger than a topographic map sheet need to be interpolated, several adjustments must be made. The contours must be joined with the contours of the same value from the neighboring sheet to retain correct topology in the map.
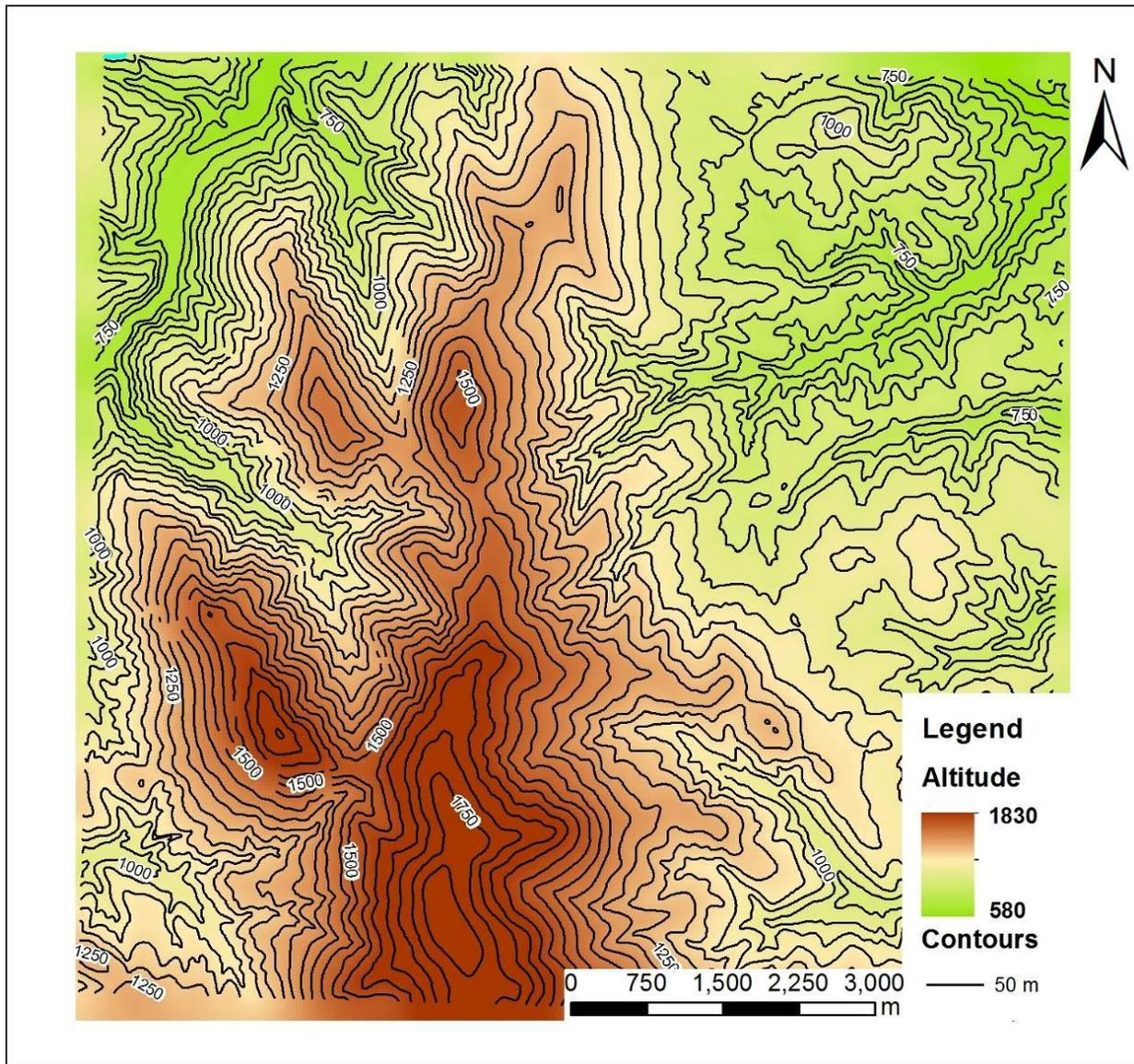
Fig. 3.6 *Altitude map obtained through interpolation of contours*

The comparison between the resolution of the TOPOGRID interpolated rasters and the ASTER dataset can be seen in Fig. 3.6

The advantage of the ASTER DEM over the contour interpolation at 38 m resolution is

clear in terms of resolution. If the accuracy specified on the ASTER GDEM website is correct, the ASTER GDEM can easily replace a DEM interpolated from contours at this resolution.

The comparison between the ASTER GDEM and the 25 m resolution DEM interpolated from contours cannot be made because the TOPOGRID algorithm created the DEM in a geodatabase format and ArcGIS automatically smooths the raster cells through a method of interpolation for better viewing. However, the resolution of the 25 m DEM is very close to the resolution of the ASTER DEM.

The altitude comparison was made by substracting the values in a raster from the values in another raster and viewing the results. Fig 3.8 shows an area of the DEM with the major differences in values. There are small areas (a few cells) where the differences exceed 100m, like the one presented in the figure, but they cover a small extent and most of the raster values are similar between the three rasters.

The altitude differences can appear from erroneus measurements in the ASTER but also from corrections in the contour interpolation like sink filling and drainage enforcement imposed by the TOPOGRID algorithm. Statistical calculations on the differences between the DEMs were not performed.

The results of the comparisons show that the ASTER DEM can replace the DEM interpolated from contours with little accuracy loss. Other reasons for using the ASTER DEM in modeling are:

- The data is available for free and the time needed for data processing can be used for other analysis tasks
- The digitizing of the contour data needs a lot of work and a suitable result can be obtained only with enough attention, good instruments and good skills in using the instruments.
- The ASTER DEM covers any area of the Earth so topographic maps do not need to be obtained for every area where the model is applied
- The processing of an ASTER DEM scene will be the same for all the scenes, so an automation of this preocessing can be carried out in the model.
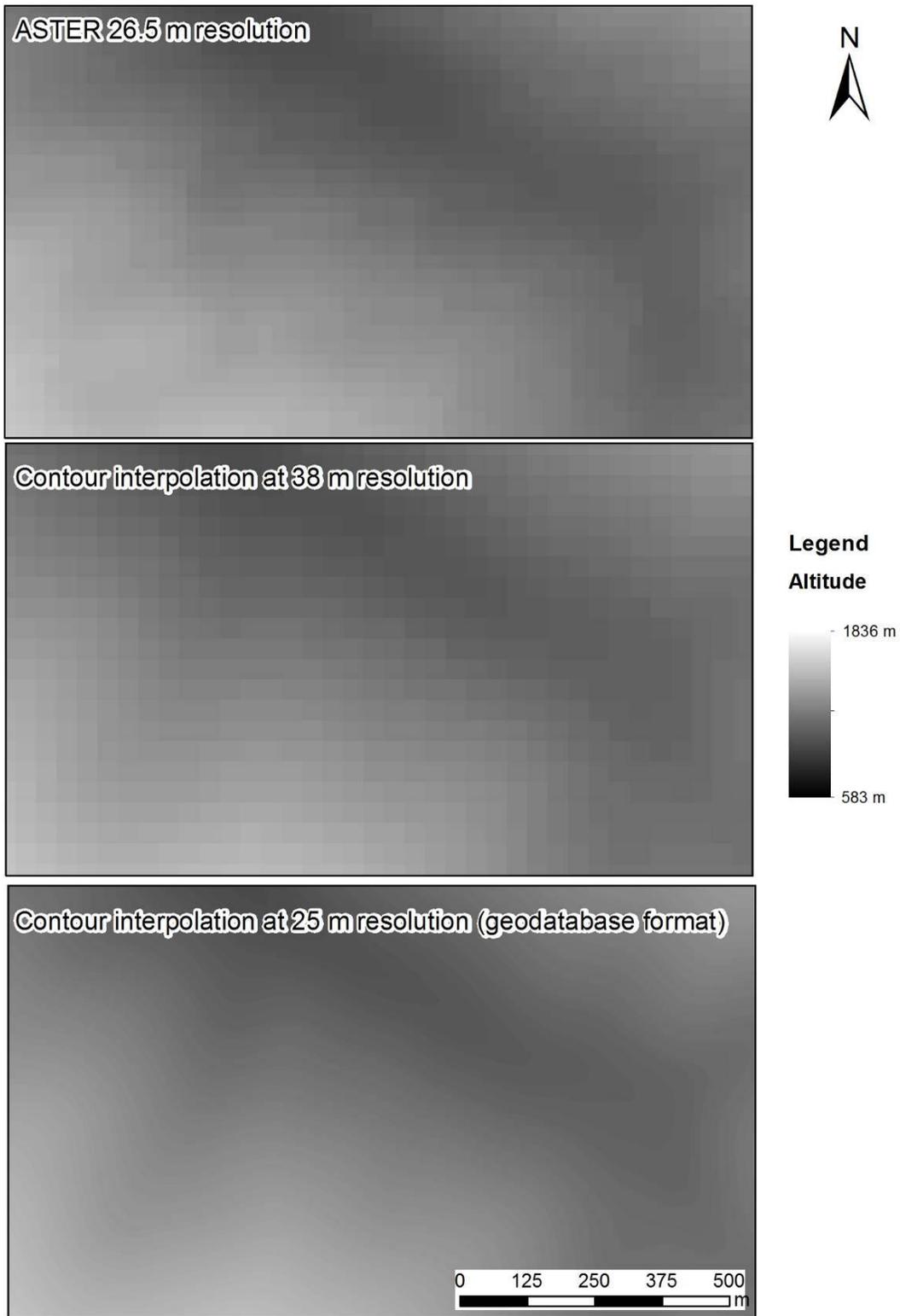
Fig 3.7 *Comparison between ASTER GDEM resolution and Contour interpolation*
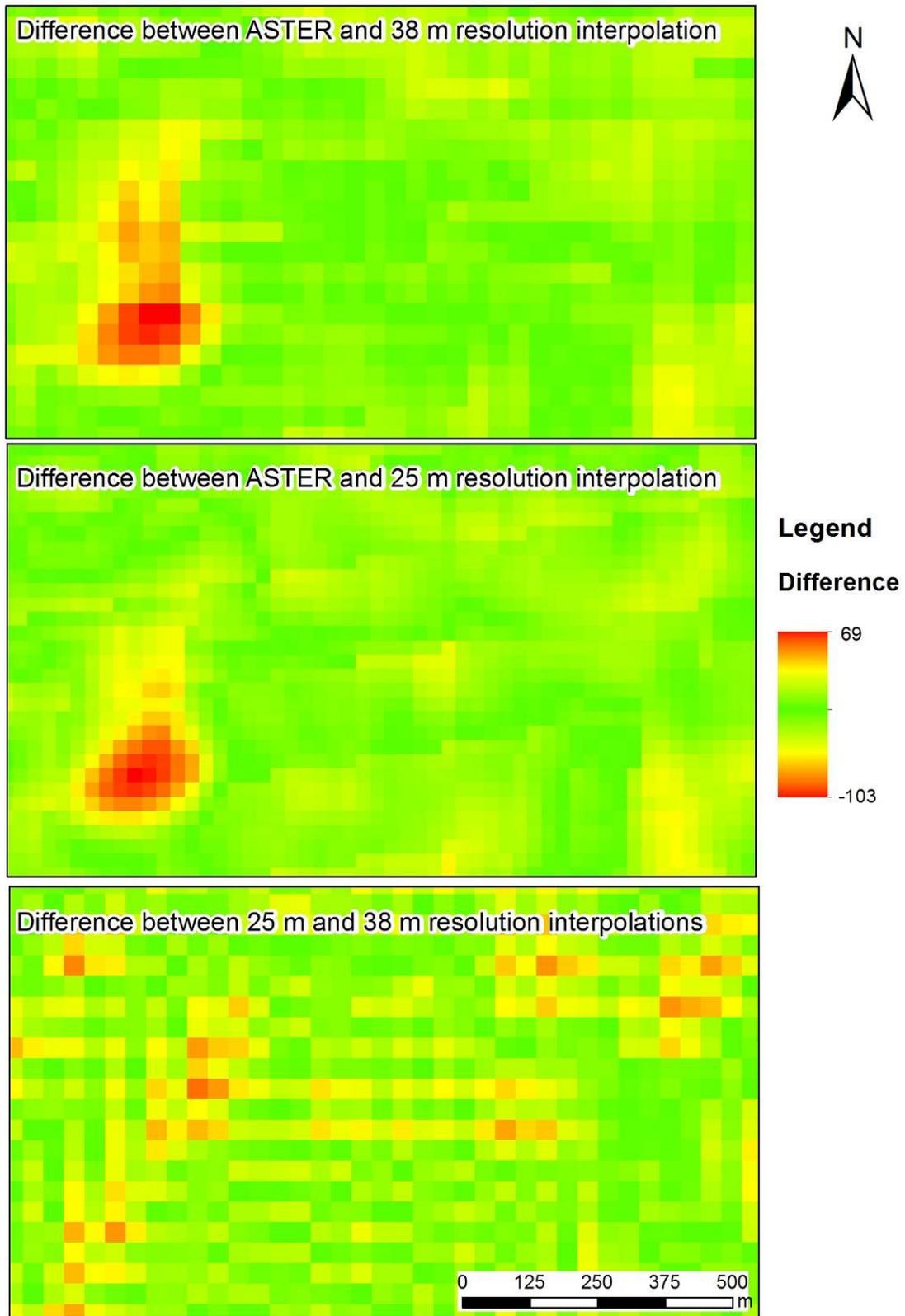
Fig 3.8 *Difference in the results from the ASTER DEM and the interpolated DEMs*

3.3 Soil data

Soil data availability in Romania

Soil data is available as soil maps. These maps represent the soil type for the surface of the ground and some characteristics of this soil type. The soil maps distributed in Romania by the ICPA (Institutul de Cercetari pentru Pedologie si Agrochimie – The Institute for Pedology and Agrochemical Research) are the base maps for Romania and they can be acquired at different scales (1:10.000, 1:50.000, 1:200.000, 1:500.000, 1:1.000.000), Thematic maps (erosion, humidity excess etc.) can also be acquired, but all these products are only commercially available. The maps come in a paper format and no GIS database for soils is currently available (as of 2010).

These maps can be converted into digital format by scanning, digitizing and correcting the topology of the result, then the study surface can be used along with the DEM by converting the digitized maps (which are originally digitized as polygons) to a raster format. Each cell will have a numeric value corresponding to a certain soil type or hydrologic soil group and these values will be used in the analysis later as needed.

Each soil type and land cover has a specific infiltration rate, but also other characteristics from these layers can be used in rainfall-runoff modeling. According to the type of soil and land cover, the runoff can have a certain speed which is needed when a user needs to determine the evolution of the discharge through time. The quantity of water in the soil that can also be obtained using the soil type and the rainfall data is also important in runoff modeling.

Soil texture refers to the size of the particles that make up the soil. Due to these characteristics, the texture has influence on the permeability, the infiltration and water retention capacity of the soil.

Three elements are important in determining soil texture: sand, which is associated with high permeability, clay, with high water retention capacity and silt (mud) with a contribution on plasticity and cohesion of the soil and on water retention capacity. The terms *sand*, *silt*, and *clay* refer to relative sizes of the soil particles. The dimensions of the particles determining the soil texture (Fig. 3.9) differ as follows (Table 3.4):

Name of the particles	Size limits in mm	Distinguishable with naked eye
gravel	larger than 1	obviously
sand	1 to 0.02	easily
silt	0.02 to 0.002	barely
clay	less than 0.002	impossible

Table 3.4 – Soil particle size

The amount of sand, silt and clay present in the soil (the dimension of soil particles) determines the soil texture.

The relative dimensions of soil particles can be seen in Fig. 3.9
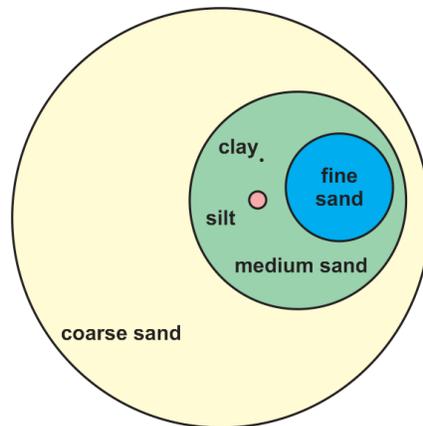
Fig. 3.9 Relative soil particle size

The soil classification system in Romania is based on the proportions of the three elements presented above (sand, clay, loam) according to the ICPA classification system. The soil texture can also be obtained using the triangular diagram (fig 3.10), which is an equilateral triangle and each side corresponds to one of the three elements defining texture or a rectangular triangle (when just the silt and loam are considered and are placed on the perpendicular edges).

Due to the specifications of the SCS method, this study used the USDA soil texture classification system was used in this study. The most popular soil texture classification diagram at an international level is probably the USDA classification system (Florea et al 2000; Lupaşcu, Patriche 2000). This soil classification system was presented along with the SCS methodology in section 1.6.

Fig. 3.10 presents a comparation between the triangular diagram of the USDA soil classification system and the ICPA soil classification system. Based on this comparation, the USDA soil groups can be correlated to soil groups from the Romanian soil maps.

(a)   (b)
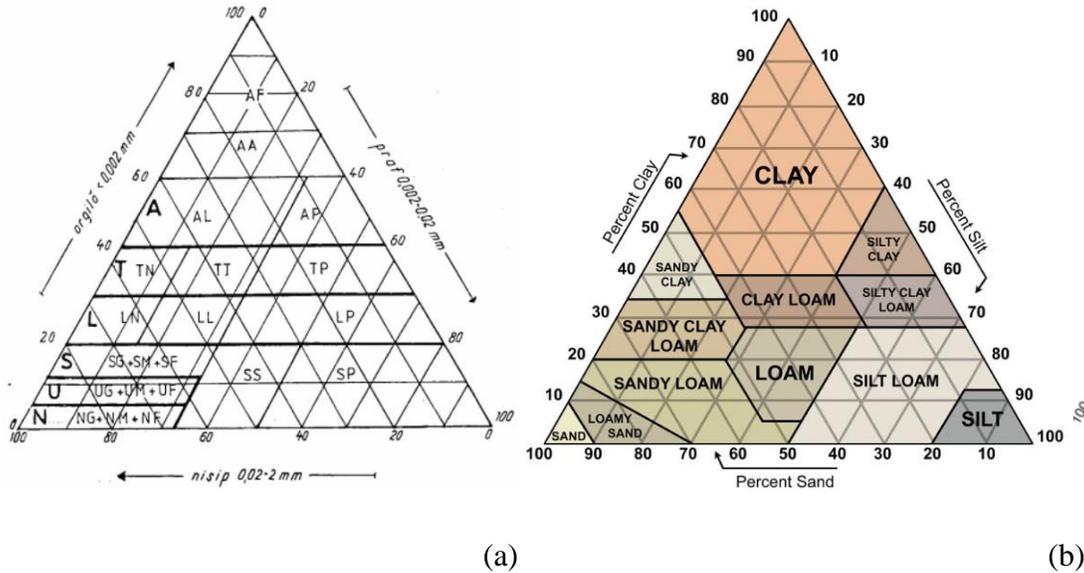
Fig. 3.10 *The Romanian (a) and USDA (b) soil classification systems diagrams*

The adaptation of the Romanian soil texture classification system and determination of HSG corresponding to each texture was done according to the method proposed by Chendeș (Chendeș 2007), using the texture classes from the ICPA.

Soil Database construction for the current model

The soil maps used for this study were digitized from the 1:200.000 soil maps. These maps represent the most important achievement of soil survey activity in Romania. The surveys were made between 1963 and 1994, but some field work had started in the 1950s. The coordination was initially assumed by Cernescu and from 1967 onward by Florea. The field work, compilation and printing were carried out at the Geological Institute and after 1970 at the Research Institute for Soil Science and Agrochemistry. (Florea 1994)

The soil map of Romania at a scale 1:200.000 comprises 50 sheets in color (Fig. 3.11) and has a general legend with 470 mapping units. Although the Romanian Soil Classification underwent several changes during the production of this map, the general legend incorporates correlation between different periods.

134

DISTRIBUTION OF SHEETS AT 1:200,000

LEGEND

20 - SHEET NUMBER

1990 - YEAR OF PUBLICATION

RISSA - BUCHAREST
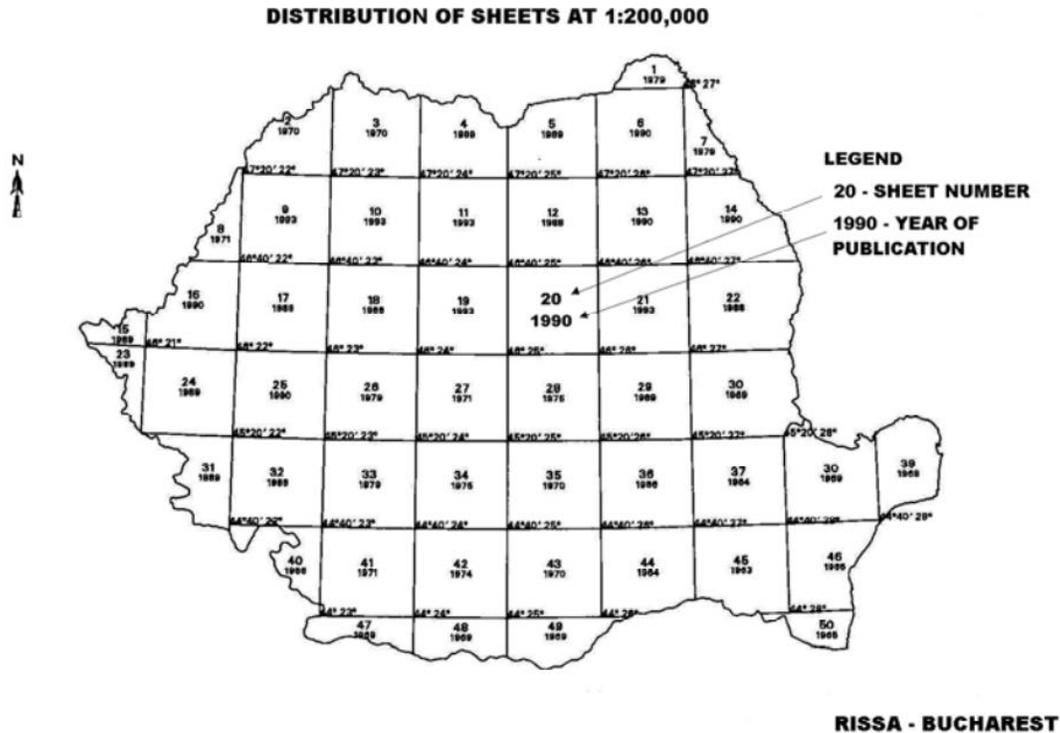
Fig. 3.11 *Soil Map Sheets for Romania (Munteanu et al. 2005)*

The area of study was digitized using GIS and a polygon layer for soils was obtained. According to the cartographic resolution rule presented in section 3.2 the accuracy that can be obtained from a map is the equivalent of 0.5 mm at map scale. In the case of 1:200000 maps, the accuracy would be of about 100m in the ideal case (scanning and georeferencing of the maps without errors and very accurate digitization). If the user takes into account the errors in georeferencing and digitization the accuracy decreases to over this value.

The soil texture extracted from this map was adapted according to the characteristics specified by NRCS for the Hydrologic Soil Groups and the corresponding HSG was assigned to each soil polygon digitized.

Figure 3.12 presents the soil types for the Someșul Cald basin upstream from Smida along with the attributes of the soils.

The attributes were collected from the information on the scanned soil maps and introduced in the attribute table of the layer. The following attributes were collected: soil symbol; soil name; soil texture and soil weight.

The USDA soil group was determined by comparation between the two soil classification systems presented earlier.

135

Soil Attributes
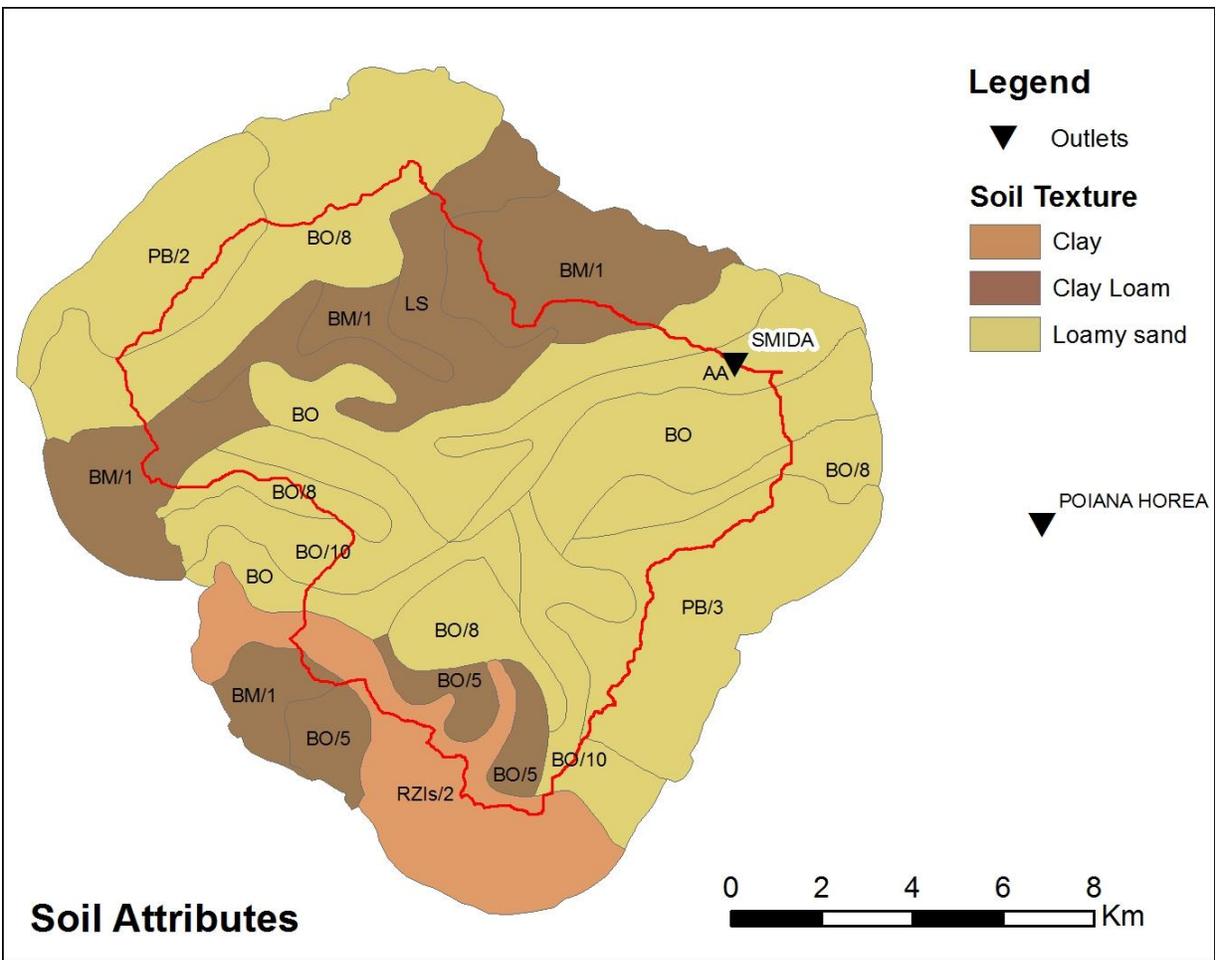
Symbol	Name (Rom)	Texture	Soil type	HSG
PB/3	Prepodzol si podzol	Sandy Clay	medium	B
BO/10	Districambosol si litosol	Sandy Clay	medium	B
RZls/2	Rendzina si roci compacte la zi	Clay	fine	D
BO/5	Districambosol si eutricambosol	Clay Loam	fine	C
BO/10	Districambosol si litosol	Sandy Clay	medium	B
BO/5	Districambosol si eutricambosol	Clay Loam	fine	C
BO/5	Districambosol si eutricambosol	Clay Loam	fine	C
BO/8	Districambosol si prepodzol	Sandy Clay	medium	B
BO/8	Districambosol si prepodzol	Sandy Clay	medium	B
BO	Districambosol	Sandy Clay	mijlocie	B
BM/1	Eutricambosol si rendzina	Clay Loam	fine	C
AA	Aluviosol	Sandy Clay	medium	B
LS	Litosol si stancarie	Clay Loam	fine	C
BM/1	Eutricambosol si rendzina	Clay Loam	fine	C
BO	Districambosol	Sandy Clay	medium	B
BM/1	Eutricambosol si rendzina	Clay Loam	fine	C
BO/8	Districambosol si prepodzol	Sandy Clay	medium	B
PB/2	Prepodzol si podzol	Sandy Clay	medium	B
BO/8	Districambosol si prepodzol	Sandy Clay	medium	B
BO	Districambosol	Sandy Clay	medium	B

Fig. 3.12 *Soil characteristics of the Someşul Cald basin*

3.4 Land use data

Land use databases for Europe

Land use is another important factor determining the way the surface runoff appears. Land use maps can be obtained from different data sources available commercially (cadastre maps, urban plans of different types). These maps can also be in vector (polygons representing areas with a certain soil type) or raster (each cell value corresponds to a certain land use) format. Like in the soil maps, the vector maps have to respect some topology rules: polygons don't overlap (because a certain point cannot have two different land uses at the same time) and there are no empty spaces (because any point on the land has a certain land use).

Like in the case of altitude, possibilities of obtaining land use data from the internet exist. Such a dataset is available for Europe under the Corine Land Cover project.

The Corine Land Cover program is the most important activity of mapping the land use and land cover based on satellite images and regional land use databases. CORINE Land Cover is a compilation of national land cover inventories, which are integrated into a seamless land cover map of Europe. The resulting European data-base is based on a standard methodology and nomenclature.

The Corine Land Cover project was adopted by the European Commission in 1985 (Directorate General "Environment") then managed by the European Topic Centre of the European Environment Agency in 1993. The aim of Corine Land Cover is to provide information on land cover and on the state of the environment in the European Union. Corine Land Cover is a cartographic tool which covers every national territory where the survey is undertaken. CORINE Land Cover databases are obtained through computer assisted interpretation of satellite images acquired in 1990, 2000 and 2006, offering the possibility to describe the geographic distribution of specific land cover changes in a georeferenced approach (Bossard et al. 2000).

The European Environment Agency (EEA) has country specific information e.g., on built-up and related areas, open land, agricultural land, forests, wetlands, and water bodies in the Corine land cover database. The classification is very detailed. Built-up and related areas include for

example continuous urban fabric, discontinuous urban fabric, industrial units, port areas and airports. Open land includes bare rocks, sparsely vegetated areas, beaches, dunes, etc. Data is available for 24 European countries, most of whom are members of the EU (countries outside the EU include Bulgaria, Croatia and Romania). (Matleena Kniivila 2004)

CORINE land cover (CLC) describes land cover (and partly land use) with a three-level nomenclature of 44 classes. CLC was elaborated based on the visual interpretation of satellite images (Spot, Landsat TM and MSS). Ancillary data (aerial photographs, topographic or vegetation maps, statistics, local knowledge) is used to refine interpretation and assign classes. The main categories at the first hierarchical level are: artificial surfaces, agricultural areas, forest and semi natural areas and water bodies. (Feuerbacher et al. 2005)

The CLC database is based on a standard production methodology characterised by the following elements: Mapping scale is 1:100 000. Mapping accuracy is 100 m. The minimum mapping unit for the inventory is 25 ha for areas, and 100 m for linear elements.

The CORINE Land Cover database was used for the land cover data in this study. The land use classes were used in calculation of their effect on flow patterns and automatic extraction of characteristics influencing flow (e.g. interception, surface roughness, water storage capacity).

Land use database sources and processing for this model

The land use database for each catchment was created using data from the CORINE Land Cover 2006 database.

To obtain the land use database the following steps were necessary:
- Download data for the area from the EEA CORINE Land Cover website – the CLC2006 dataset is available for free on the EEA website and datasets for each country can be downloaded in different formats. The model created in this study uses the shapefile vector format for the landuse dataset.
- Project data to the same coordinate system with the rest of the data – the CLC2006

database is originally available in the ETRS89 / ETRS-LAEA coordinate system, a true area coordinate system for Europe. The rest of the datasets are created in the Stereo70 coordinate system, the main coordinate system used in Romania. The GDAL tools were used for reprojecting the data from one coordinate system to another.

- Clip the extent of the basin from the database – once the land use dataset is projected to the chosen coordinate system, a Clip operation is applied and the dataset is clipped by the extent of the study area.

The land use for the Someşul Cald basin along with the attribute table can be seen in Fig. 3.13. The land use for all the other basins was presented in section 1.6

The classification on the three levels can be seen in the attribute table and the map colors correspond to the third level of classification.

Land Use
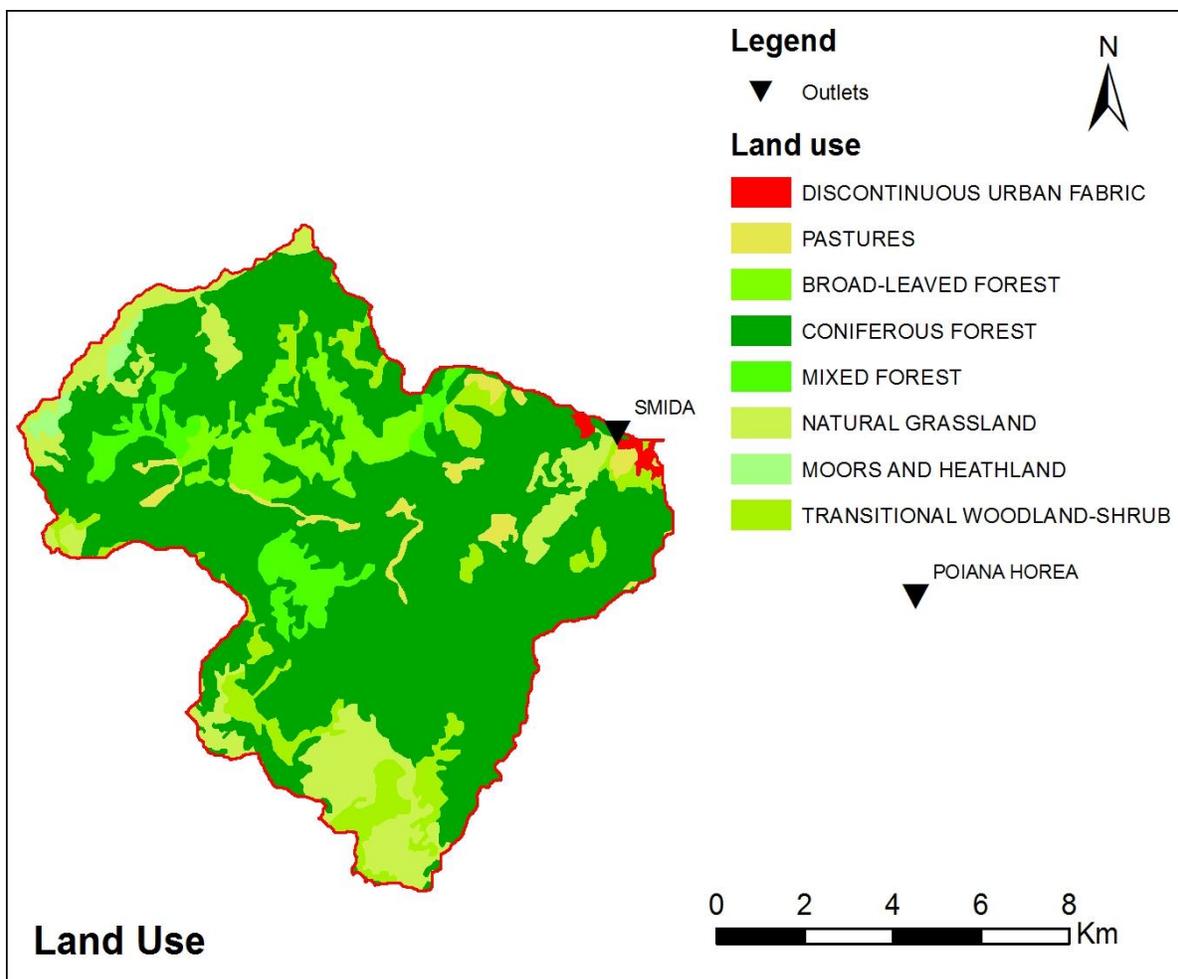
CODE_06	LABEL1_EN	LABEL2_EN	LABEL3_EN
112	Artificial surfaces	Urban fabric	Discontinuous urban fabric
112	Artificial surfaces	Urban fabric	Discontinuous urban fabric
231	Agricultural areas	Pastures	Pastures
231	Agricultural areas	Pastures	Pastures
231	Agricultural areas	Pastures	Pastures
311	Forest and semi natural areas	Forests	Broad-leaved forest
311	Forest and semi natural areas	Forests	Broad-leaved forest
312	Forest and semi natural areas	Forests	Coniferous forest
312	Forest and semi natural areas	Forests	Coniferous forest
313	Forest and semi natural areas	Forests	Mixed forest
313	Forest and semi natural areas	Forests	Mixed forest
321	Forest and semi natural areas	Scrub and/or herbaceous vegetation	Natural grasslands
321	Forest and semi natural areas	Scrub and/or herbaceous vegetation	Natural grasslands
322	Forest and semi natural areas	Scrub and/or herbaceous vegetation	Moors and heathland
324	Forest and semi natural areas	Scrub and/or herbaceous vegetation	Transitional woodland-shrub
324	Forest and semi natural areas	Scrub and/or herbaceous vegetation	Transitional woodland-shrub

Fig. 3.13 *CORINE Land Cover data for the Someşul Cald basin*

3.5 Rainfall data

If the data presented until now have an important role in the evolution of surface runoff, the runoff is generated by the rainfall falling on the ground.

The rainfall can be considered uniform or spatially distributed according to the necessities, the data availability and the size of the area under study.

Rainfall data can be obtained from different sources: rainfall maps are available at meteorological stations or can be created by interpolation of data from different measurement points on the surface under study.

3.5.1 Climate databases available for Romania

Some rainfall data from different rain gauges are available for free on the internet from different providers. The most popular provider for rainfall and other types of climatic data is the National Climatic Data Center in the USA (NCDC). NCDC is an organization located in Asheville, North Carolina that has the biggest climate data archive in the world. NCDC offers a large number of global datasets, including daily and hourly rainfall maps (Federal Climate Complex 2010).

Another free dataset with weather information is the European Climate Assessment & Dataset project (ECA&D). Climate datasets are usually archived at the national meteorological institutes. The need for exchanging high resolution observational data formed the motivation for the EUMETNET-European Climate Support Network (ECSN) to start the European Climate Assessment & Dataset project (ECA&D) in 1998. In this project the participating countries collaborate successfully to form an international observational dataset with daily data for a large area, including Europe, northern Africa and the Middle East. The database contains daily station data for over 3000 stations, and this number is continuously increasing. (Besselaa, 2011). Even if the data from the ECA&D project is blended and daily values are available, the number of measuring stations is quite small compared to the NCDC database.

Another source of rainfall data are the RADAR measurements. The RADAR (RAdio Detection and Ranging) represents a fixed installation using electromagnetic waves and their reflection from different objects to determine their position relative to the antenna. The meteo radar can be used for determining the location, movement and type of precipitation and for estimating

its future changes in position and intensity. Modern Doppler radars can not only detect the intensity of rainfall but also the movement of rainfall systems. The information coming from the radar are analyzed to identify the structure of storms and bad weather potential (National Meteorological Administration 2011).

Weather stations possess their own climate variable database and these values are stored into tables maintained at the National Meteorological Administration. According to the NMA website, repeated access to historical data from the National fund is limited. However, access for individuals and companies outside the National Meteorological Administration, other than strategic users, historical data for a place of observation is available for a fee. (NMA 2011). Therefore, the weather data from the National Meteorological Administration cannot be used for scientific purposes without a fee. If this data is available it should be used instead of data from the free databases presented.

As RADAR measurements are not available for free, this study used NCDC GSOD data as the main free rainfall data source when needed. Global Surface Summary of the Day is a product archived at the National Climatic Data Center (NCDC). It is produced by the National Climatic Data Center (NCDC) in Asheville, NC and is derived from The Integrated Surface Hourly dataset. The ISH dataset includes global data obtained from the USAF Climatology Center, located in the Federal Climate Complex with NCDC. The latest daily summary data are normally available 1-2 days after the date-time of the observations used in the daily summaries.

The Global Surface Summary of day Data is available for free and contains the following elements (University of Miami 2011): Mean temperature (.1 Fahrenheit), Mean dew point (.1 Fahrenheit), Mean sea level pressure (.1 mb), Mean station pressure (.1 mb), Mean visibility (.1 miles), Mean wind speed (.1 knots), Maximum sustained wind speed (.1 knots), Maximum wind gust (.1 knots), Maximum temperature (.1 Fahrenheit), Minimum temperature (.1 Fahrenheit), Precipitation amount (.01 inches), Snow depth (.1 inches), Indicator for occurrence of: Fog, Rain or Drizzle, Snow or Ice Pellets, Hail, Thunder, Tornado/Funnel Cloud.

The NCDC database for Romania contains daily data for more than 200 stations (Fig. 3.14) spread through the country.
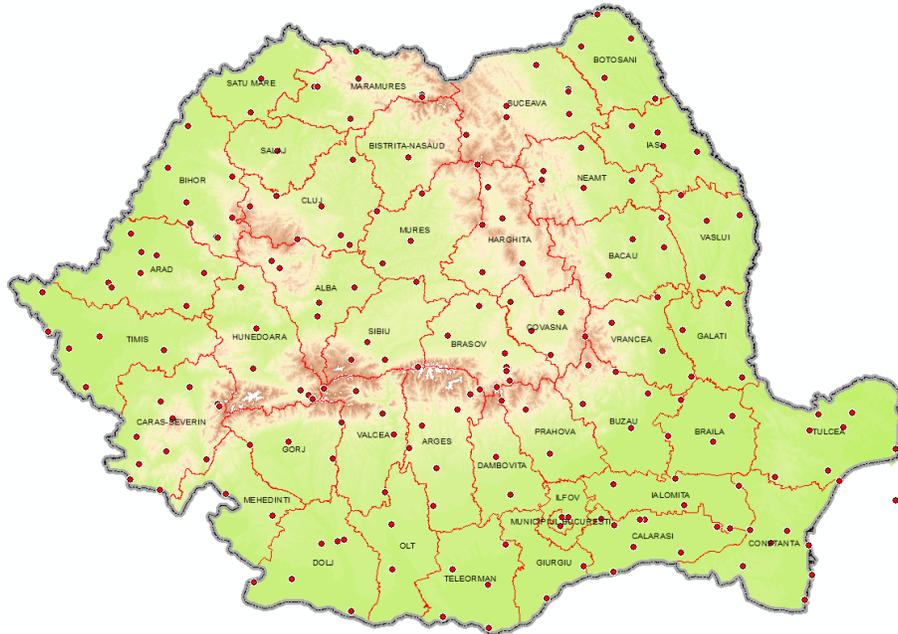
Fig. 3.14 *NCDC GSOD stations in Romania*

The data is available as tables with records for each day. A table specific for every weather station exists and the tables contain daily data since 1928 and are publicly available from ftp://ftp.ncdc.noaa.gov/pub/data/gsod/. Every element is recorded in the table in the corresponding field and the field structure can be found in the NCDC documentation. (NCDC 2010). It is important to note that the data is not continuous, some stations lack data for certain days or periods and some stations only have some of the data available in different periods.

Even if all the databases presented in the previous section should contain the rainfall data needed for the model application, the availability of rainfall data is not continuous and consistent. This is why different procedures were used to extract rainfall data from the rainfall data sources presented.

The Python scripts presented in the next section extract the data from the NCDC database and can either be called from the command line with the necessary parameters or from ArcGIS using a form to insert the parameters.

The scripts can extract any parameters from the NCDC dataset as presented in the

previous section and can sum them or calculate the mean for a given period or for the whole year.

The NCDC database from Romania is not very complete and some stations lack data for long periods. To give a perspective on the situation, two extracts from the NCDC database tables are given here for the Cluj-Napoca station, one of the biggest meteorological stations in Romania. The first one is for the period 1.04. 2000 - 15.04. 2000 (Table. 3.5) and the second one is for the period 12.08.2007 – 26.08.2007(Table. 3.6)

Table 3.5 NCDC database extract for Cluj-Napoca, 1-15.04.2000

STN---	YEARMODA	TEMP	DEWP	VISIB	WDSP	MXSPD	PRCP	SNDP
151200	20000401	49.8/18	33.1/18	6.4/18	3.7/18	7.8	99.99	999.9
151200	20000402	49.9/18	41.5/18	7.5/18	3.2/18	11.7	99.99	999.9
151200	20000403	52.1/19	37.6/19	5.9/19	3.4/17	7.8	0.01E	999.9
151200	20000404	54.1/21	44.1/21	8.2/21	2.9/20	3.9	0.06F	999.9
151200	20000405	54.9/20	50.9/20	6.9/20	3.6/19	11.7	99.99	999.9
151200	20000406	41.0/21	38.0/21	6.6/21	16.2/21	23.3	99.99	999.9
151200	20000407	38.0/20	20.3/20	8.0/20	14.1/20	27.2	0.03A	999.9
151200	20000408	40.9/20	27.4/20	7.7/20	8.5/20	17.5	0.03E	999.9
151200	20000409	38.0/20	22.9/20	7.8/20	6.5/19	11.7	0.00I	999.9
151200	20000410	42.5/21	25.2/20	6.4/21	4.1/21	11.7	0.00I	999.9
151200	20000411	46.4/20	32.2/20	6.2/20	2.9/20	5.8	0.00I	999.9
151200	20000412	52.5/19	35.9/19	6.0/19	4.3/18	9.7	0.00I	999.9
151200	20000413	55.5/20	42.0/20	6.5/20	2.0/19	5.8	0.00I	999.9
151200	20000414	59.2/18	45.2/18	8.4/18	2.4/18	5.8	0.03E	999.9
151200	20000415	60.5/20	43.5/20	7.3/20	3.3/17	9.7	0.00I	999.9

Table 3.6 NCDC database extract for Cluj-Napoca, 12-26.08.2007

STN---	YEARMODA	TEMP	DEWP	VISIB	WDSP	MXSPD	PRCP	SNDP
151200	20070812	63.1/24	58.1/24	6.2/24	4.1/24	11.7	1.41F	999.9
151200	20070813	65.2/24	57.3/24	6.7/24	3.5/24	5.8	0.02F	999.9
151200	20070814	66.5/24	59.8/24	6.4/24	6.7/24	13.6	0.03F	999.9
151200	20070815	69.0/24	56.6/24	6.2/24	5.4/24	7.8	0.00F	999.9
151200	20070816	69.4/24	57.9/24	5.8/24	3.6/24	7.8	0.00F	999.9
151200	20070817	71.5/24	58.8/24	5.3/24	3.1/24	7.8	0.00F	999.9
151200	20070818	69.8/24	62.0/24	5.2/24	4.9/24	13.6	0.00F	999.9
151200	20070819	68.7/24	62.3/24	5.6/24	4.0/24	13.6	0.02F	999.9
151200	20070820	72.5/24	61.8/24	5.9/24	4.5/24	9.7	0.00F	999.9
151200	20070821	72.6/24	62.1/24	6.7/24	5.7/24	15.5	0.04F	999.9
151200	20070822	75.0/24	63.2/24	4.7/24	4.9/24	13.6	0.00F	999.9
151200	20070823	78.6/24	64.1/24	5.5/24	4.6/24	9.7	0.00F	999.9
151200	20070824	79.6/24	65.3/24	5.9/24	2.1/24	5.8	0.00F	999.9
151200	20070825	73.3/24	65.8/24	5.4/24	4.6/24	11.7	2.24F	999.9
151200	20070826	66.7/24	63.4/24	5.0/24	3.6/24	9.7	0.20F	999.9

Where:

STN--- Meteorological station code

YEARMODA - Year Month Day

TEMP - Temperature (°F) / Number of observations used in calculation

DEWP - Dew Point (°F) / Number of observations used in calculation

VISIB - Visibility (miles) / Number of observations used in calculation

WDSP - Mean wind Speed (knots) / Number of observations used in calculation

MXSPD - Maximum sustained wind speed (knots)

PRCP - Total precipitation (rain and/or melted snow) in inches

- Precipitation flag (char) has the following meaning:

- o A = 1 report of 6-hour precipitation amount.

- o B = Summation of 2 reports of 6-hour precipitation amount.

- o C = Summation of 3 reports of 6-hour precipitation amount.

- o D = Summation of 4 reports of 6-hour precipitation amount.

- o E = 1 report of 12-hour precipitation amount.

- o F = Summation of 2 reports of 12-hour precipitation amount.

- o G = 1 report of 24-hour precipitation amount.

- o H = Station reported '0' as the amount for the day, but should be considered as incomplete data for the day.

- o I = Station did not report any precipitation data for the day --it's still possible that precipitation occurred but was not reported.

- o SNDP - Snow depth in inches

As 99.99 signifies a missing value, the extract from year 2000 only contains valid precipitation amounts for only 5 from the 15 days shown. Over the whole year only 120 values were valid measurements at this station and none of the reported values was over 1 inch per day. The less important meteorological stations have even less data available so this method should not be used for the older data.

In the case of the extract from 2007, the precipitation measurements could be valid for all

of the 15 days presented here. Over the whole year, 363 values were valid measurements. Therefore, the data from the more recent years can be used to extract precipitation amounts. As a comparison, the ECA&D table of values for the same period of time can be used (Table 3.7)

Table 3.7 ECA&D Dataset for Cluj-Napoca compared to the NCDC GSOD dataset, 12-26.08.2007

STAID	SOUID	DATE	RR	Q_RR	PRCP	PRCP(.1mm)	DIFF
902	100902	20070812	360	0	1.41F	358.14	1.86
902	100902	20070813	4	0	0.02F	5.08	-1.08
902	100902	20070814	8	0	0.03F	7.62	0.38
902	100902	20070815	0	0	0.00F	0	0
902	100902	20070816	0	0	0.00F	0	0
902	100902	20070817	0	0	0.00F	0	0
902	100902	20070818	0	0	0.00F	0	0
902	100902	20070819	4	0	0.02F	5.08	-1.08
902	100902	20070820	0	0	0.00F	0	0
902	100902	20070821	14	0	0.04F	10.16	3.84
902	100902	20070822	0	0	0.00F	0	0
902	100902	20070823	0	0	0.00F	0	0
902	100902	20070824	0	0	0.00F	0	0
902	100902	20070825	572	0	2.24F	568.96	3.04
902	100902	20070826	56	0	0.20F	50.8	5.2

Where:

- STAID: Station identifier

- SOUID: Source identifier

- DATE : Date YYYYMMDD

- RR : Total precipitation from ECA&D (rain and/or melted snow) in 0.1 mm

- Q_RR : Quality code for RR (0='valid'; 1='suspect'; 9='missing')

- PRCP : Total precipitation from NCDC (rain and/or melted snow) in inches

- PRCP (.1 mm): Total precipitation from NCDC (rain and/or melted snow) in 0.1 mm

- DIFF – Difference between ECA&D and NCDC precipitation values (0.1 mm)

For the 15 days presented the difference between the values from the NCDC and the ECA&D dataset are all under 0.52 mm so values from the datasets will be taken as valid. The ECA&D database has a more climate database for European countries but the stations where the data is available are less than the ones from NCDC. Another problem with automating data

processing from the ECA&D database is the fact that the station locations for the meteorological stations are not available in a GIS format. The ECA&D uses station IDs that are different from the World Meteorological Organization station IDS so the assignment of climate data to a spatially distributed layer is not possible through an automated script.

Therefore, the NCDC database was used to extract climate variables through an automated script.

3.5.2 Rainfall data processing for this model using Python scripts

I created some Python scripts to process the NCDC data and extract the needed information, if available, in a spatially distributed manner. The scripts can be used to process the NCDC information to obtain:

- a feature dataset with attribute values representing the GSOD database measurements for a certain day
- a feature dataset with attributes for a single measurement for a day
- a feature dataset with the mean value for a certain field during a given period of time as an attribute
- a feature dataset with the sum of values for a certain measurement during a given period of time as an attribute
- a feature dataset with the sum of values for a certain measurement during a year as an attribute
- a feature dataset with the mean value for a certain measurement during a year as an attribute
- a raster dataset with the Kriging interpolation of any of the elements extracted (the interpolation is for a single field in the NCDC tables)

If certain data is not available for the selected day or period the script will ignore the station where the data is missing and use only the stations with data in processing. The final dataset will only contain the stations with existing data for the selected period.

The scripts can be used to extract any kind of climate variables from the NCDC dataset, not just rainfall data. As the tables in section 3.5.1 show, the temperature is the most complete column in the NCDC GSOD database in the Apuseni Mountains area. For this reason, the application example for the scripts processing NCDC data uses the mean daily temperature as the data source. The rainfall data does not contain continuous measurements in the older NCDC datasets (before 2007) so for applications in this period the data had to be extracted from other sources.

The processing of data in the cases when NCDC data is not available was done using data directly from different weather station or using only the important weather stations in the interpolation.

For the representation of the daily precipitations in the northern part of the Apuseni Mountains I have used an algorithm based on the Kriging method presented by Crăciun in 2010. The Kriging method was developed in 1960 by the French mathematician Georges M Atheron, basing it on the master dissertation of G. Krige (mine ingineer). The method was initially used in detecting gold deposits but later it was applied in different other areas, among which climatology. The hypothesis of the method is that there is a self correlation on small distances of the values of the interested spatial variable.

The algorithm includes the relief as a fundamental factor for spatial distribution of precipitation values. This model may be included in the category of a residual Kriging, a tendency model. The succession of the operations that are the basis of this algorithm are synthesized as a graphic in Fig. 3.15 (Crăciun, 2010).
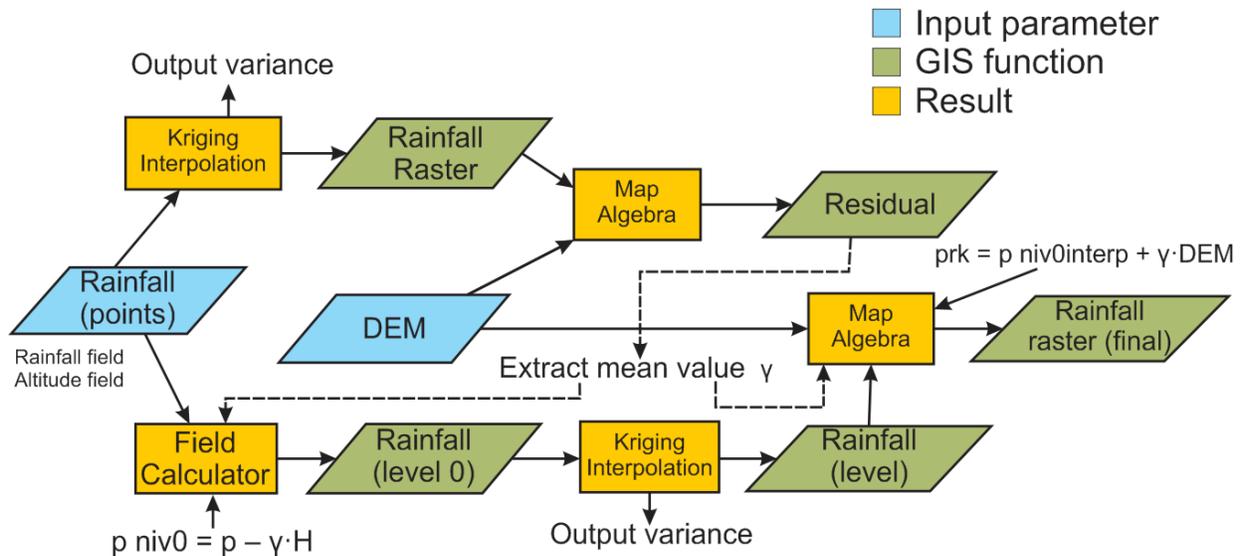
Fig. 3.15 – Interpolation workflow for the Residual Kriging method (Crăciun 2010)

The main phases of the algorithm for conceiving the precipitation map using this procedure are the following:

a) interpolation of point values of precipitaions measured at meteorological stations/pluviometric points, using the Kriging method;

b) determining a residual value which should illustrate the way in which the precipitations are influenced by altitude:

1 transforming the digital model of the terrain in mm;

2 establishing the difference between the digital elevation model in mm and the precipitation layer that resulted after the kriging interpolation;

3 re-conversing the layer in m;

4 obtaining the gradient through difference between the digital elevation model and the layer obtained in phase 3.

c) creating a field in the attribute table of the point type layer concerning the measured precipitations and adjusting the precipitations at level 0 by appling the following relation:

$$p \; niv0 = p - \gamma \cdot H \qquad\qquad (3..1)$$

-where:

p – quantity of measured precipitations (mm)

149

γ – mean value of the gradient determined in phase b

d) interpolation of the new field using the kriging methid

e) creation of the final map by application of the following relation:

$$p\ rk = p\ niv0interp + \gamma \cdot DEM \qquad\qquad (3.2)$$

- where:

Prk – precipitations interpolated by Residual Kriging method

pniv0interp – precipitations adjusted to level 0 and interpolated by kriging method

DEM – Digital Elevation Model

Using this algorithm datasets of rainfall that take relief into account can be obtained. These maps ofer better accuracy then the datasets where the altitude is ignored and therefore only the values in the points and the distance between these influence the interpolation.

Application Example: Mean temperature for 6-21.06.2000 in the Apuseni Mountains

The NCDC database processing scripts were used to calculate the mean temperature at the middle of June 2006 in the Apuseni Mountains area. The database from NCDC consisting in .op files (contain values separated by space for each record) was downloaded for the following weather stations in the area (Fig. 3.16)
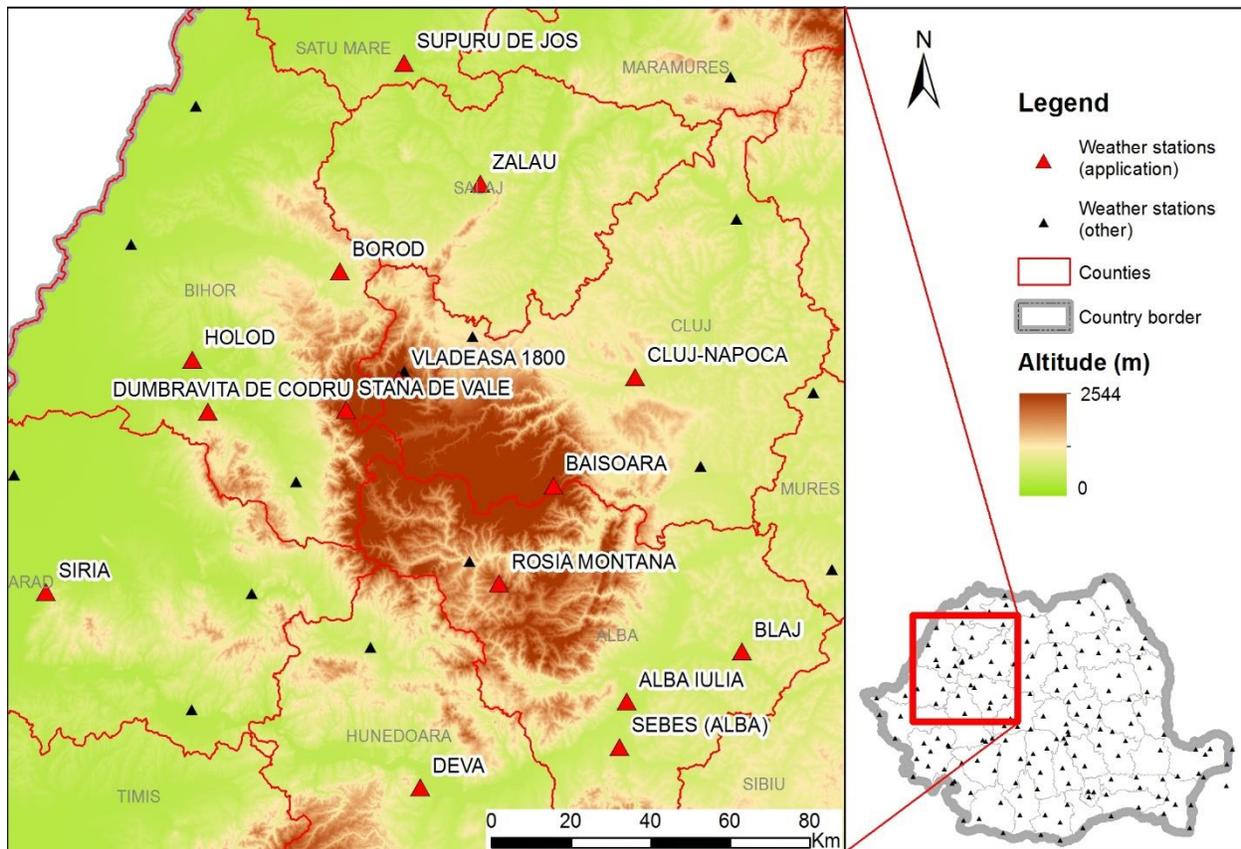
Fig. 3.16 *Weather stations used for the application*

The data downloaded was stored in a folder given as a parameter to the script that processes the data. The syntax for the script is the following:

```
ProcessMeanPer.py in_NCDC_Data_Folder, in_Start_Date,
in_End_Date, out_Dataset, in_Station_Locations, in_Field
```

The script can be called directly from a command line or a form in ArcMap can be created to call the script directly from there.

The script does the processing in the background and gives messages about the progress and the data extracted:

```
Start Time: Sun Jan 15 12:21:14 2012
Running script ProcessMeanPer...
Process mean of values for a period from same year
Period: 6/1/2000 : 6/21/2000

Getting the records for the period
Processed 25 lines
```

151

```
Processed 50 lines

Creating point feature class. Field: TEMP

Inserting points in feature class
Point Object
Point: 47.467 ; 22.783
Value: 12.3888888889
Point: 47.183 ; 23.083
Value: 12.2777777778
Point: 46.983 ; 22.6
Value: 12.8611111111
Point: 46.783 ; 22.117
Value: 14.4166666667
Point: 46.683 ; 22.617
Value: 6.66666666667
Point: 46.767 ; 22.8
Value: 4.90740740741
Point: 46.783 ; 23.567
Value: 19.4391534392
Point: 46.633 ; 22.167
Value: 12.3333333333
Point: 46.533 ; 23.317
Value: 7.38888888889
Point: 46.25 ; 21.65
Value: 11.6111111111
Point: 46.317 ; 23.133
Value: 9.98148148148
Point: 46.067 ; 23.567
Value: 15.2222222222
Point: 46.183 ; 23.933
Value: 14.1944444444
Point: 45.867 ; 22.9
Value: 13.6666666667
Point: 45.967 ; 23.55
Value: 14.1851851852

Raster Interpolation of result

Completed script ProcessMeanPer...
Succeeded at Sun Jan 15 12:22:35 2012 (Elapsed Time: 1 minutes 21
seconds)
```

After the script is run the datasets that were created can be automatically added to ArcMap (if the script is run from within the application). The output consists of two datasets: a raster dataset and a feature class or shapefile.

The first dataset that is output is the shapefile containing the points along with the data stored as an attribute (Fig. 3.17).
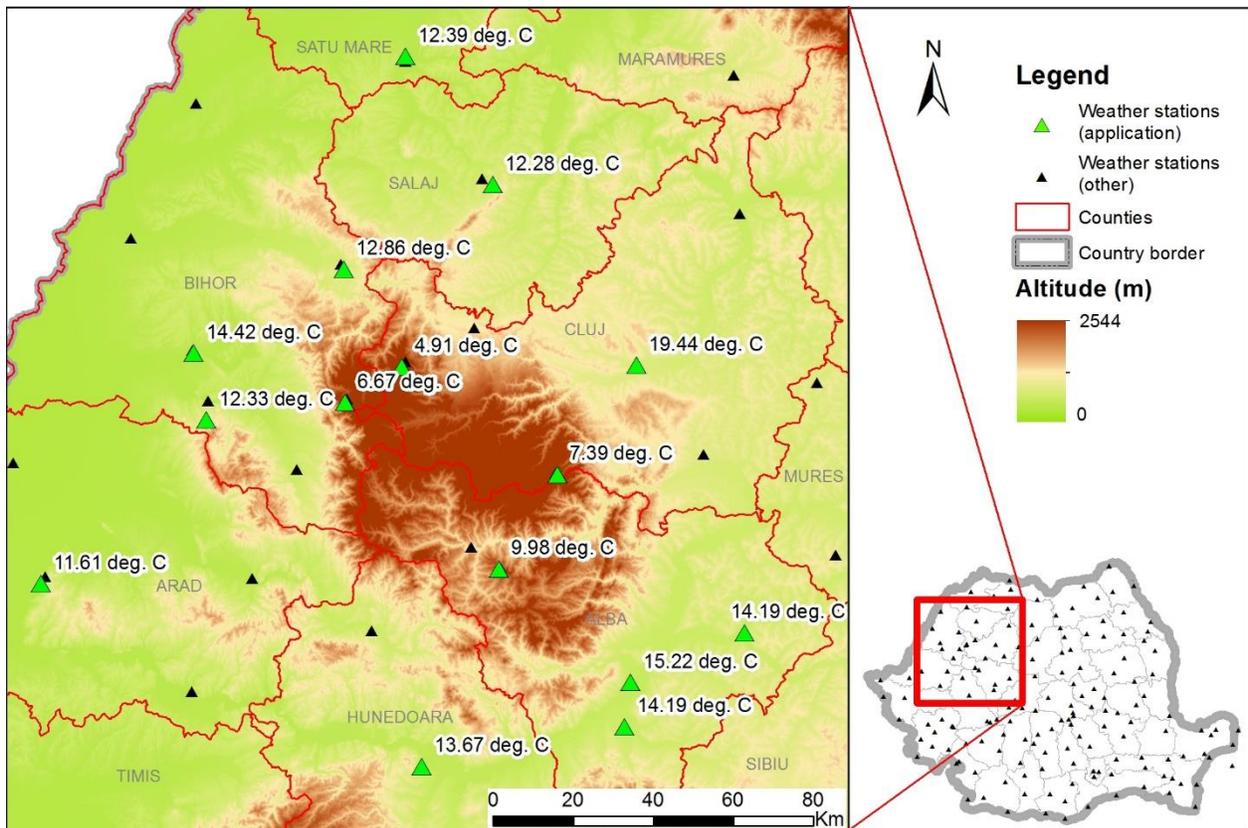
Fig. 3.17 *Temperature values obtained for each station*

The script continues with the interpolation of the result and creates a raster dataset with the values of the interpolation. For this application the raster created can be seen in Fig. 3.18
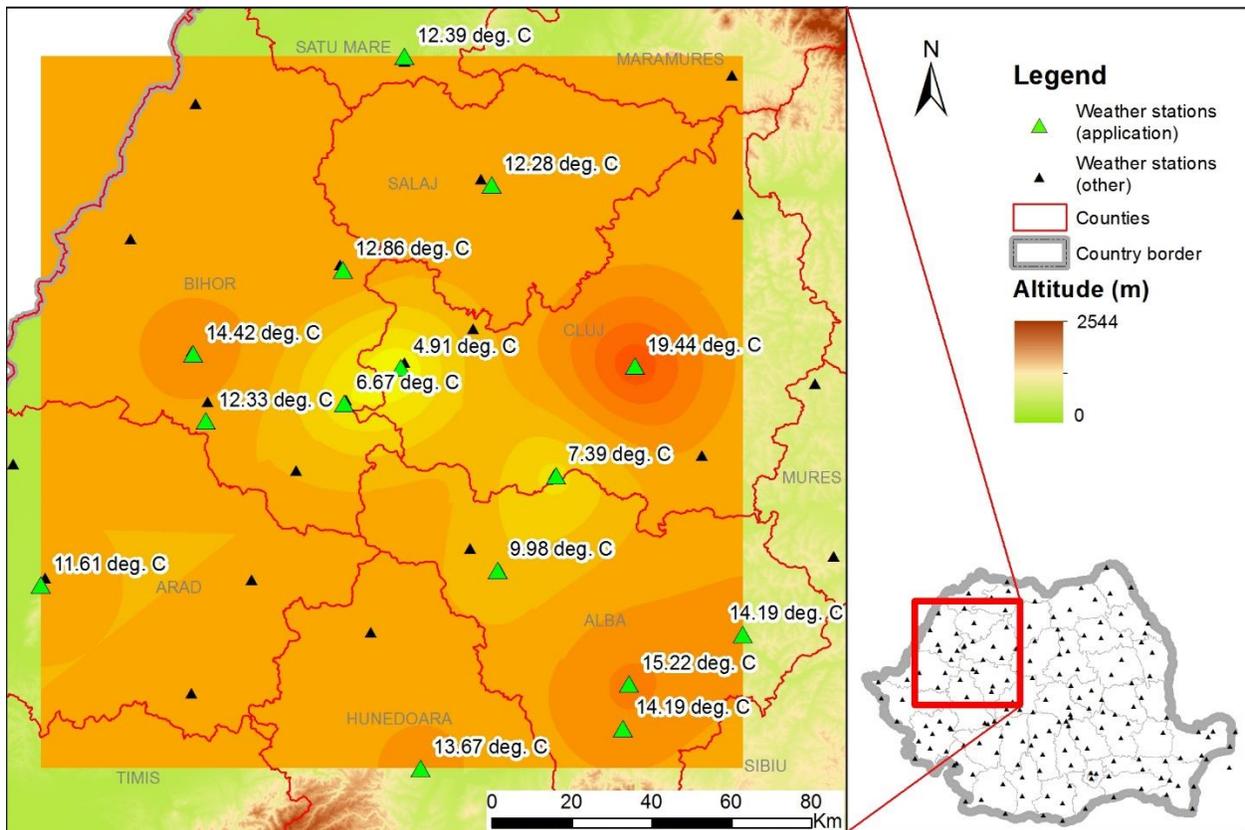
Fig. 3.18 Kriging interpolation of temperature values obtained

After the interpolation is finished the analyst can use the dataset created by the script for the next steps of the analysis. In the current case, the obtained layers can be used as inputs to the model created.

4. Spatial discretization according to the drainage network

Flow estimation at a point of interest on a watercourse involves two processes: modeling surface flow generation and routing of water from surface runoff into watercourses. Regardless of the method of calculation, drainage hydrograph will be available only in certain points on the surface of a catchment.

As water moves downstream, the hydrograph behaves like a wave and will be attenuated (flow time of water through the hydrograph rises and discharge falls). The process describing the water movement and surface water concentration through the watershed is called flood hydrograph routing. Flood routing offers the opportunity to appreciate the magnitude and arriving time of maximum flow based on upstream measured or estimated flood conditions. (Smith et al. 1998)

In case of an automated model, the points where flow calculation will be executed should be chosen according to the characteristics of the study area. The model created in this paper will automatically choose the outlet for each sub-basin in a study area.

After the watersheds are delimited, GIS software can be used to obtain basins with certain characteristics in the area. In terms of quantity several types of classification can be made based on river basin morphometry.

Morphometry is defined as the measurement of form. Morphometric studies in hydrology were made by Horton and Strahler in the 1940s. Their goal was to obtain properties of streams using measurements of different attributes for these streams.

One of the first attributes to be quantified is the hierarchy of stream segments according to a classification system. Classification of river networks was always in the attention of researchers; depending on the purpose, basic criteria taken was the configuration of stream connections on the terrain or a series of dimensional elements (basin area, length of streams, depth, direction, discharge, river position relative to the main collector etc). (Zăvoianu 1978)

Researchers studied the possibilities of stream classification according to these criteria.

One of the first attempts at stream ordering was carried out by Gravelius (1914) (quoted by RE Horton 1945), who believes that the greatest collector is the first order from source to mouth. Its direct tributaries have order 2, courses that flow into one of order 2 has order 3, etc. This analysis makes no distinction of watercourses depending on length, flow or pool area.

Horton proposed, while studying infiltration and surface run-off generation in 1945, a law of composition of streams reversing Gravelius' system. After Horton's system, the first order is assigned a watercourse consisting of concentrated flow. A second order stream is one that receives at least one tributary and has only first order tributaries. When a 2nd order stream meets another 2nd order stream a stream of order 3 is formed etc..

Strahler took this system in 1952 and created his own system for classifying river networks. The Strahler classification system (Fig. 4.1) starts from the basic water course considered a first order stream. A second order stream occurs as a result of the union of two first order streams. A stream of order three is the confluence of two second order streams and so on. If a stream receives a lower order tributary it does not change its order. Studying rivers classified by the Strahler classification system is observed that the physical and geographical conditions for streams that have the same order are similar in terms of catchment area, the average length of the water network, average slope, average water flow rates, etc. (Zăvoianu 1978)

Another classification system is the Shreve stream ordering method. Shreve proposed a systen which takes into account all tributaries and provides a realistic picture on the size os a basin in terms of the number of stream segments it contains. The Shreve method accounts for all links in the network. As with the Strahler method, all exterior links are assigned an order of 1. For all interior links in the Shreve method, however, the orders are additive. For example, the intersection of two first-order links creates a second-order link, the intersection of a first-order and second-order link creates a third-order link, and the intersection of a second-order and third-order link creates a fifth-order link. (Tarboton et al. 1991)

Because the orders are additive, the numbers from the Shreve method are sometimes

referred to as magnitudes instead of orders. The magnitude of a link in the Shreve method is the number of upstream links.
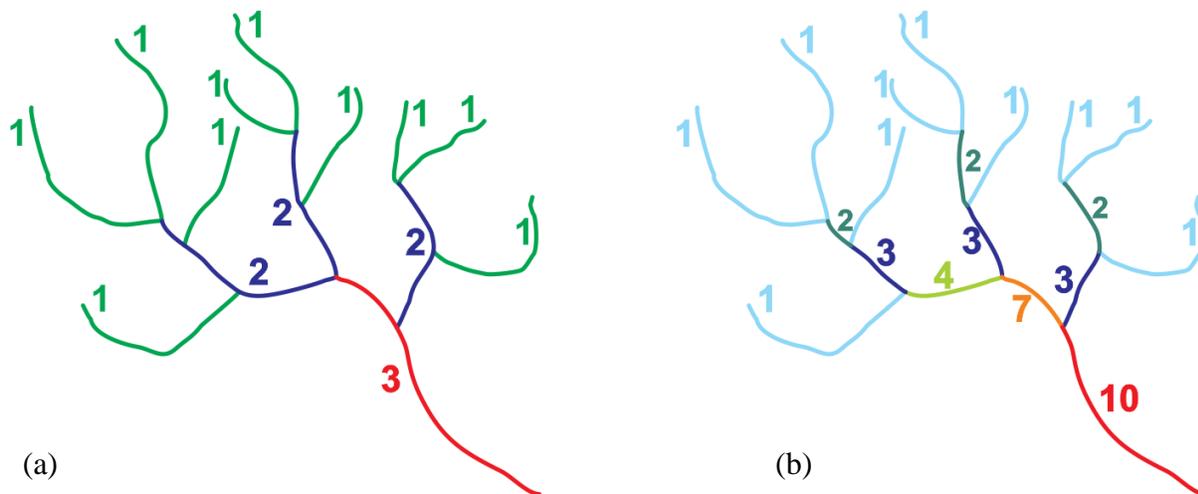
Fig. 4.1 *Strahler stream ordering (a) and Shreve stream ordering (b)*

Since the contribution of surface flow to generate large discharges is achieved by concentration of runoff from the slopes and accumulation of flow generated by several catchments, the Strahler system is suitable to divide a given area. Basins of order 1 are those that have the important quantity of flow caused by runoff. Basins of any order greater than 1 are created by concentrating the water flow from river tributaries of lower order plus flow from interbasin areas.

If the hydrograph from a catchment of a certain order is known, this could contribute by translation and flow accumulation to the discharge of the higher order catchment (depending on travel time of water between the basin of lower order and outlet of the basin of higher order).

In order to achieve the necessary calculations for flow translation and accumulation the ordering of streams hidrographs in each basin is necessary. A Strahler classification of river basins and determination of a ranking based on water accumulation could facilitate this. The problem is represented in GIS as an ordered graph for a network of drainage basin areas and catchments corresponding to each segment of the network.

Besides the classification of basins in the area of study it is enough to know every segment of the drainage network and the next downstream segment. A structure of these segments

is equivalent to an ordered graph from computer science that can be made useful design flow calculations.

The next section will describe a model created generation of Strahler order for each watershed in an entire area.

4.1 The need for spatial discretization in surface runoff modeling

Determination of catchments in an area and their classification

To automatically calculate flows at certain points the area needs to be split into catchment areas and interbasin areas in which discharge calculations can be computed. The first division of the catchment areas can be achieved using the watershed tool, available in any GIS package. This function divides the entire study area in catchments and sub-basins.

To determine river basins and to make classifications the GIS functions specific for these tasks were applied. Functions were available in the ArcHydro extension for ArcGIS presented in section 2.2.

Arc Hydro is a data model and set of tools for integrating geospatial and temporal information about water resources, a model that can run within the ESRI ArcGIS geographic information system. Although implemented in a commercial GIS,the data model and tools are in the public domain and are freely available as an extension to ArcGIS. (Maidment 2004)

Arc Hydro was developed by a consortium of data providers and users in the field of water resources, coordinated by the Centre for Research in Water Resources at the University of Texas at Austin. Arc Hydro was designed to facilitate the operations required by a hydrologist to obtain essential datasets for modeling. ArcHydro not model the functioning of hydrological processes, but provides only the basic environment for the modeling.

Using ArcHydro, a user can obtain data sets for primary operations in hydrological modeling: the direction and accumulation of flow, flow path length, definition of streams according to the accumulation of water, stream segmentation and determination of flow of water through them and watershed delineation.

4.2 Spatial discretization of a catchment according to stream orders

Spatial discretization procedure in the study areas for this model

The module for determination of the discretization based on the drainage network used several of the functions available in ArcHydro. The model obtains all the catchments from a certain area corresponding to a specified stream order according to the Strahler classification. The workflow implemented for this purpose is shown below.

The model has several components achieving the following needs: set the environment for the operations, extract information on the streams, catchments and outlets, determine the Strahler stream order corresponding to each catchment outlet and extract catchments with the requested order.

Configuration of the system for running the commands

This configuration is required to run any ArcHydro commands and creates the locations where the ArcHydro tools will save the data. The structure of the configuration function component is shown in Fig. 4.2. The system configuration requires a DEM as a parameter and locations will be created in the same folder with this DEM.

The folders and the geodatabase resulting from the application of the configuration tool are the locations where the results of all operations from the ArcHydro tools will be automatically stored.

In addition to this configuration, a proper configuration of the environment for the ArcGIS geoprocessing tools is necessary. These tools do not belong to the ArcHydro extension, so their configuration will not be discussed here.
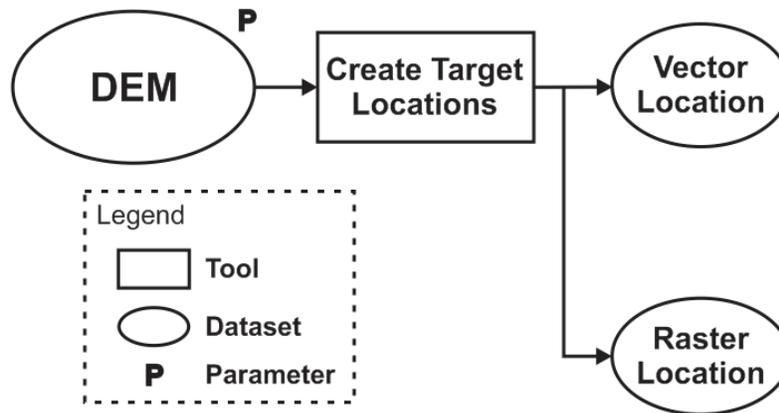
Fig. 4.2 – *Creation of target locations for the ArcHydro tools*

Determination of streams, catchments and outlets

This is the most consistent component of the model created using most of the basic functions available in ArcHydro. The structure of this component can be seen in Fig. 4.3.

Operations taking place in this part of the model that were presented in chapter 2.5 are:

Fill sinks - fill small sinks that appear as an error in recording mode or generation of DEM's. Thus, the following hydrological features will run correctly without problems occuring due to erroneus sinks.

Flow Direction – Direction of flow is used for all other operations related to water runoff

Flow Accumulation - The following standard operation in any catchment analysis of is the Flow Accumulation, which creates a raster with cell values equal to the upstream cell count (number of cells that contribute flow to the cell). Flow Accumulation is used later to define streams and catchments.

Stream Definition – Further on the branch for automatic extraction and segmentation of streams, Stream Definition function is used. This function creates a raster with a NoData value for areas without a stream and "1" where a stream that is considered permanent is defined.

ArcHydro implemented the most common method for obtaining streams. First a grid with surface area drained by each point is calculated, then each cell corresponding to an area larger than a predefined value is assigned a value of 1. For the best results, this value should be obtained from a graphical representation of slope compared catchment area or on the basis of the fact that the starting points of the streams correspond to the transition from convex slope profiles to concave

160

profiles, as shown by Tarboton and his collaborators (Tarboton et al., 1991, cited by Montgomery). Because the other methods are difficult to achieve automatically or need the intervention of an analyst in determining the initial points of the streams, the model keeps the default method defining streams according to contributing area.

The presented method is not best suited to represent natural variability of the landscape caused by geology, altitude and other factors. For this reason Peckam (Peckham, 1998, quoted by Hengl) presented a method that could be more robust: creation of a raster with Strahler stream orders for the whole surface of the DEM, then removing flow vectors with an order below a certain value. This method could be applied when the comparison between the results obtained in the field with the GIS results does not lead to the desired similarity.

Stream Segmentation – For segmentation of streams between the confluences this tool may be used. The method takes streams in raster format and creates a raster that has the value of each cell according to the segment limited by two confluences or a confluence and a free end.

Drainage Line Processing – This is the method used by ArcHydro to convert representation of streams from a raster format to a vector format. The method converts the stream segment raster obtained with the previous tool into a polyline vector layer and creates a network of rivers in vector format.

Catchment Grid Delineation – This function defines each catchment and creates a raster where each cell has a value corresponding to the first segment that drains water from the cell. After delineation the catchments are obtained in a raster format, they can be converted to polygons to perform vector analysis on them.

Catchment Polygon Processing – Conversion of the catchment raster obtained in the previous step to a polygon layer is achieved with this function. The result is a polygon layer with each polygon corresponding to one of the catchments determined earlier. These polygons are created with correct topology so they do not overlap and have common sides with their neighbors without any space between neighboring polygons.

Drainage Point Processing – Last operation from this component of the model is getting close points. For each catchment or interbasin area (which will then belong to a larger order catchment) the outlet positioned in the lowest point of elevation is determined. The cell in this point will drain all the water in the catchment. Points, stored in a point-layer will receive the

appropriate attributes of catchments that they close and main stream segments draining those catchments.

After running this model the user will have all catchments and interbasin areas along all the corresponding streams and outlets in vector format. In addition to this, the basic raster analysis (flow direction, flow accumulation, DEM reconditioning) will be executed.

Determination of the Strahler/Shreve stream order for outlets

To enable the discretization according to the order of river drainage network is necessary to know the stream order for each outlet obtained above. The order attribute that will be assigned to each outlet is in fact the stream order of the main stream draining through the outlet. This is required because the watershed delineation functions require points representing the specific outlets for the watersheds that are delineated.

Since the functions of ArcGIS only allows for ordering of streams in raster format and the modeling was done in ArcGIS, an artifice was required. To determine the order attribute for outlets, the raster with stream orders was previously generated and then the raster cell values representing the stream orders were taken as attributes for outlet points. The structure of this component can be seen in Fig. 4.3.

Processes taking place in this part of the model are:

Stream Order - Streams in raster format created in the previous component are retrieved and ordered according to the Strahler system. Performing stream ordering function needs a digital elevation model used to determine the direction of water flow through each segment. The result of applying this function is a raster which has the value of each cell belonging to a stream of water equal to the order of the corresponding stream.

Extract Values To Points – Orders for the outlet points previously created and given as a parameter to this component are obtained by extracting values from the stream order raster and storing these values in an attribute of type Integer.

After running all the tools from this component the outlet of each catchment in the study area has a stream order attribute corresponding to the main stream in the catchment.
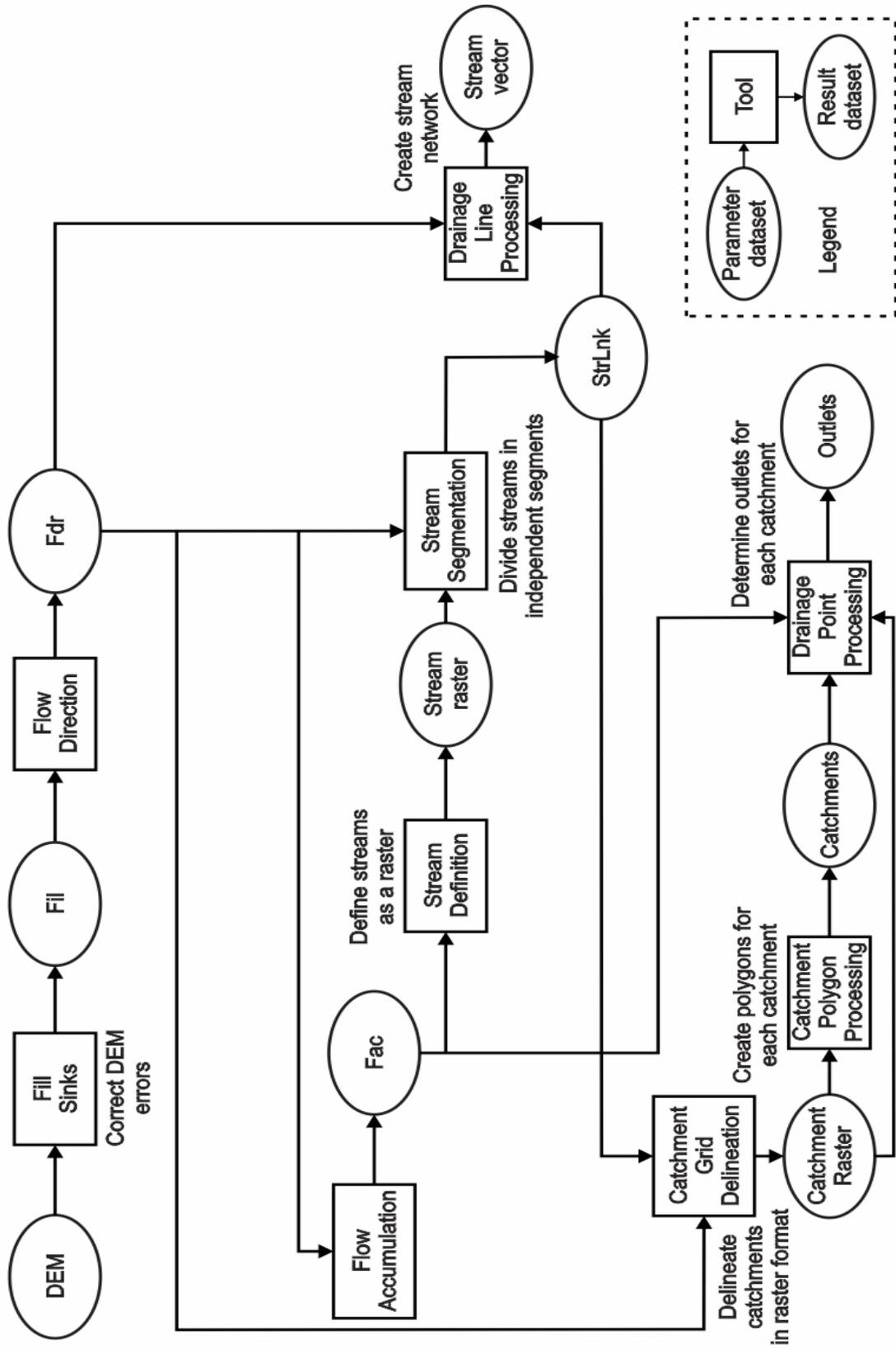
Fig. 4.3 – *Determination of streams, catchments and outlets*
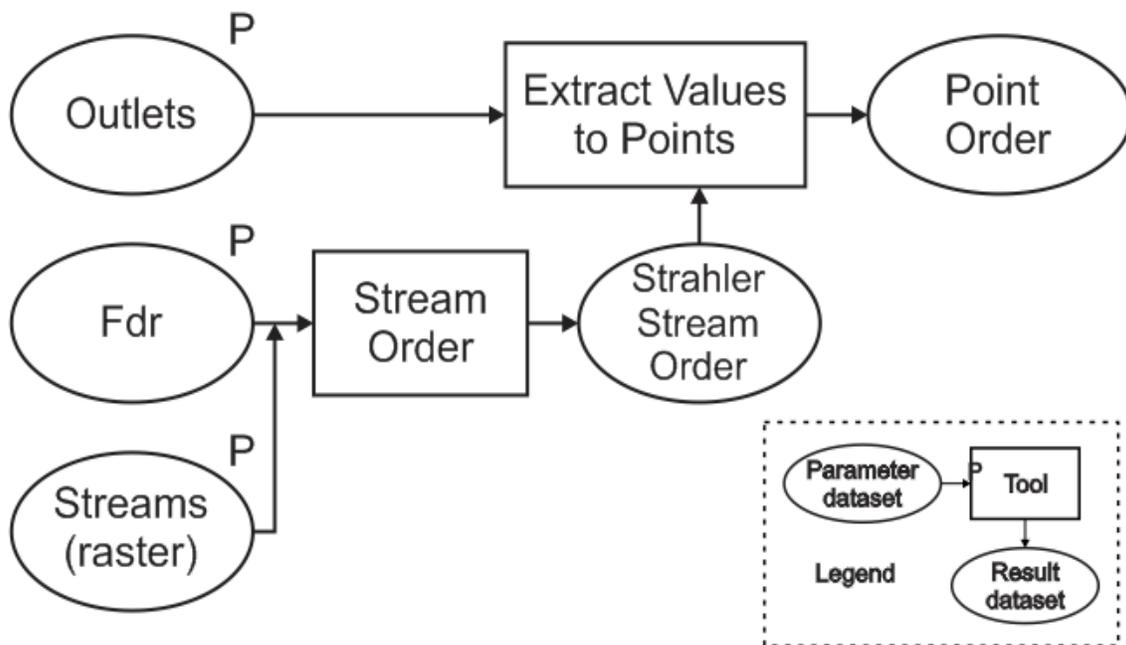
**Fig.4.4** – *Extraction of order attribute for drainage points*
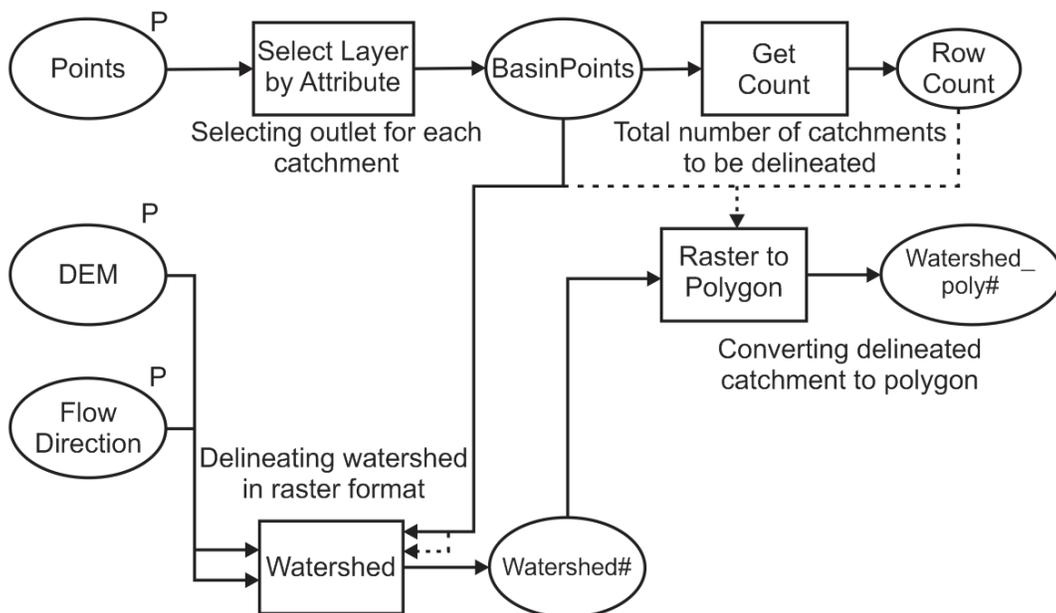
Fig.4.5 – *Determination of the stream order for basins*

Creating catchments with the required order

The last part of the model developed uses the outlets with the order stored as an attribute to create corresponding catchments for each outlet. The catchment will be stored in polygon format, and have an attribute equal to the order of the main stream that drains water from the basin. The structure of this component can be seen in Fig. 4.5

This component of the model is more complex because the ArcHydro 1.4 Extension for ArcGIS 9.3 does not offer the possibility to automatically delineate watersheds for each point in a set of points in another way than by calling a tool from the toolbar with user intervention. Because the aim is to automate processes for obtaining these watersheds, a model that extracts the watersheds without user intervention was created.

The model presumes creating a catchment for every outlet from the outlet layer that corresponds to the condition of having the desired order. The need for iteration through the outlets arises and this iteration is performed automatically by Model Builder. The iteration functions offered in the Model Builder interface for ArcMap were used for this purpose. The model presumes creating a basin for each outlet from the area and it iterates each drainage point to create all the catchments.. The iteration function from ArcGIS Model Builder was used for this purpose.

The "Get Count" function in this component provides data on the number of points corresponding to the chosen stream order from which watersheds will be delineated. The model will run once for each point and create a polygon representing the area drained by that point.

Processes taking place in this part of the model are:

Select Layer By Attribute - Select the appropriate item with the ID number corresponding to the iteration number of this component. The first run will select the item with ID 1; the second will select the item with ID 2 and so on, until the end of the points of the layer. Because points were determined automatically, IDs are already in an ascending order and will be consecutive, so iterations will run correctly.

Watershed - Creates basin corresponding to the selected point from the current iteration. Basin will be stored in raster format with raster name corresponding to the number assigned in the

iteration (ie. equal to the ID of the outlet point). Raster cell value will be equal to the order of the catchment taken from the attributes of the outlet.

Raster To Polygon - Converts the catchment created as a raster in a polygon and adds it to a layer of this type. Basin will have the attribute corresponding to the Strahler order of the main stream within it.

After running all four components the user will have all of the basins in the study area in polygon format and know their order. Using the available NextDownId ArcHydro function, the user can order catchments depending on the route of water flow through them. The order will be represented in computer memory as a directed graph that will facilitate translation of the hydrographs generated upstream to downstream basins in order to minimize the amount of calculations required to determine hydrographs at each confluence.

The models presented used the digital elevation model in raster format to get the catchments in a specific area of interest and stored an attribute for each catchment corresponding to the order by Strahler classification. These catchments can be ordered by flow accumulation and the next catchment downstream can be obtained for each of the catchments, forming a directed graph.

Based on this graph a calculation can be made for hydrographs in each catchment and then the upstream hydrographs can be translated and added to the total accumulation to get a final discharge hydrograph.

4.3 Results of spatial discretization in the study areas

The spatial discretization method was applied for each of the study areas presented in the first chapter and the results of the discretization were used as parameters to the discharge hydrograph model.

The results of the discretization for the Zlatna basin are presented in the following section. The location, altitudes and other characteristics of the basin presented can be seen in chapter 1 and fig 1.7

The first step is the component that determines the streams, catchments and the outlet for each catchment. The base layer id the DEM for the basin that can be seen in the Fig. 4.6
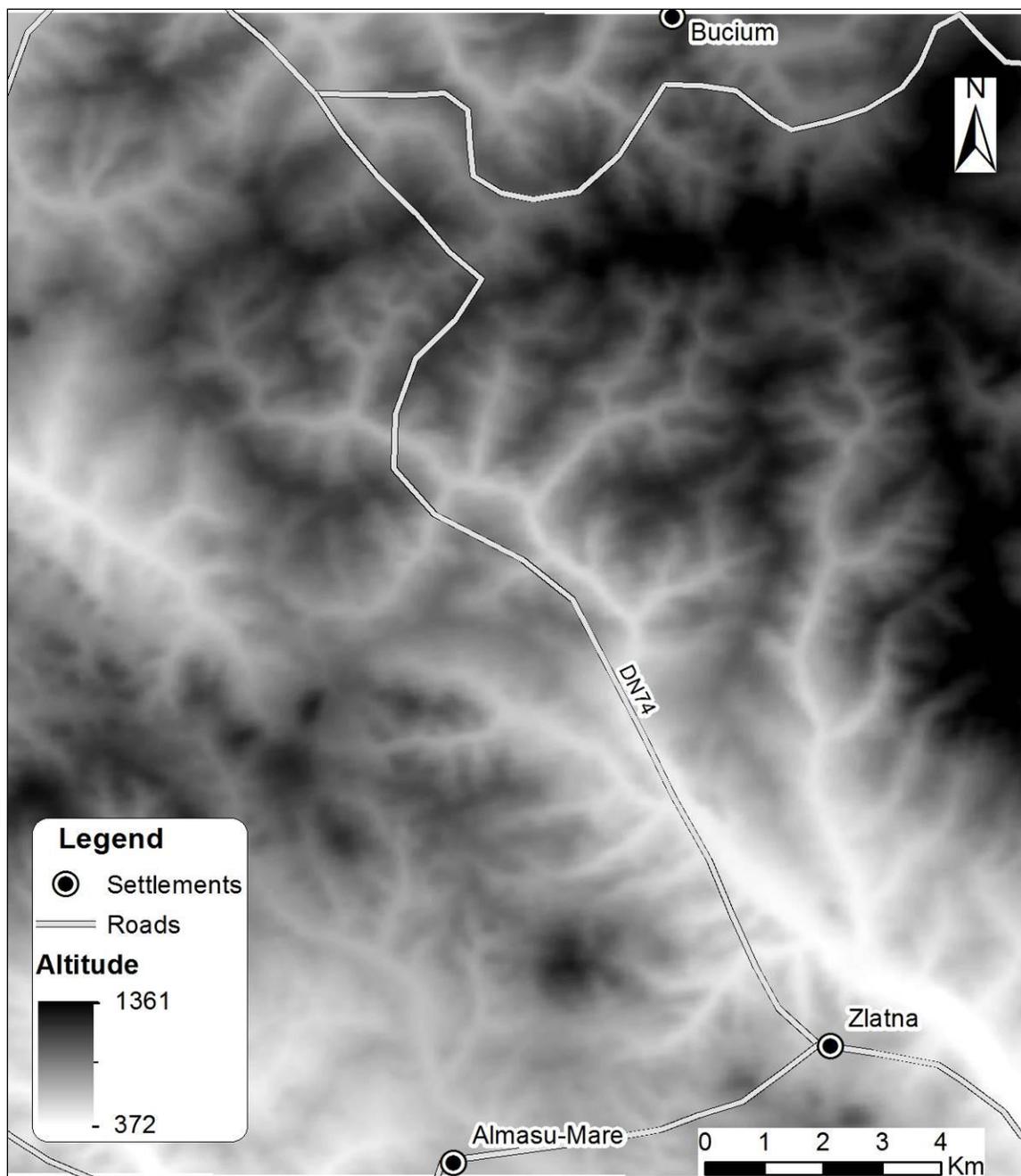
Fig. 4.6 *DEM in application area for spatial discretization*

The first module is used for obtaining the streams and catchments in the study area in a raster format. The user can select a number of cells to define a stream and every cell with the flow accumulation over this value will be considered a stream.

In this case, a value of 8500 cells lead to the determination of 27 drainage points from which 13 were on the main stream in the area and the others were on secondary catchments or on the limit of the study area.

167
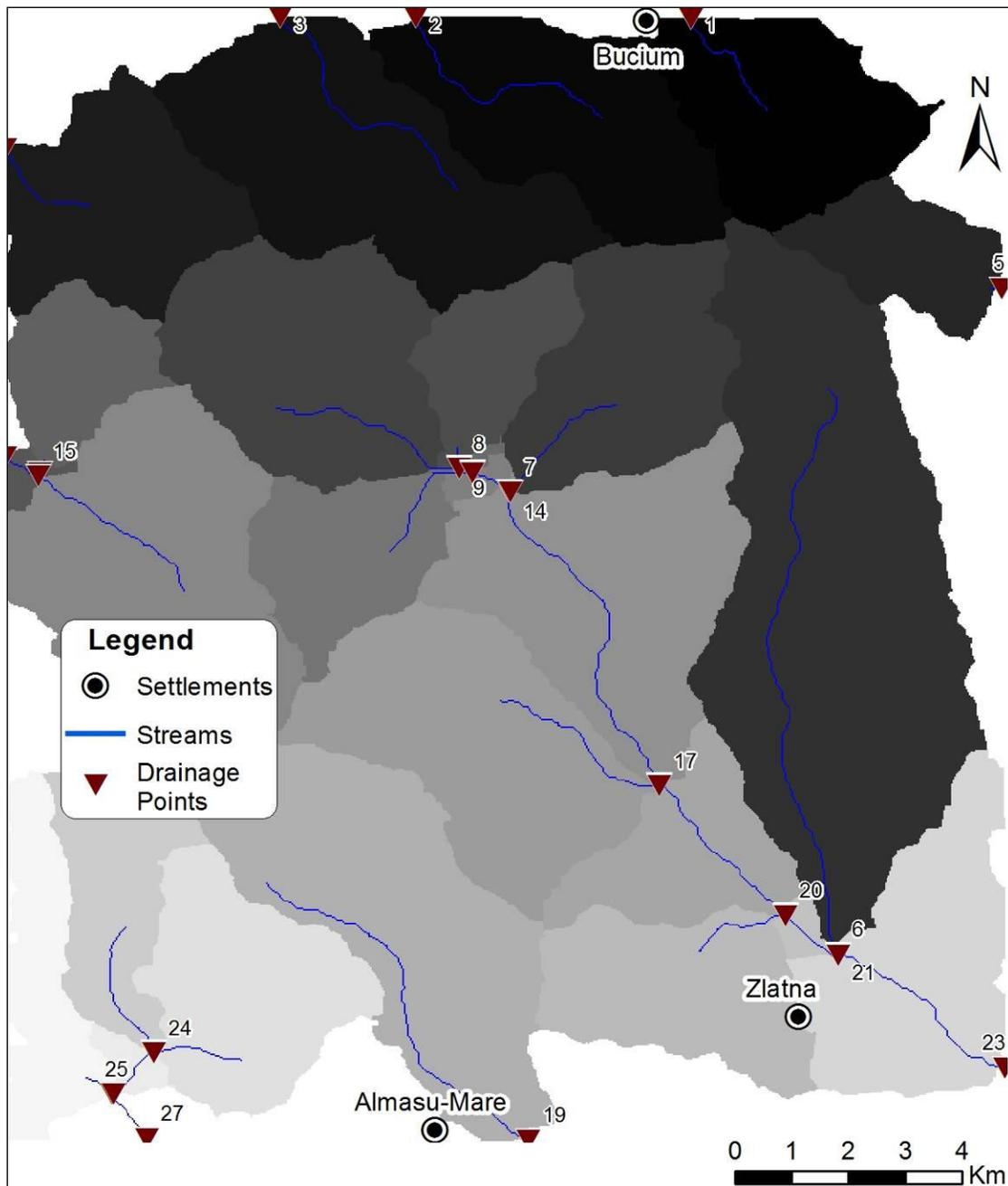
Fig. 4.7 *Streams, drainage points and catchments*

The second part of this module is used for determination of the catchments in a polygon format, as seen in Fig. 4.8. The catchments are presented and numbered along with the corresponding drainage points.

After this module ends, the user can choose between applying the hydrograph model in any of the drainage points, continuing with the ordering of streams or obtaining the adjoint catchments in a polygon format. In this case, I continued with stream orders for the entire area.
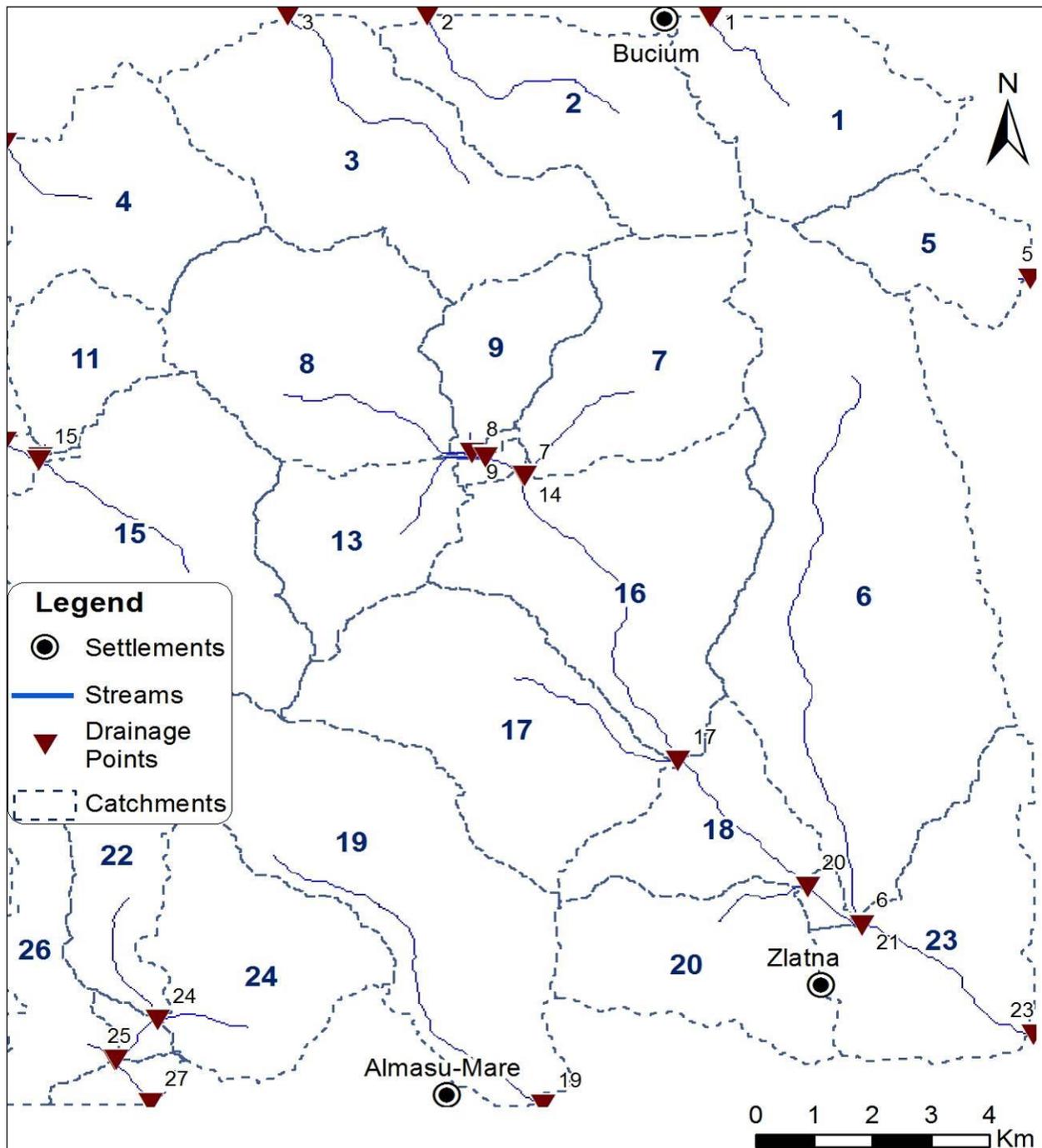
Fig. 4.8 *Catchments in polygon format and drainage points*

After the catchments are delineated in polygon format, the stream ordering is the next step in the application. The stream ordering can be processed using the Strahler or the Shreve method, and the user has to choose between these two. In this case I chose the Shreve method because it gives a more clear perspective on the size of a catchment. The result of the stream ordering can be seen in Fig. 4.9
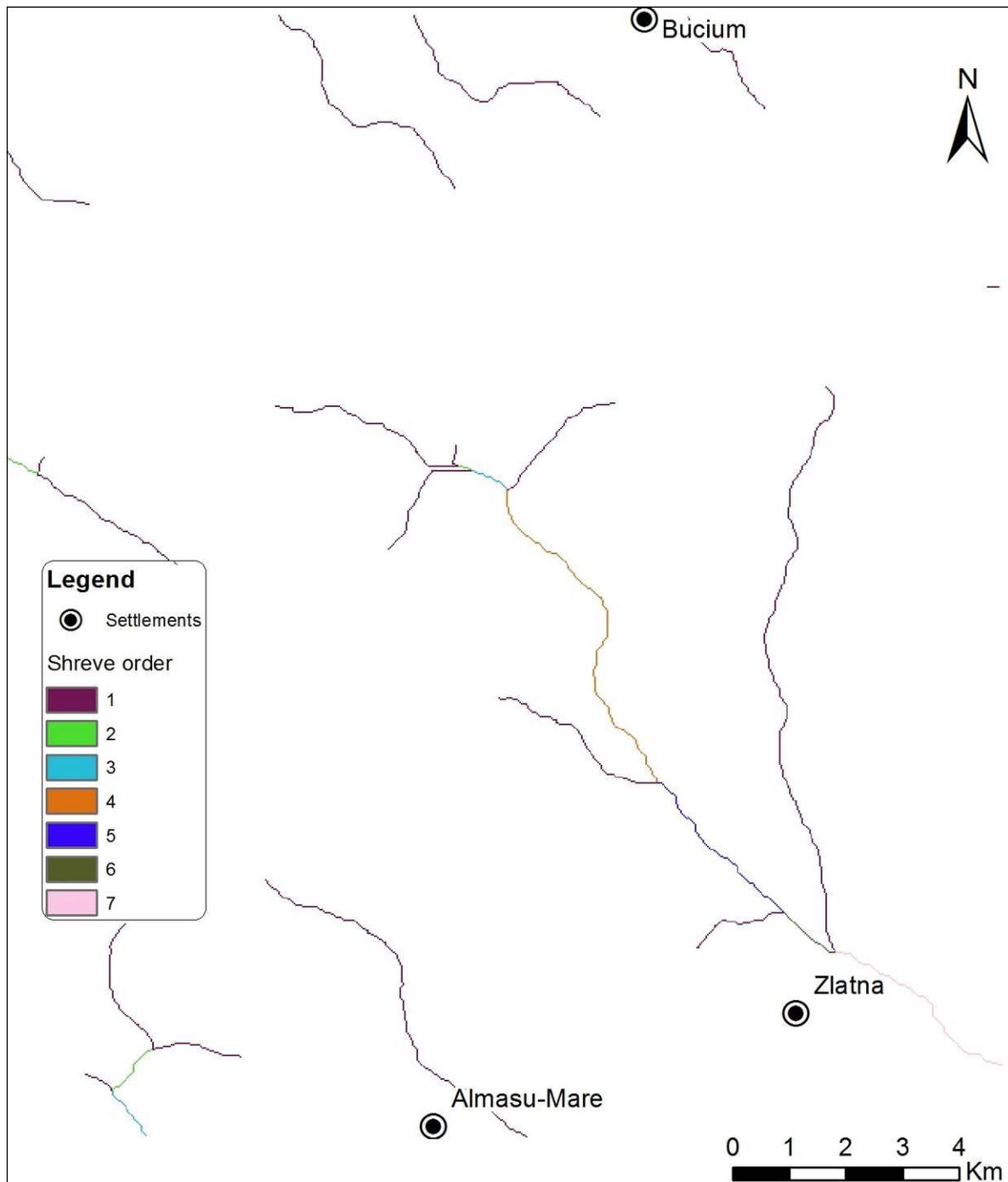
Fig. 4.9 *Stream orders for the study area*

Due to the fact that the value of 8500 cells for stream definition was not chosen according to measurements in the field or other methods that determine stream origins based on terrain characteristics, the results can be different from the reality in the field. The topographic maps from Romania are also quite old so they cannot be used as a reference, so a study in the field or based on aerial images is recommended.

Once the stream orders are known, their values can be transferred to the drainage points using the second part of the module. Each drainage point corresponding to the outlet of a catchment will receive the value corresponding to the stream order as an attribute.

The result of this component can be seen in Fig. 4.10.
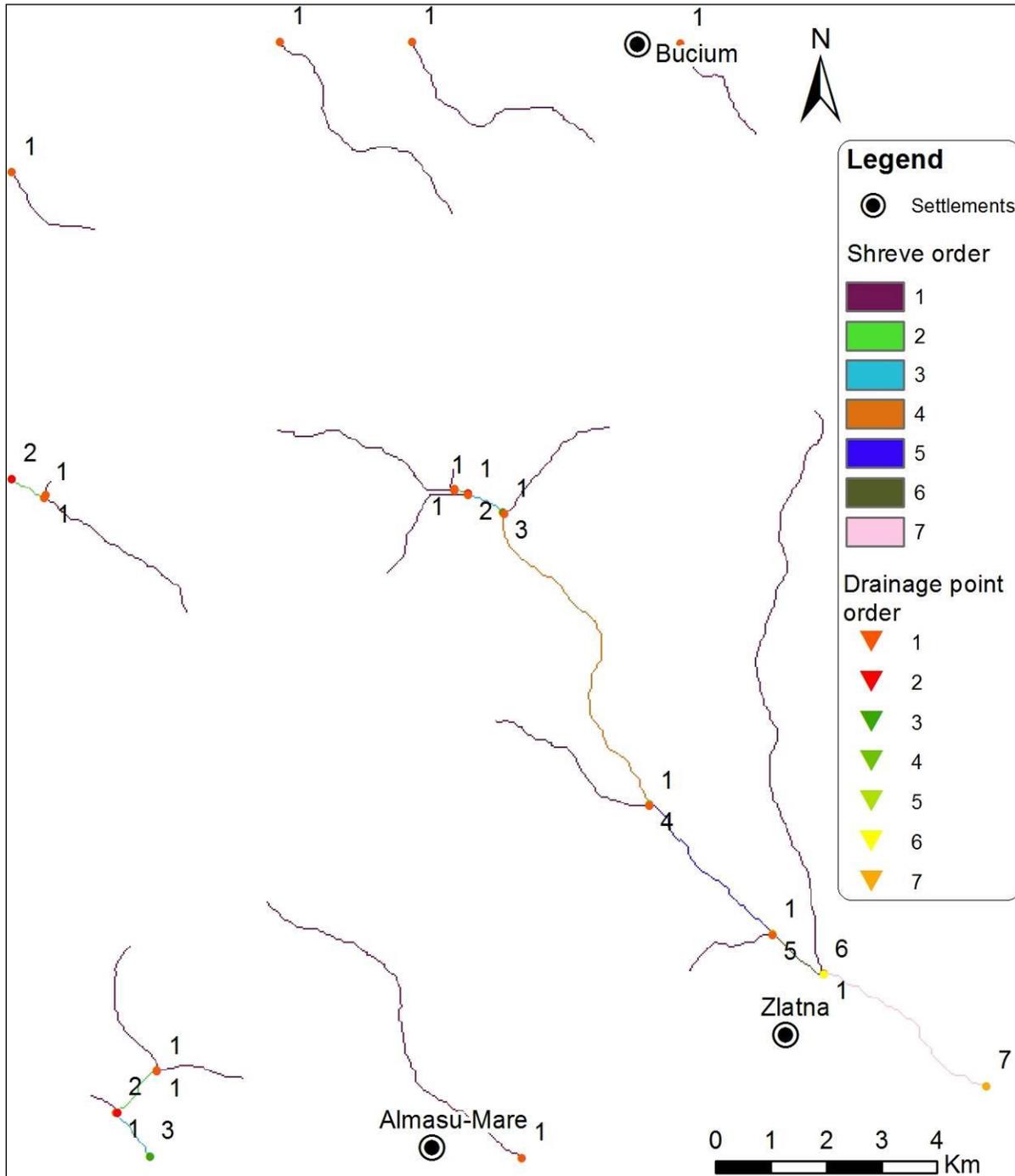
Fig. 4.10 *Drainage point order for study areas*

Another option for the analyst is the assignment of the stream order from the main stream to each catchment. Each catchment receives an attribute corresponding to the order of the main stream in the catchment.

This can be useful when the batch delineation functions are applied and offers a very good perspective on the flow accumulation from the catchments in the area. An example for the study area can be seen in Fig. 4.11

The characteristics of the Shreve stream irderung can be clearly seen in the figure. Seven first order catchments form a 7th order catchment near the Zlatna town.
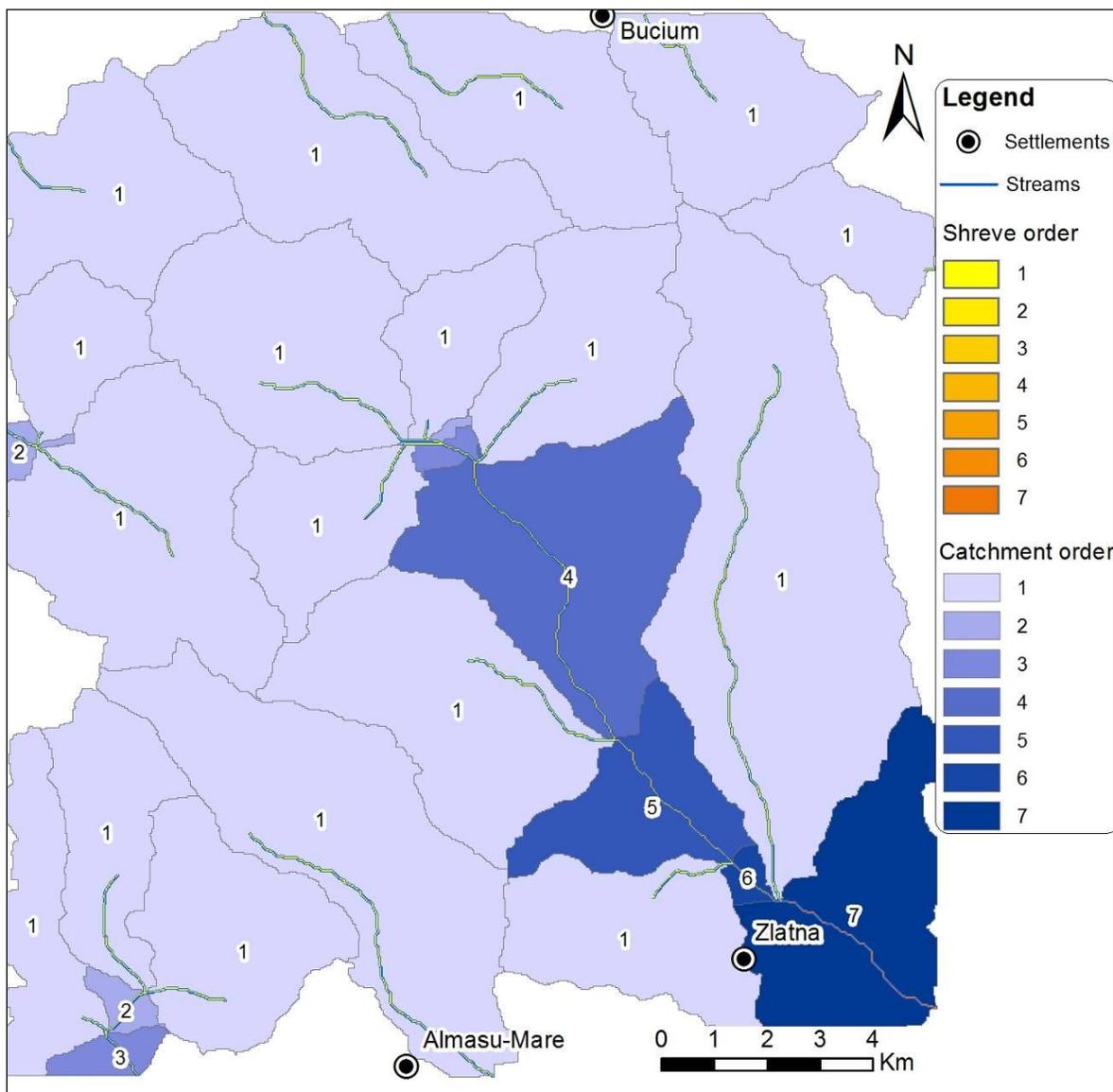
Fig. 4.11 *Catchment orders for the study area*

The last module is the iterative module used for calculation of the complete watersheds along with the stream order of the main stream. This component is useful in the determination of the entire areas for application. Once the catchment where the hydrograph will be estimated is chosen from the previous components (either from the catchment order or from the drainage point order), the user can get the watershed corresponding to that catchment and clip all the datasets to the required area for the final model application. The result of this module can be seen in Fig. 4.12

The higher order watersheds are marked with line fills and labels according to their stream order.
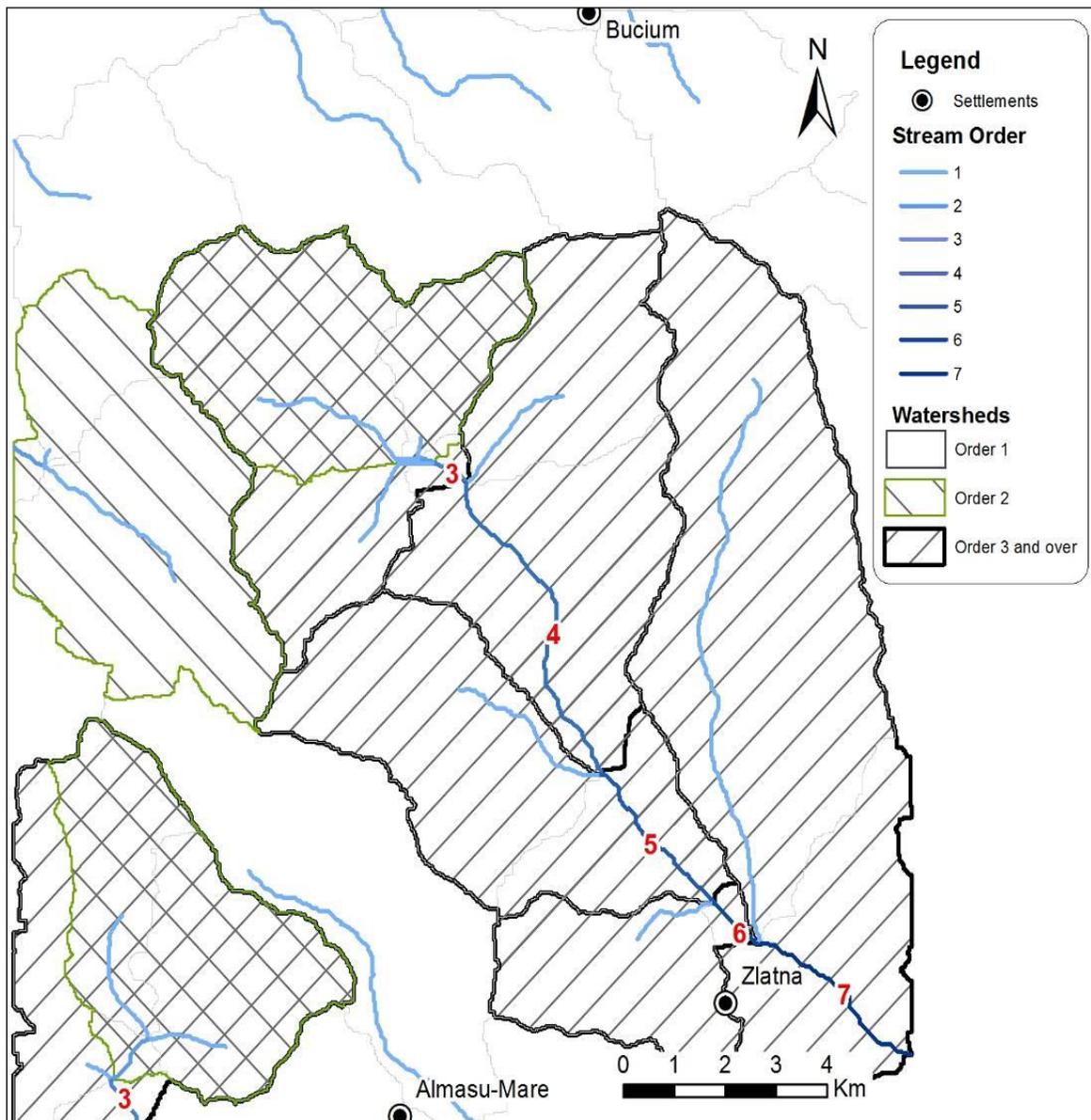
Fig.4.12 *Watershed orders obtained with the model*

After all the modules were applied the analyst has the spatial extent of the watersheds in a digital format with their stream order stored as an attribute. The next step in modeling Is the choice of the catchment where the discharge hydrograph is desired and the application of the model in that catchment.

Along with the delineation of the watersheds where the hydrograph model will be applied, the ArcHydro tools also assign an attribute called NextDownId representing the next downstream catchment. This attribute creates a structure similar to an ordered graph and it can be used to calculate hydrographs downstream based on the routing of the hydrographs already calculated upstream. An example of the NextdownId connections can be seen in the Fig. 4.13.
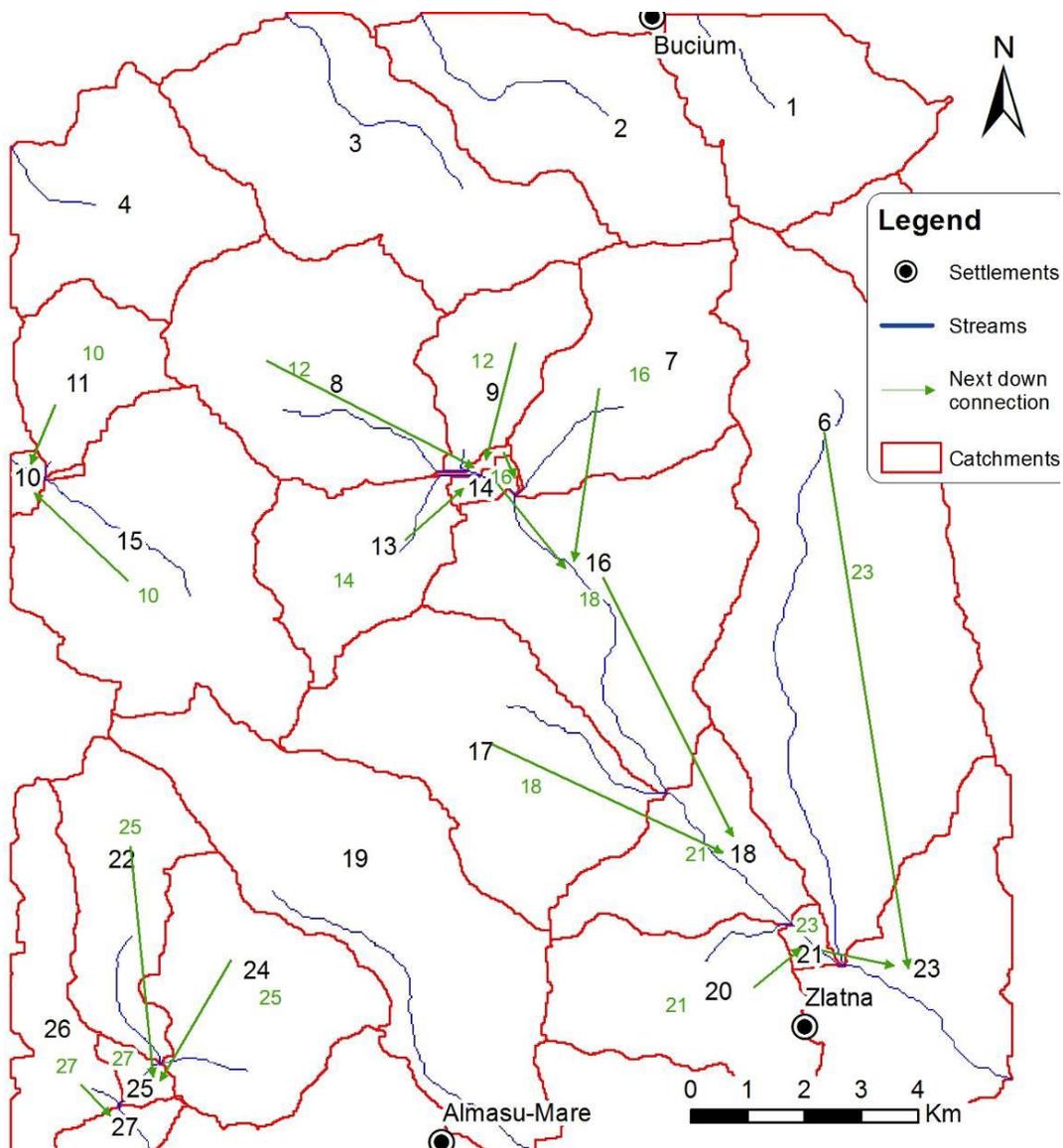
Fig. 4.13 Connections between watersheds using the NextDownId

In the present case, if the flood hydrograph for the ampoi river in the Zlatna city is needed, the best option would be the sixth order catchment, because the seventh order catchment receives a tributary from outside the Zlatna city and the fifth order catchment is at a longer distance upstream.

The catchment where the discharge measurement station can also be determined by spatial selection functions. This catchment will be used in the validation section for every basin.

5. GIS algorithm for modeling surface runoff in small catchments

This chapter aims to strengthen a methodology based on the use of Geographic Information Systems on flood modeling.

Emphasis will be on:

• estimation of the amount of water available for runoff according to prior knowledge of the soil moisture conditions;

• determination of runoff coefficients for various rainfalls;

• integration of runoff on the slopes and hydrograph flow determination in different sections of the basin;

5.1 Conceptual model

The research team from the Faculty of Geography made several studies on flood modeling using GIS and model development for this purpose (Domniţa, Crăciun, Haidu, 2009; Domniţa, Crăciun, Haidu, Magyari-Sáska 2010) in the small basins of the Apuseni Mountains. The basis for determining the runoff hydrograph algorithm was implemented in GIS and ultimately resulted in a model built with ArcGIS Model Builder and in a generated Python script based on this model.

The conceptual model developed is based on ideas presented in a recent study with the research team (Domniţa et al. 2009). The concept is based on time-area type models presented in section 1.5 The basic idea is to model and combine results using known and tested methods in hydrological applications that address different needs for the time-area method. The model contains four components presented in the form of GIS modules that calculate the following characteristics: the water depth available for runoff, a runoff coefficient, the travel time from each cell in the raster representation to the outlet and a module for calculating the discharge (Domniţa et al. 2010).

The travel time from each cell to the outlet is calculated according to the runoff speed in each cell on the flow path. Then, the travel time raster is divided in areas according to equal travel time intervals called isochrones. The areas for each of these isochrones can be plotted on a time-

area diagram to offer information on the areas that contribute to discharge in each time interval. The discharge in each cell is calculated and then routed through a linear routing procedure to the outlet, generating information related to the discharge variation through time. Results from this variation are integrated in a runoff hydrograph, which is the final result of the model.

The model was created using Python scripts based on the arcpy library included with ArcGIS presented in section 2.2 loosely coupled to SAGA GIS for calculation of the runoff speed and MATLAB for interpolation and display of the results. The main components of the model are also created as Python scripts for easier usage. Basic parameters of the ArcGIS component of the model are the following thematic layers (Fig. 5.1): DEM (raster), soils (polygon), land use (polygon), rainfall (raster) for antecedent rainfall for five days (raster), runoff speed (raster). Runoff velocity is calculated using SAGA GIS using the Curve Number index from one of the modules implemented in ArcGIS and Manning's roughness coefficient n obtained from the interpretation of the land use layer.

The result of the ArcGIS component of the model is a table that represents discharge corresponding to each isochrone. The table generated is sent for processing to a MATLAB script. The MATLAB script considers the duration of the rain as a parameter and makes the calculations necessary for accumulated flow at the outlet of the basin in each time interval. This variation is then interpolated and displayed as the final hydrograph. Water routing is therefore linear and performed depending on the speed of flow through each cell on the surface of the basin.
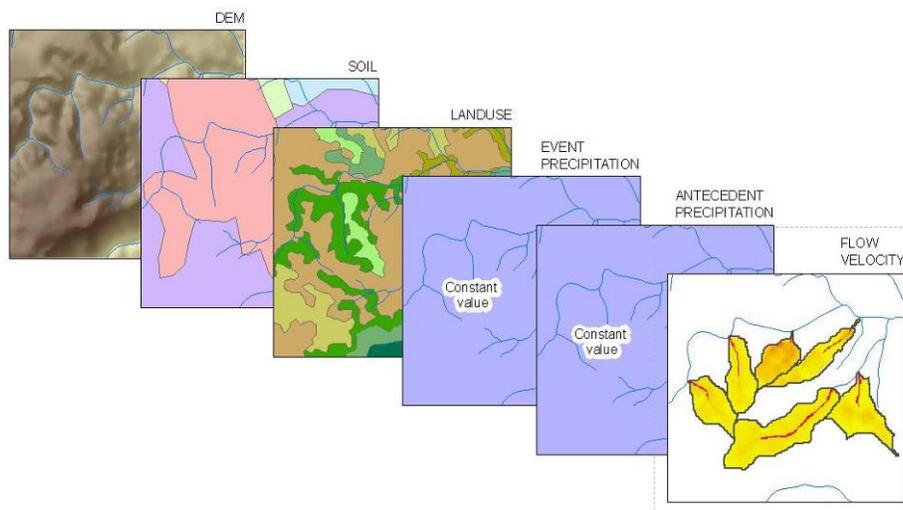
Fig. 5.1 Layers used as parameters in runoff model

Figure 5.2 presents the coupling between ArcGIS and the other components used in the application of the model.
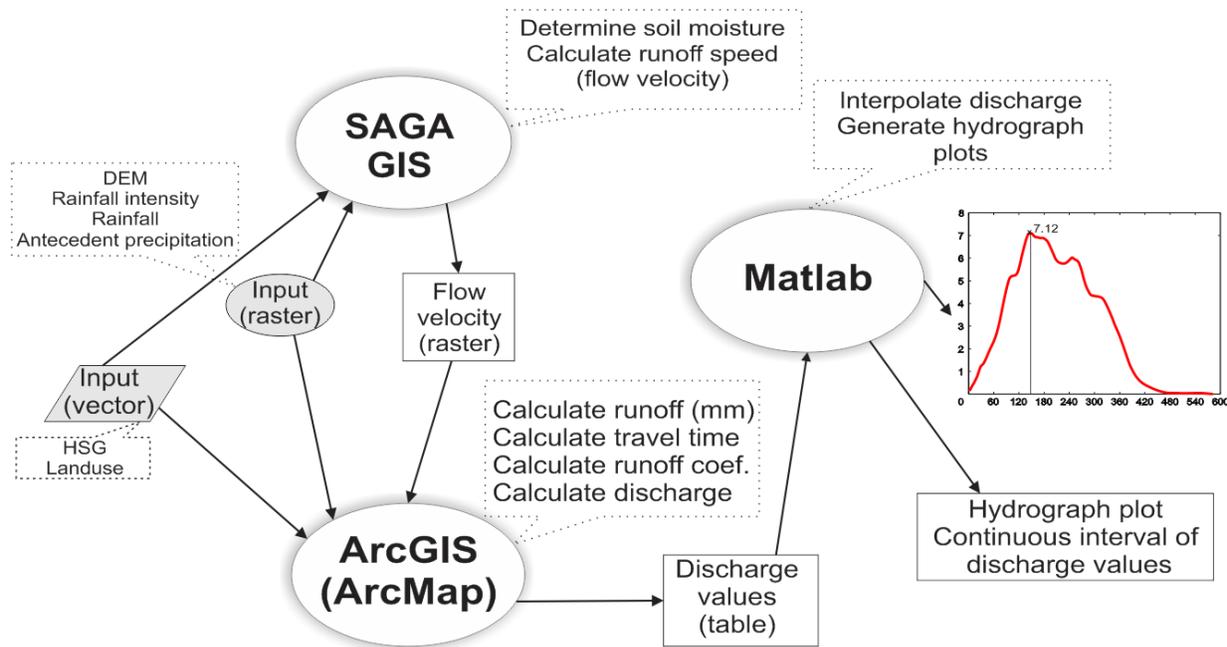
Fig 5.2 Coupling between software products used in model development

The ArcGIS component of the model was built using a modular approach and the outputs of each module become inputs for the next module in the workflow. This means that a module can be used independently if it is needed for another purpose. A diagram of the model implemented in ArcGIS can be seen in Fig. 5.3 and the model components can be identified in this figure.

The algorithms presented correspond to the following workflow:

- Estimation of available runoff depth from rainfall according to the antecedent soil moisture and soil characteristics
- Determination of spatially distributed runoff coefficients for the rainfall, later used in discharge determination
- Runoff integration and discharge hydrograph determination in different sections of the basin
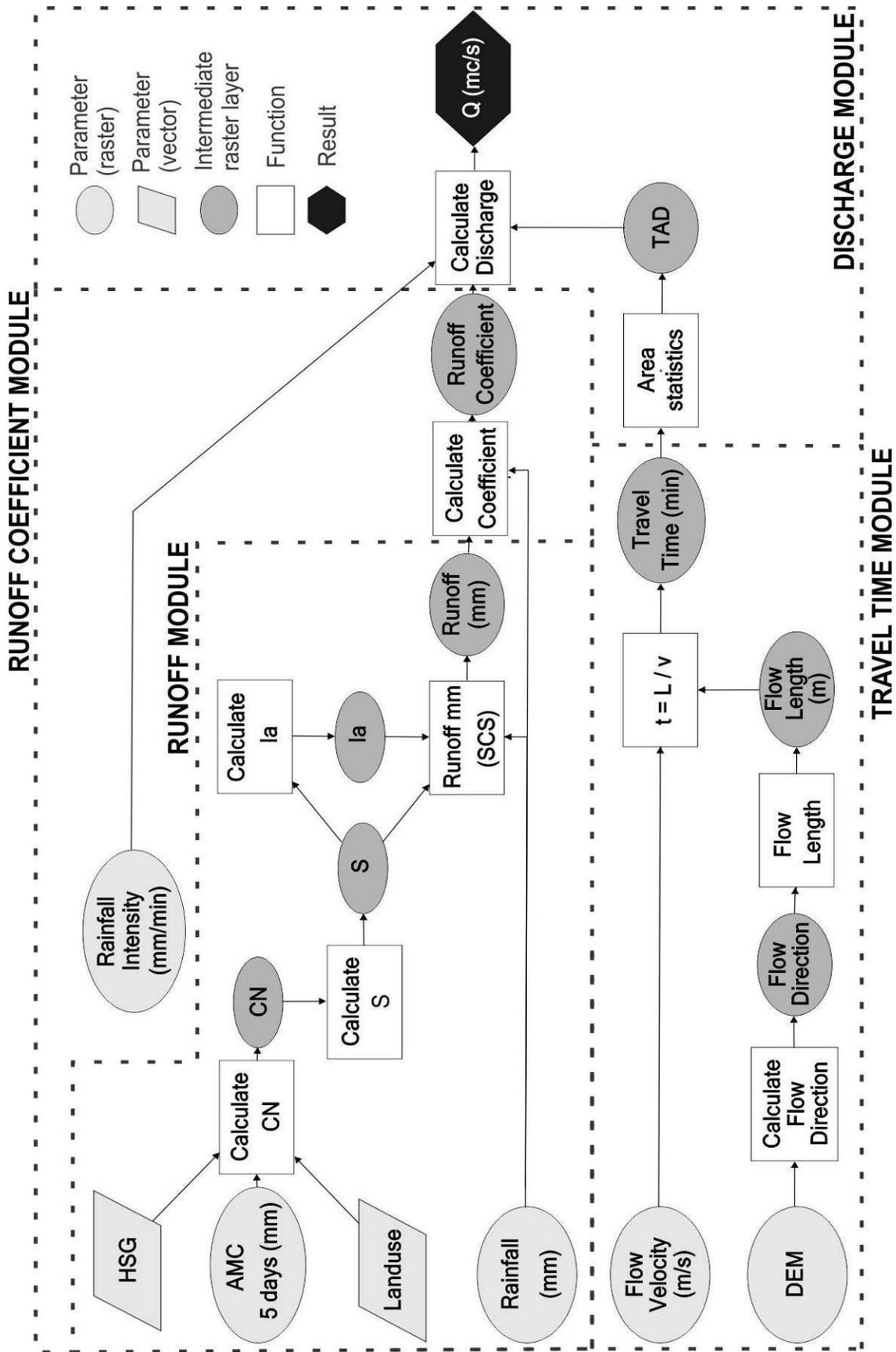
Fig. 5.3 Discharge calculation model diagram

179

5.2 Implementation of the conceptual model using Python and ArcGIS

Determination of the runoff depth from available data layers using the SCS Curve Number Method

The first component, used to calculate the runoff depth in mm (Q) from rainfall distributed through the basin (P, mm), is based on the SCS-CN (Curve Number) method presented in section 1.5. The Curve Number (CN) index is an empirical adimensional parameter used in surface hydrology to estimate the runoff. This module uses the polygon layers (Land Use, Hydrologic Soil Group - HSG) and raster layers (Digital Elevation Model - DEM, rainfall, rainfall for the previous. 5 days) as input parameters.

The Soil Conservation Service-curve number (SCS-CN) method (see section 1.5) is one of the most popular methods for computing the volume of surface runoff for a given rainfall event from small agricultural watersheds. The method is widely used everywhere and included in a large number of hydrological models and continuous modeling schemes.

The method is simple, easy to understand, and useful for ungauged watersheds. The method accounts for major runoff producing watershed characteristics: soil type, land use/treatment, surface condition and antecedent moisture condition (Ponce and Hawkins, 1996; Mishra and Singh, 2003; Mishra et al., 2004 2005).

Ponce and Hawkins (1996) published a critical examination of this method in which they clarified its conceptual and empirical basis, delineated its capabilities, limitations and uses. They concluded that the hydrologic methods developed by the Natural Resources Conservation Service (NRCS, formerly the Soil Conservation Service (SCS)) were originally developed as agency procedures and did not undergo journal review procedures (Ponce and Hawkins 1996). Therefore, the CN often considered far too reliable and the method is not as accurate as a physically based runoff calculation method.

Even if the method is based on empirical principles, the small number of parameters and the ease in obtaining these parameters makes this method a very good choice for the purpose of this study. The accuracy of the SCS-CN method has been criticized many times (Fennessey, 2001; Ponce and Hawkins 1996) but the little data available and the difficulty of taking field measurements calls for a simple method even if the accuracy of the results is not the best.

In this study, the SCS curve number method was used for the design storm in order to determine the runoff depth for a certain rainfall and then the runoff coefficient.

The model has the same limitations as the SCS Curve Number method, so it can only be used at positive temperatures when surface runoff is not from snowmelt, ice, sleet, or rain on frozen ground.

The model gives intermediate results on the CN index, initial loss (Ia) and maximum potential water retention (S). The SCS method is applied to raster cell level, according to equation (USDA-SCS 1985):

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \qquad (5.1)$$

where:

- Q – runoff depth (mm);
- P – rainfall (mm);
- I_a – initial abstraction (evapotranspiration, interception, other losses);
- S – maximum potential water retention;

The potential maximum retention is based on the CN, which is obtained from rainfall amount for the previous five days (Antecedent Moisture Condition - AMC), hydrologic soil group (HSG) and land use.

Equations for the initial losses (Ia) and the maximum potential losses (S) are:

$$S = \frac{25.400}{CN} - 254 \; ; \qquad I_a = 0,2 \cdot S \qquad (5.2)$$

The CN index is calculated from the tables available in the literature (presented in section 1.5) and its calculation uses the land use, hydrologic soil group and antecedent moisture conditions layers as parameters.

Land use values are automatically extracted from the Corine Land Cover database (CLC2006) as presented in section 3.4..

The soil data was obtained from the 1:200000 soil maps as presented in section 3.3

The combinations of land use and hydrologic soil groups and corresponding Curve Numbers can be seen in annex 1. By using the land use and soil layers, a raster map of the CN values for AMC II is obtained.

CN values for dry soil (AMC I) or saturated soil (AMC III) can be calculated using the following formulas (5.3,5.4):

$$CN_I = \frac{(-75 \cdot CN_{II})}{(-175 + CN_{II})}, \text{ AMC I (dry soil)} \tag{5.3}$$

$$CN_{III} = \frac{(175 \cdot CN_{II})}{(75 + CN_{II})}, \text{ AMC III (saturated soil)} \tag{5.4}$$

Another way to obtain the CN values for different moisture conditions is using the table (Table 5.1) provided by the Soil Conservation Service (USDA-SCS 1985*)*

Table 5.1 - Conversion table for Curve Numbers (CN) from Antecedent Moisture Condition Class II to AMC Class I or Class III

AMC II	AMC I	AMC III		AMC II	AMC I	AMC III
0	0	0		60	40	78
5	2	13		62	42	79
10	4	22		64	44	81
15	6	30		66	46	82
20	9	37		68	48	84
25	12	43		70	51	85
30	15	50		72	53	86
32	16	52		74	55	88
34	18	54		76	58	89
36	19	56		78	60	90
38	21	58		80	63	91
40	22	60		82	66	92
42	24	62		84	68	93
44	25	64		86	72	94
46	27	66		88	75	95
48	29	68		90	78	96
50	31	70		92	81	97
52	32			94	85	98
54	34	73		96	89	99
56	36	75		98	94	99
58	38	76		100	100	100

The Antecedent Rainfall for 5 days parameter contains a raster with the cumulative

precipitation for the previous five days. This raster is used to determine the AMC for each cell in the raster representation of the basin. If the AMC is not II, the script will automatically calculate the new CN value for the corresponding AMC.

Once the CN is obtained by this method, the script applies the SCS-CN runoff formulas (1, 2) and returns the raster with the runoff depth (mm) in each cell of the basin. This raster is then used to calculate the runoff coefficient for each cell in a raster.

Runoff coefficient

The dimensionless runoff coefficient α is the proportion of precipitation that contributes to surface runoff. This coefficient takes into account initial losses (for example, water stored in depressions), the continuous losses (eg infiltration) and thus takes into account hydrodynamic effects encountered while surface water flows overland. (Butler 2000)

The coefficient is a value that can vary with time and this method only calculates an overall runoff coefficient for the whole rainfall duration. The coefficient may be lower at the start of rainfall and higher after the soil becomes saturated, but it is calculated as a mean runoff coefficient in this case. A method which includes periodic calculation of the runoff coefficient at a specified time interval will be developed in the future in order to improve the accuracy of the model.

In our case, the runoff coefficient (α) was calculated as the ratio of runoff depth to total rainfall (5.5). Calculated flow depth is used to determine the runoff coefficient raster according to the following formula:

$$\alpha = Q / P \qquad\qquad (5.5)$$

The method used in this calculation is similar to a previous algorithm for obtaining the runoff coefficient depending on soil moisture based on GIS functions (Crăciun et al 2009).

The script can be called directly from a console or run from ArcGIS with a form where the parameters are input. The form of the command used for running this script is:

```
python Runoff_coef (<Rainfall_Raster>, <Runoff_mm_Raster>,
<Result_Folder>)
```

Parameters represent the following:
-   <Rainfall_Raster> - Raster layer representing the rainfall in mm

- <Runoff_mm_Raster> - Raster with runoff in mm calculated with the previous script
- <Result_Folder> - Folder where the results will be stored

Travel and concentration time

Travel time (t) for the path from each cell in the catchment to the outlet was calculated from the velocity of flow. Velocity is obtained using the tool called "Isochrones - Variable Speed" from SAGA GIS and used in conjunction with DEM-derived data to calculate the travel time through the basin.

Besides the CN, the algorithm uses Manning's n number, calculated according to the table in annex 2, as a parameter.

Using the raster with flow velocity in each cell, a weight corresponding to the travel time through the cell can be calculated as the inverse of the runoff speed through the cell (5.6).

$$Tt = 1 / v \qquad (5.6)$$
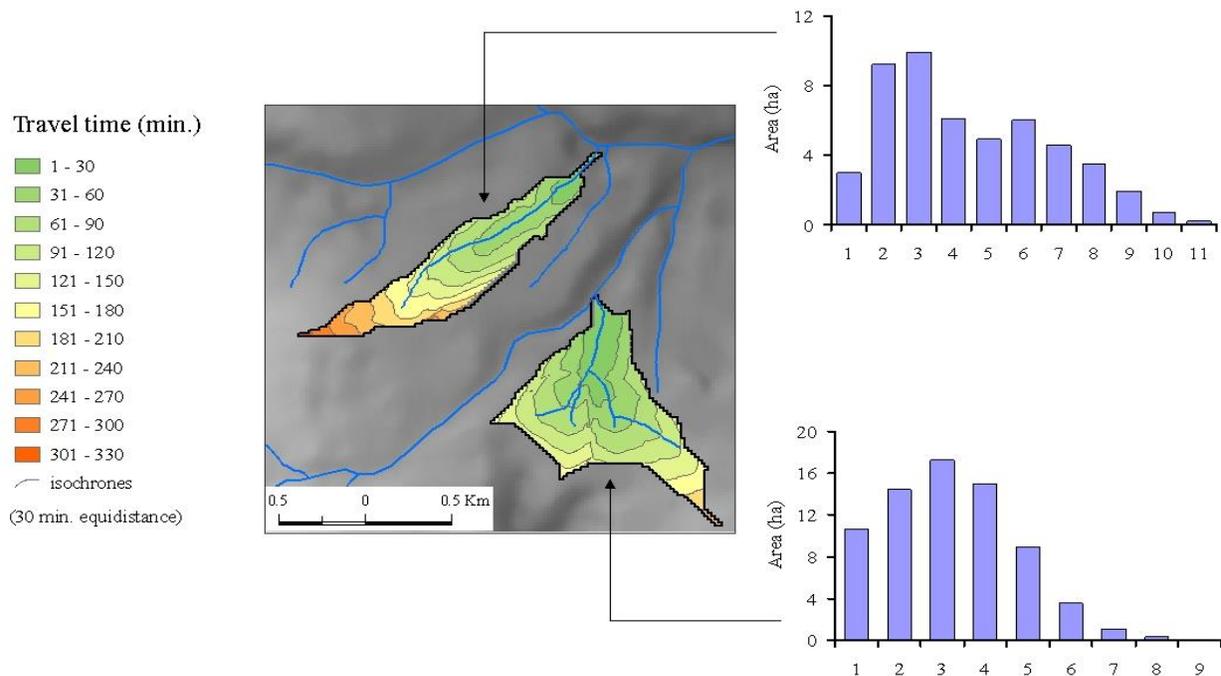
Fig. 5.4 *Spatial distribution of travel time and time-area diagram (30 minutes)*

184

Travel time to the basin outlet is computed adding travel time for each cell on the flow path. Depending on the direction of the flow, the distance of water flow through a cell may vary. To calculate a travel time of the basin as accurately as possible I used the Weighted flow length function that calculates the length of flow depending on the direction of flow and a weight raster defining the impedance or resistance to move through each cell, in this case the travel time of each cell.

The raster of calculated travel times is reclassified in equal intervals (isochrones, in this case 1-minute) (Fig. 5.3). This reclassification is made to determine a best estimate of the variation in time for calculated flow. If computational efficiency and fast processing is required, the intervals can be made longer than one minute, at the expense of accuracy in results.

Flow calculation

The component for flow calculation is used to calculate the flow generated by precipitation in each cell of the study basin representation. After determining the drainage area contributing to each isochrone total discharge can be determined for each isochrone by summing the discharge in each cell. Values of dischage generated for each isochrone are stored in a table.

The results of the previous components are used in this discharge calculation module. The parameters used are the same parameters used in the other modules: Digital Elevation Model (raster), soils (polygon), land use (polygon), Rainfall (raster) Antecedent Precipitation for 5 Days (raster). The discharge is calculated by applying the rational formula at raster cell level to determine maximum flow generated in the cell. The method was adapted for Romania and the scientific literature in Romania includes a number of applications of this method. One of the papers where this method is presented is the one elaborated by Diaconu C et al (1994):

$$Q_{max} = 0,167* S * i *\alpha \qquad\qquad (5.7)$$

Where (parameters have been adapted for the use of a cell in the raster)

Q_{max} - raster cell maximum discharge (m^3/s)

S - raster cell surface (ha)

i - average rainfall intensity (mm / min)

α - runoff coefficient

By adding flow from all cells within each isochrone cummulated discharge is obtained for each time interval (1 min) and is stored in a table of discrete values of the flow. Hydrograph is displayed at the end is a continuous variation of the flow, not only their discrete values. To calculate the accumulation of runoff during rain, these values should be interpolated and transformed into a continuous range of values, which then will be displayed as a graph.

The addition of maximum discharges from each isochrone corresponds to a linear routing of flow towards the outlet of the catchment. The discharge through each cell is accumulated towards the outlet according to the rainfall duration (5.8)

$$Q_i = \begin{cases} \displaystyle\sum_{j=0}^{i} Q_j, & i < d \\ \displaystyle\sum_{j=i-d}^{i} Q_j, & i > d \end{cases}$$

(5.8)

This procedure is similar to the Unit Hydrograph summation (superposition) procedure because all the calculations made are linear and it does not consider flow attenuation, storage or spatial variation of runoff causing nonlinear conditions. Due to the small size of the catchments and the consequent similar soil and land cover characteristics, linear routing was considered sufficient for the purpose.

Another reason for not using more complex equations for flow routing is the fact that the model is a single-event model, not a continuous one. The discharge values are calculated at specific time intervals without considering the previous conditions in the catchment rather than continuous calculations through time.

The script can be called directly from a console or run from ArcGIS with a form where the parameters are input. The form of the command used for running this script is:

```
    python  Discharge_calc  (<  RunoffCoef  >,  <  RainInt  >,
<Result_Folder>)
```

Parameters represent the following:
  - < RunoffCoef> - Raster layer representing the runoff coefficient from the previous script
  - <RainInt> - Rainfall intensity in mm/min, floating point value
  - <Result_Folder> - Folder where the results will be stored

Hydrograph generation

The final hydrograph is generated in MATLAB software.
MATLAB is a development environment for applications in mathematics and engineering where complex mathematical calculations can be made. The program can perform calculations introduced from the keyboard or run an entire script that contains a list of calculations. Once it is created, named and saved a script can be used as any other function available in MATLAB (Lindfield, Penny 2001).

Tabular data with the flow are imported into MATLAB and used to calculate flow accumulation and route flow to the outlet. To automate the procedure I created scripts that import data from the table created by ArcGIS and perform all calculations necessary to display the hydrograph.

During the rainfall period (given as a parameter in calculdischargevar.m) the discharge generated and the flow accumulation are calculated simultaneously. After the rainfall ends only the existing runoff from the basin is routed until the time of concentration is reached (there is no water left in the basin).

Flow routing is linear and based on the previously calculated travel time. The end result of this script is a list of discrete values representing the total accummulated flow at the outlet of the

187

catchment for each time interval.

Discrete values are then interpolated using a spline interpolation (splinedischarge.m) to determine the continuous variation of discharge over time. The result is then displayed as a plot representing the basin discharge hydrograph.

The main MATLAB script can be called from MATLAB using the following form:

Result = apply_all('<Discharge_table>',<Rain_duration>);
Where the parameters are:
- <Discharge_table> - Discharge table from previous python script
- <Rain_duration> - Rainfall duration in minutes (integer value)

The script returns both the graphical representation of the hydropgraph and the final values for accumulated discharge (stored in the Result variable). If the final values are needed for further operations in Excel, they can be exported using the xlswrite command from MATLAB.

5.3 Conceptual model for discharge calculation in small basins based on the shallow water equations

This conceptual model is based on the possibility of combining libraries from several GIS products using the necessary methods from each one. The model uses some of the functionality implemented in ArcGIS Model Builder for the first models to create data sets necessary for running the runoff simulation module in GRASS GIS (r.sim.water). The GRASS module r.sim.water was briefly presented in the previous section (chapter 8).

To call this module in GRASS GIS the layers for the needed parameters of the module must be obtained. We recall that the parameters are: the excess of precipitation (calculated from the intensity of rain and soil infiltration rate in mm / h), flow gradient vector (parameters dx, dy) and surface roughness given by Manning's n coefficient. The derivation of these parameters can be seen in the model diagram (Fig. 5.4).

Some parameters can be obtained by using some components of the models presented above:

• The CN is obtained in the first conceptual model before determining surface runoff with the SCS method

• Manning's n coefficient is used in the first conceptual model to calculate the flow velocity in SAGA GIS

• Outlets are calculated in the module for the spatial discretization of an area using ArcHydro tools. For their calculation a component of the method presented in section 2.2 can be used

• Soil infiltration rate can be obtained from the soil map used for calculating CN

Other parameters required for the GRASS module can be calculated using GRASS GIS. The most important parameters are the corresponding flow gradient vector (dx and dy) which are calculated using the r.slope.aspect module in the GRASS GIS library.

To apply this conceptual model a user interface is required where the user can choose the necessary parameters. Due to the possibilities offered by the Python language this interface can be created in any GIS product. In this case the interface will be created for ArcGIS because it allows the use of its interface components when setting up a script. Using ArcGIS the number, name, type or limits of parameter values can be imposed in order to run the model correctly without errors.

Conversions between the GRID type of raster data in ArcGIS and a raster data type recognized by GRASS GIS is made directly using the Open Source GDAL library or the r.in.gdal function that exists as a GDAL wrapper in the library of GRASS GIS.

Outlets in shapefile or geodatabase feature class format can be imported in the same way using the OGR library or the library function r.in.ogr from GRASS GIS.

Next, the result of calling the GRASS GIS module obtained as a table with values for flow and water level can be taken and interpolated using the interpolation functions in SciPy. Interpolation result can then be displayed as a hydrograph using the MatPlotLib library (Tosi 2009), and a component that allows the creation of graphical interfaces in Python (e.g. wxPython).

189

Parameter (raster)
Parameter (vector)
Intermediate Raster Layer
Function
Result

Q (mc/s)

= hidrograph

Interpolation

Time Series with discharge in points

Flow Routing in GRASS GIS

Rainfall Intensity (mm/min)

Infiltration (mm/h)

CN

n Manning

Flow Direction

Slope dx dy

Outlets

Calculation of Infiltration Rate

Calculation of CN

Calculation of n Manning

Calculation of Flow Direction

Calculation of slope and derivatives

Determination of Outlets

HSG

AMC 5 days (mm)

Landuse

DEM

190

Fig. 5.5 Conceptual model diagram

Implementing this model would provide results based on a function of flow accumulation and routing calculated using nonlinear equations from the field of hydraulics.The SAGA GIS runoff speed is calculated at maximum flow rate generated for input data and do not vary over time. Unlike linear routing, this method would take into account the amount of water accumulated during time and how this affects water quantity and flow speed, so flow rate calculation would make the flow rate vary not only in space but also in time.

However, maps of flooded areas can not yet be achieved without the availability of accurate data on the profiles of riverbeds and other elements along the river and a high resolution DEM. Using an external program such as HEC-RAS (see chapter 8) can automatically do all the mathematical calculations required to create the flood maps.

6. Application examples

6.1 Application for very small basins in the Râşca Village

Six subwatersheds located in the Râşca village, close to housing and private properties were selected for the study.

The rainfall data used in this example is derived from a storm that occurred in 27 June 2009 in the area. The precipitation recorded at the Vlădeasa meteorological station and considered constant over the study had the value of 54mm. The antecedent moisture condition are in the AMC II category, because the rainfall recorded in the 5 previous days sums 36.2 mm.

The study area is located in the NE part of the Apuseni mountains in the Gilau massif. The altitude in the area ranges from 869 to 1152 m. Most of the land in the area is occupied by agriculture, but there are zones with mixed forests or pastures. The study area contains only soils characterized by a loamy sand texture that can entirely be included in the B hydrological soil group according to the SCS (Soil Conservation Service) classification. The hydrograph is presented for six sub-basins of the Râşca basin, with the area ranging from 0.5 to 1.2 km^2.
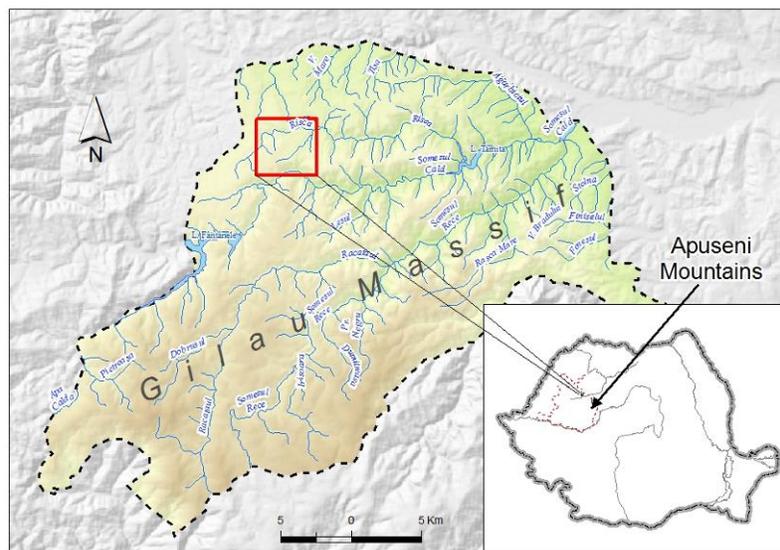
Fig. 6.1 *The location of the studied watershed*

The calculated flow velocity using the SAGA Isochrones – Variable speed algorithm varied from 0.05 m/s to 3.5 m/s. The time of concentration for six basins that were analyzed in the Râşca area vary between 157 and 330 minutes for the rainfall data used. The discharge was calculated

192

for all six sections in studied area, and the results are presented in Fig 6.2.

The travel time (t) from every cell in the watershed to the outlet was calculated according to the flow velocity. The flow velocity (V) for each cell and the DEM from which the flow length grid is obtained are used to calculate the travel time.

The calculated travel time raster is reclassified in 1 minute time intervals (isochrones). This reclassification is carried out for determining the best estimate of time variation for the calculated discharge. The reclassified raster can be presented as a time-area diagram to give a better idea on the area contributing runoff to the outlet at each time step (Fig. 6.2)
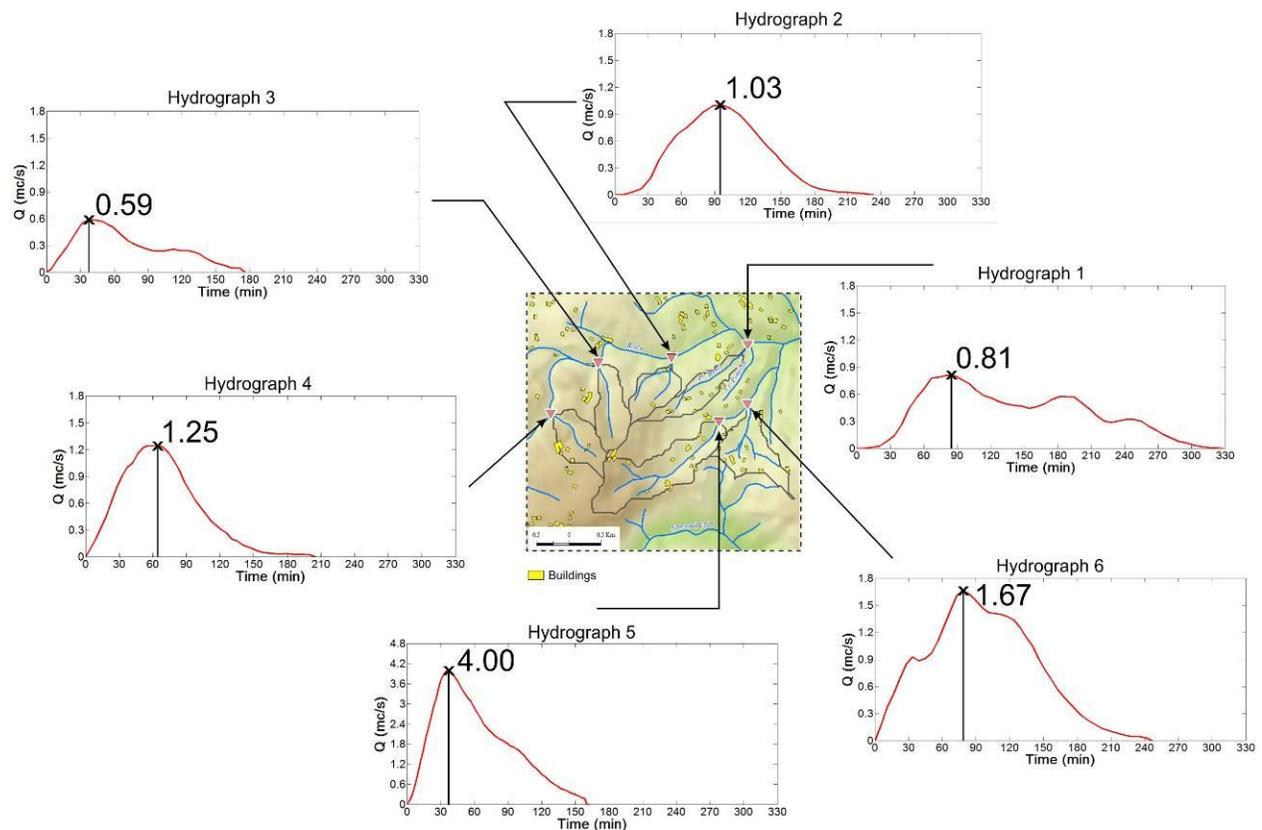
***Fig. 6.2*** *Runoff hydrograph example calculated for six sections of the Râşca basin*

The maximum discharge calculated in this case was about 4 m^3/s in the subwatershed corresponding to hydrograph 5.

The specific maximum discharge (q_{max}) was also calculated based on the maximum values of the estimated discharges and the subwatershed surface for each of the six basins in the area. (Table 6.1).

Table 6.1. Maximum discharge characteristics in the six subwatersheds from the study area

Watershed	F (km^2)	I (%)	Q$_{max}$ (m^3/s)	q$_{max}$ (l/s/km^2)
1	0.51	10.04	0.81	1577.63
2	0.44	18.2	1.03	2369.39
3	0.54	20.03	0.59	1082.23
4	0.52	17.1	1.25	2411.13
5	1.23	16.77	4	3254.23
6	0.72	9.38	1.67	2325.81

where:
F - watershed area
I - slope
Q$_{max}$ - maximum discharge
q$_{max}$ - specific maximum discharge

These results show that the watershed surface is not necessarily the main factor in determining the maximum discharge for a rainfall. The first four basins have a smaller maximum discharge.

The subwatershed corresponding to hydrograph 3 has the smallest maximum discharge because its area is mostly forested and the CN has a smaller value. The subwatershed corresponding to hydrograph 1 has a small maximum discharge because of the long shape and the small slope that causes a low flow speed and a long concentration time.

The highest hydrograph peak and the smallest concentration time can be seen in watershed 4. This is caused mainly by the long drainage length through the subwatershed and the fast flow speed through the drainage channel. Subwatershed 6 also has a significant discharge because of the quick concentration of runoff caused by its shape.

After this study we concluded that the settlements from basins corresponding to hydrograph 5 and 6 are the most vulnerable to flash floods in case of torrential rainfall.

20 torrential rainfalls from year 2008 and 2009 were examined, and the inhabitants from the villages in the study areas confirmed an important discharge, but a quantitative result could not be determined.

A quantitative evaluation of the water level in streams will be created considering the profile morphometry and the cross section of interest which can be measured in the field.

6.2 Application for small basins in the Mărișel Village

The Mărișel village is situated in the southwestern part of Cluj county in Romania, in the Gilău mountains. The basin in study is situated between 660 and 1270 m altitude and includes the

main part of the Mărișel village. The village has a population of 1.670 persons (2002) and the properties are spread through the whole area (fig 6.3 a).

The landuse mainly consists of agricultural land between significant areas of forests and pastures (fig. 6.3 b).
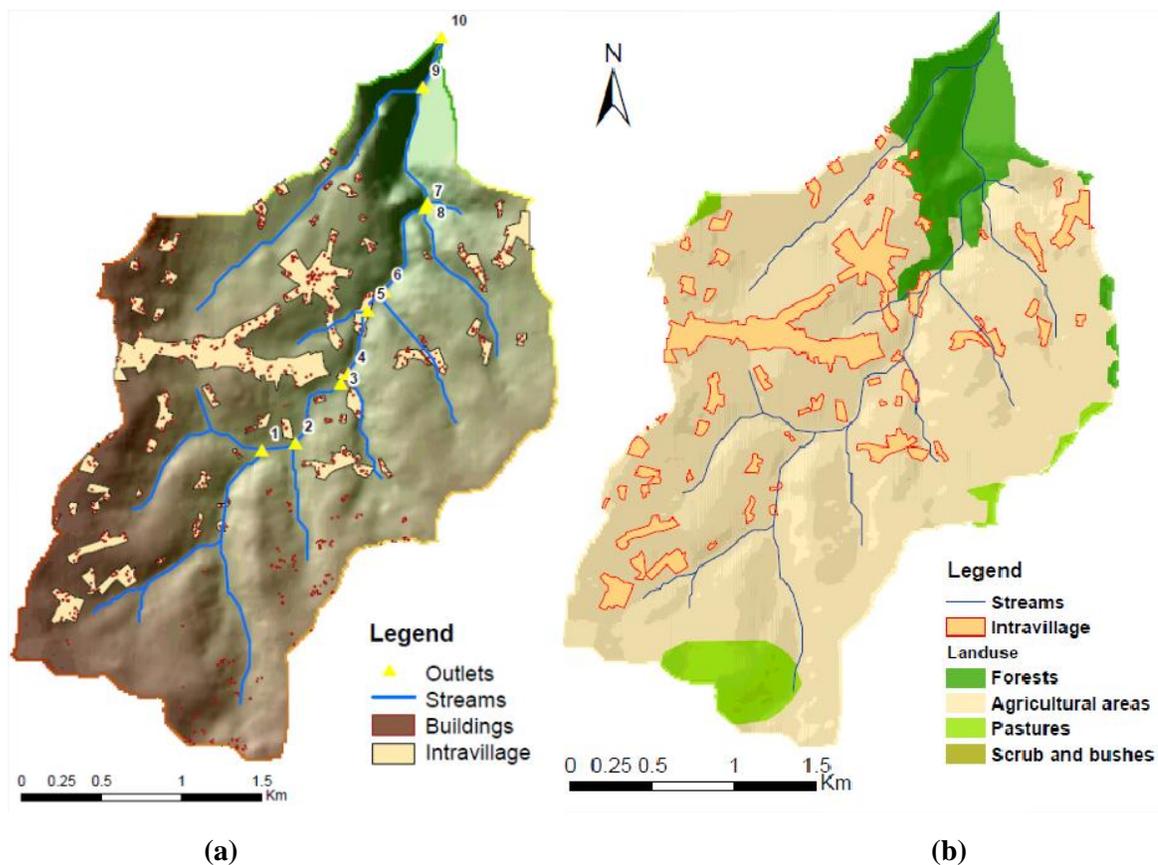
Fig. 6.3 *The Mărișel basin – inhabited areas (a) and landuse (b)*

Ten sub-basins were chosen from the basin near inhabited areas and are presented in fig. 6.3 b. The area of the largest basin is 8.1 km² and the areas of the others range from 1.96 to 6.92 km².

The rainfall data corresponds to a storm in 21.07.2001 with a depth of 55.4 mm. The antecedent rainfall of 46.4 mm corresponds to the AMC 2 moisture condition. The value was considered constant through the whole basin and obtained from the Băișoara weather station. The runoff depth calculated with the runoff depth script presented in section 1.1 has the values shown in fig. 6.4 a. The SCS CN values in the basin varied between 60 and 78. The runoff coefficient varied between 0.03 and 0.26.

(a)   **(b)**
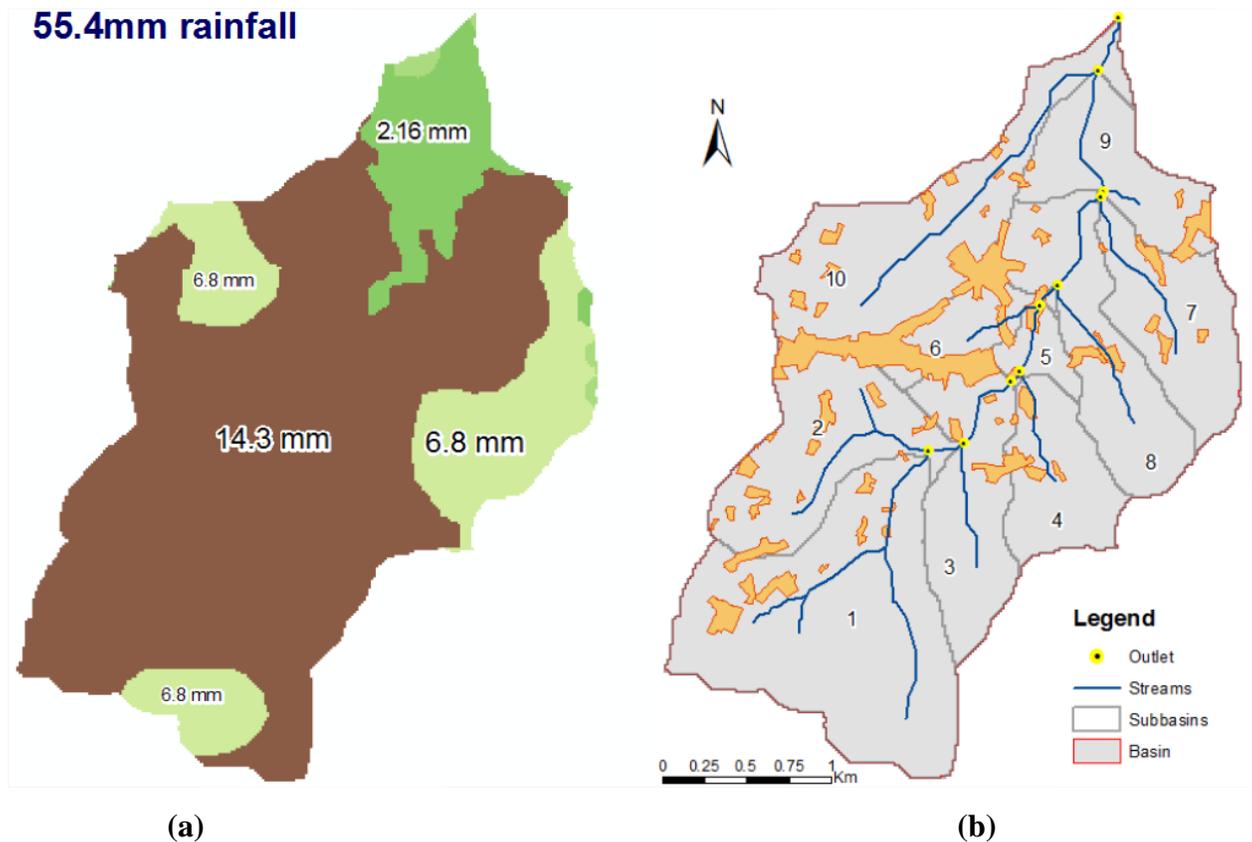
Fig 6.4 Runoff depth (mm)(a) and sub-basins (b)

The time of concentration in the sub-basins varies between 490 and 605 min and can be seen in fig 6.4. The rainfall duration was taken as one hour. Specific data on storm durations is not available at this spatial scale and the closest weather station where this data is available was at Cluj-Napoca in the Transylvanian plateau, at a significantly lower altitude. Therefore, the value for the rainfall intensity when calculating the discharge was 0,92 mm/min.

The discharge obtained for three of the hydrographs is presented in fig. 6.5 along with their location in the basin. The clear difference in discharge between the three hydrographs can be noticed in the figure and corresponds to the location of the sub-basins.

All ten rainfall hydrographs obtained are presented in fig. 6.6. As the figure shows, the maximum discharge obtained was 13.05 m^3/s for the largest basin (basin 10) after 176 min from the start of the rainfall.

The data from the application of this model show good results compared to discharge measurements in basins of similar size and shape in the same region as this basin, although measurements from a station corresponding to any of the basins from this study were not available.

196
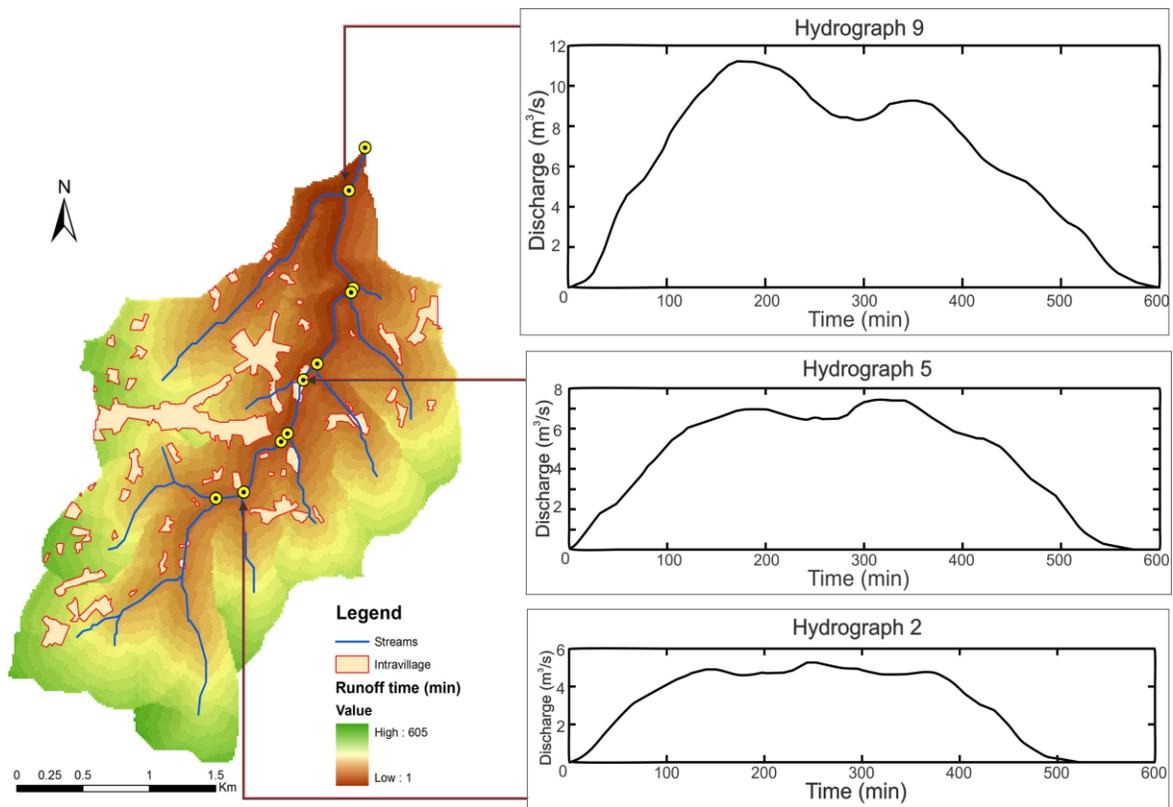
**Fig 6.5** *Hydrograph examples for three of the sub-basins*
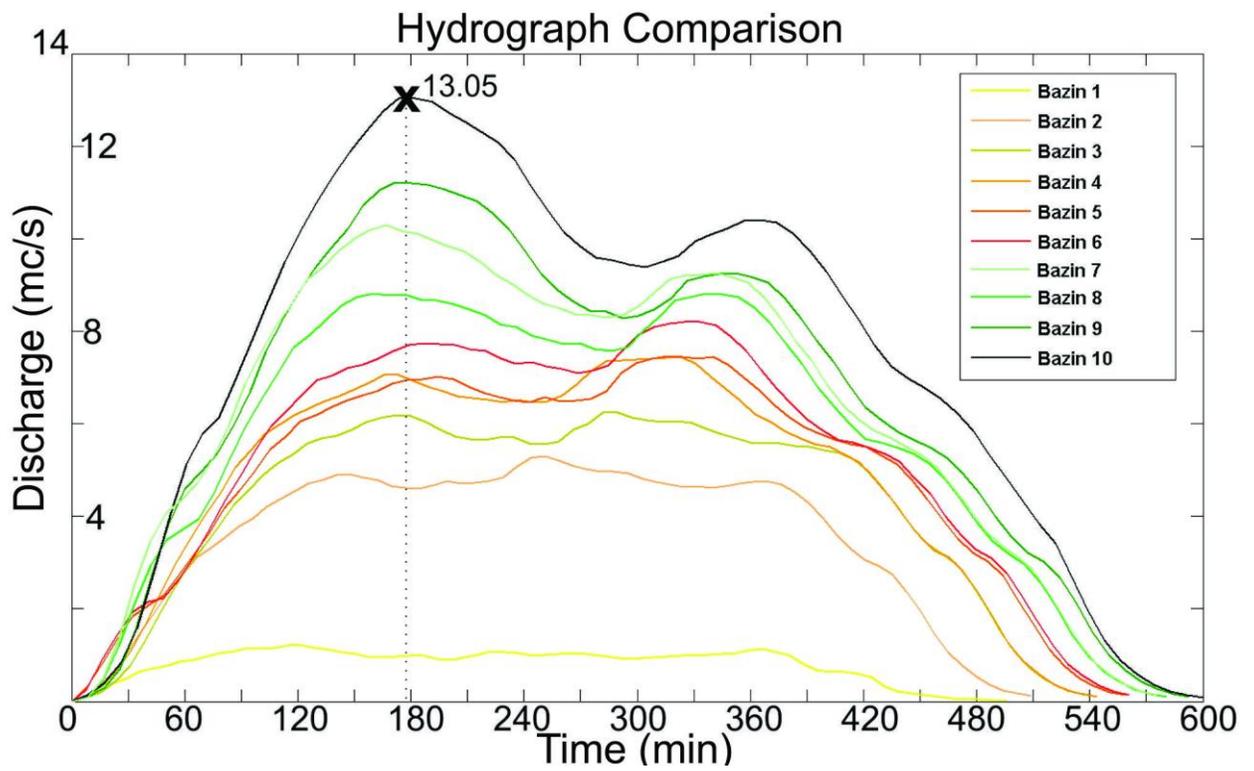
Fig 6.6 *Hydrograph comparison for all the sub-basins*

197

7. Results and validation using measured data

7.1 Validation procedures. General aspects

A validation chapter is necessary in order to establish a procedure to potentially confirm the GIS model of discharge generated by surface runoff. Validating the methodology that constitutes the basis of a research study is a very important phase in a scientific paper.

To gain confidence in the reliability of models for predicting streamflow and yields of its constituents, sensitivity analysis, calibration and validation have been considered essential (Grayson and Bloschl, 2000; White and Chaubey, 2005). In several studies reported before the new millennium, watershed models were calibrated and validated at one site, i.e. the drainage outlet of a watershed (Refsgaard, 2000; Qi and Grunwald, 2005). This approach has been considered problematic for extrapolating watershed response to all other locations within the watershed because of the uniqueness of model parameters for individual watersheds (Beven, 2000). This is the reason for which validation will be dome in a number of catchments in the Apuseni Mountains, Romania.

The validation contributes to creating an image on the accuracy of the results obtained through modeling. It is known that all the hydrological models that make forecasts have certain errors in results. The successive calibrations of a model have minimizing these errors as a purpose. However, it is important that the limitations of a chosen modeling approach are taken into account when doing the calibration and validation. Any model can have different sources of error and these sources should be taken into account both in calibration and validation procedures and in model usage.

Differences between recorded data and simulated model output arise basically from four sources of uncertainty (Abbott, Refsgaard, 1996):

1 Random or systematic errors in the input data, i.e. precipitation, temperature and evapotranspiration etc. used to represent the input conditions in time and space over the catchment.

2 Random or systematic errors in the recorded data, i.e. the river water levels, groundwater heads, discharge data or other data used for comparison with the simulated output.

3 Errors due to non-optimal parameter values.

4 Errors due to an incomplete or biased model structure.

Any differences between simulated and recorded data are due to the combined effect of all four error sources. Only error source 3 can be minimized during the calibration process. The measurement errors (sources 1 and 2) are the limitations that cannot be overcome no matter how good the model is. A model can and should not be designed to give results with a better accuracy than the input data. Calibration should therefore be made to minimize error source 3 until it becomes insignificant compared to the other sources.

In the present study for validation of the GIS model for flash flood simulation the employed procedure was to compare the flash floods hydrographs obtained by modeling to hydrographs obtained by measurements at hydrometric stations. But we must keep in mind that the GIS model estimates only the surface runoffs generated by rainfall, and it does not model river base flow, hypodermic runoff or groundwater flow. The model can thus be applied only to very small basins or to hillslopes and torrents where the surface runoff is very small compared to that resulted from heavy rains.

The validation procedure is necessary because it reveals the imperfections of the model, thus helping to apply the corrections. It also reveals the problems of the input data (spatial-temporal reduced resolutions of hydro-meteorological data, insufficiently detailed cartographic database, up to date/obsolete cartographic database etc.). Concerning the hydrologic modeling of the hydrographic basin, the validation process is difficult because a large number of basins have to be taken into consideration.

7.2. Comparison to existing flash flood models developed in Cluj Napoca

Another model developed in Cluj Napoca in the PhD thesis of A.I Crăciun in 2010 uses the same base concepts that are used in the model I implemented in this work. There are, however, some differences, especially in the approach on infiltration and soil moisture. The following section will present a direct comparison between the results of the scripts presented in this work and the results presented in the thesis of Crăciun (2010).

The application is made on the same basins using the same parameters as inputs. The study areas are the basins of Beliș upstream from Poiana Horea and Someșul Cald upstream from Smida.

Event presentation

In the second half of July 2005 the entire country faced abundant precipitations for several days in a row, and in the interval 11-18th of July 2005 some of the greatest flash floods occurred. As it is mentioned in the paper of Arghius V. (2008), the area of Apuseni Mountains was characterized by a pronounced atmospheric instability that had triggered strong storms. This author explains that the west and the north of the country was hit by a cyclone in the 10th of July, the occlusion of which was produced in the next day and persisted for one or two days, and this had as a consequence in the 12th of July in both studied basins.
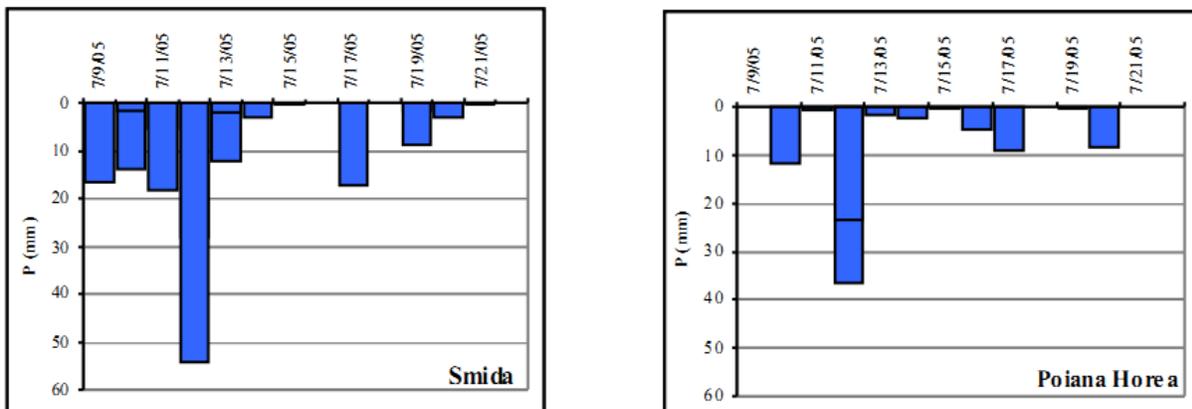
Fig. 8.1 Rainfall in the Beliş and Someşul Cald basins (Crăciun, 2010)

The quantities of precipitation surpassed 50-60mm/24h in the 12th of July, especially at the meteorological stations situated on higher altitudes (Vlădeasa, Băișoara) or at pluviometrical posts Smida and Poiana Horea (fig 8.1). The maximum intensity of rains had even values above 1mm/min. The maps of daily precipitations are presented, created according to the interpolation methodology from the previous chapter and taking into consideration the following meteorological stations: Vlădeasa, Băișoara, Huedin, Zalău, Cluj-Napoca, Dej, Turda, and the pluviometrical stations of Smida and Poiana Horea.
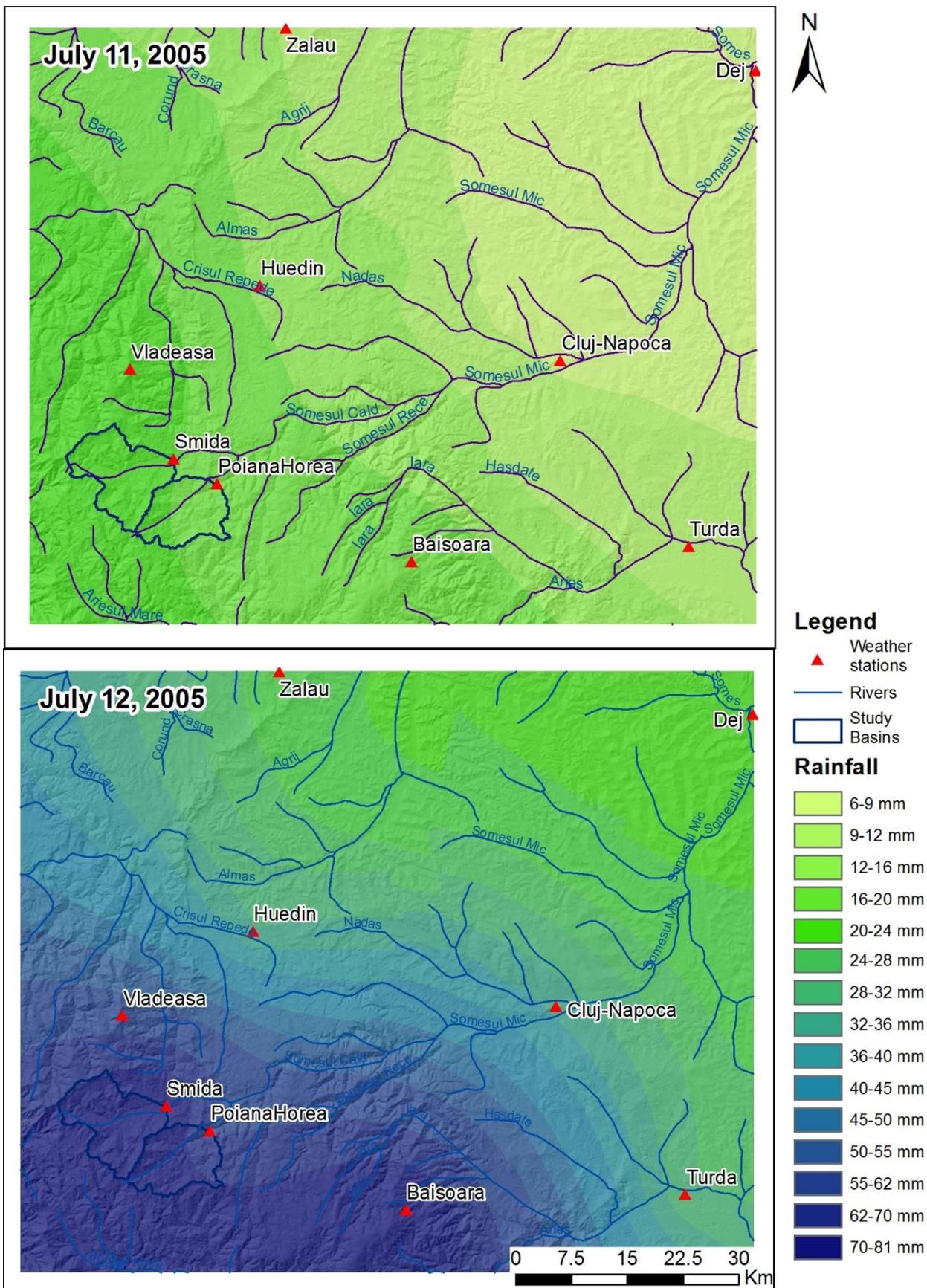
Fig. 8.2 - Rainfall 11-12 Jul 2005

July 13, 2005

July 14, 2005
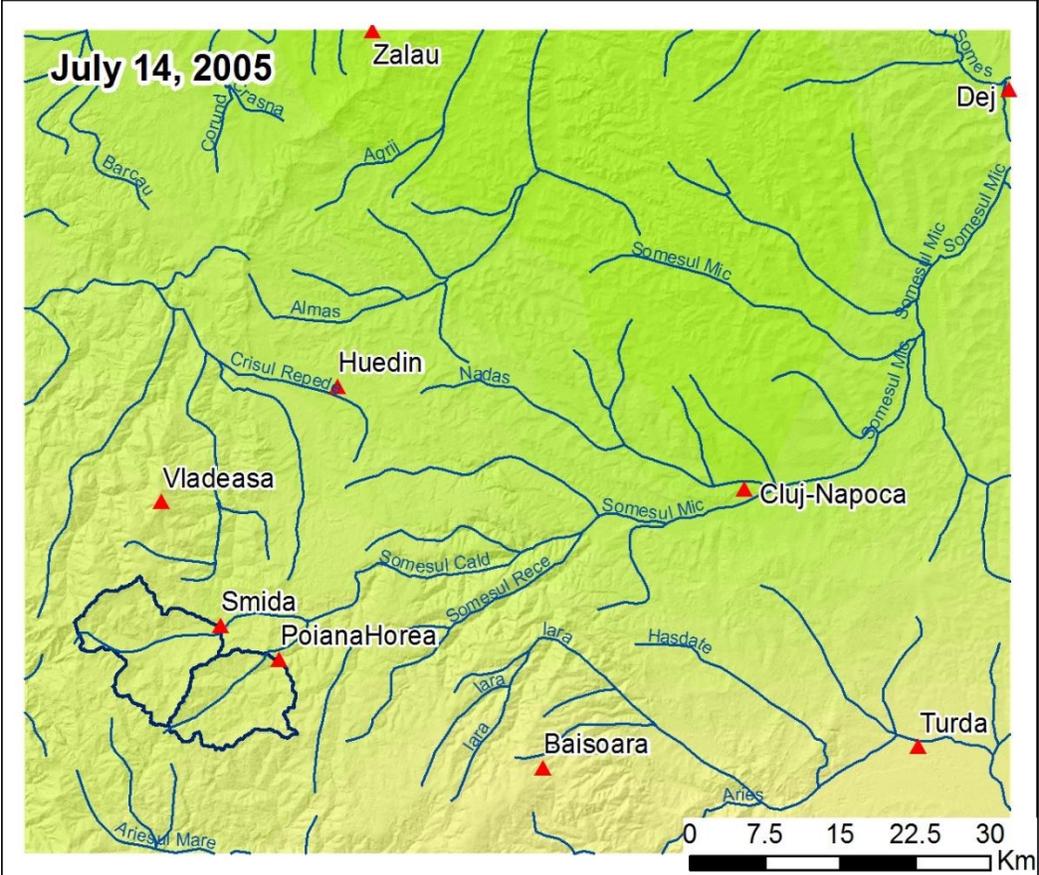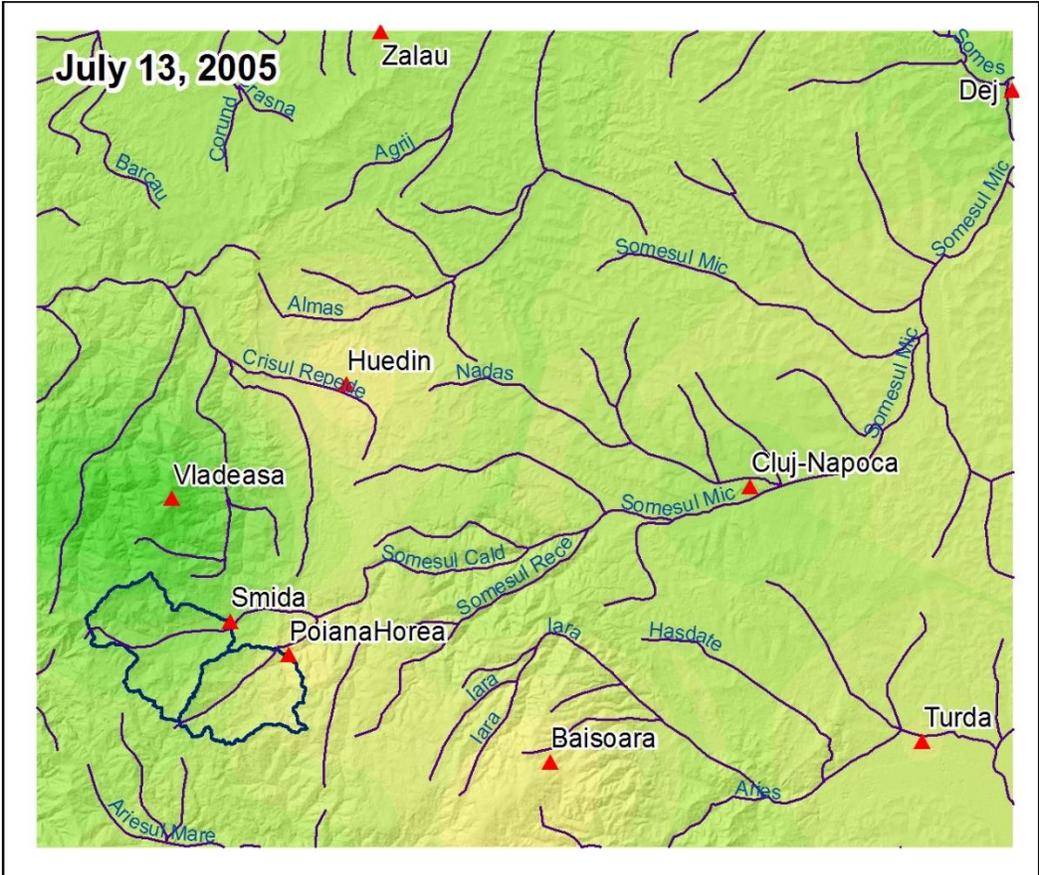
Legend

▲ Weather stations

— Rivers

☐ Study Basins

Rainfall

☐ 0-2 mm
☐ 2-4 mm
☐ 4-6 mm
☐ 6-9 mm
☐ 9-12 mm
☐ 12-16 mm
☐ 16-20 mm
☐ 20-24 mm
☐ 24-28 mm
☐ 28-32 mm

Fig. 8.3 - Rainfall 13-14 Jul 2005

Beliş Catchment

The hydrograph in the case of the Beliş catchment shows a fast increase of the discharge (Fig. 8.4) from 4.36 m³/s in the 12th of July at 6:00 to 16 m³/s at 14:00. The rise is more than 11.5 m³/s in 8 hours, a quite steep rise corresponding to a flash flood in the area. The discharge increase is probably due to some high intensity rainfall, especially in the lower areas of the catchment.

As the temporal resolution of the measurements is not high (only 5 measurements in the day of the maximum discharge) there is not much information on the variation of the discharge in time. Due to this impediment, the hydrograph shape cannot be estimated with sufficient accuracy. More frequent measurements would make this task easier.
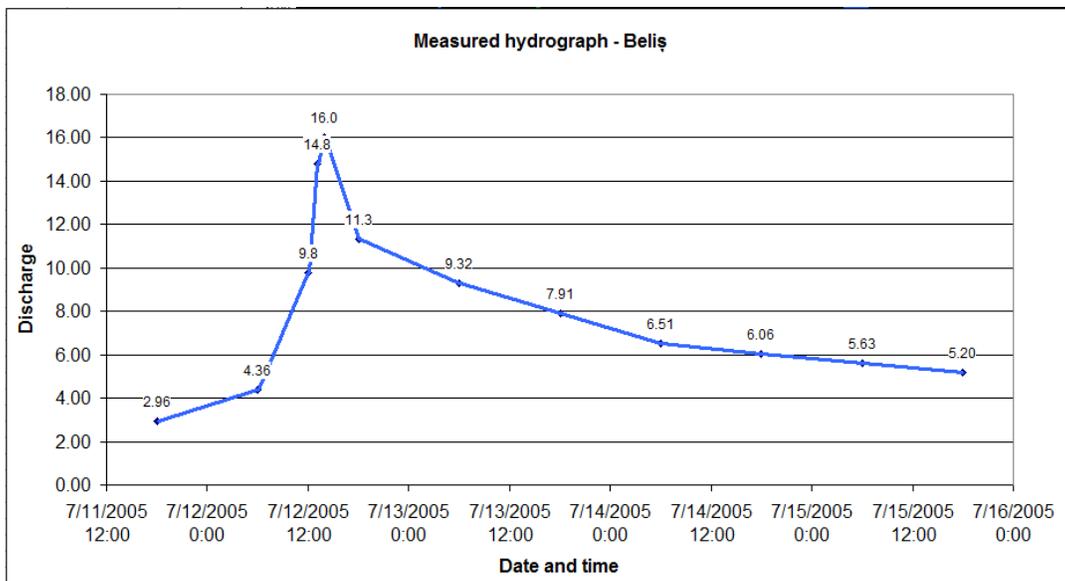
Fig. 8.4 Measured hydrograph for the Beliş catchment (m³/s)
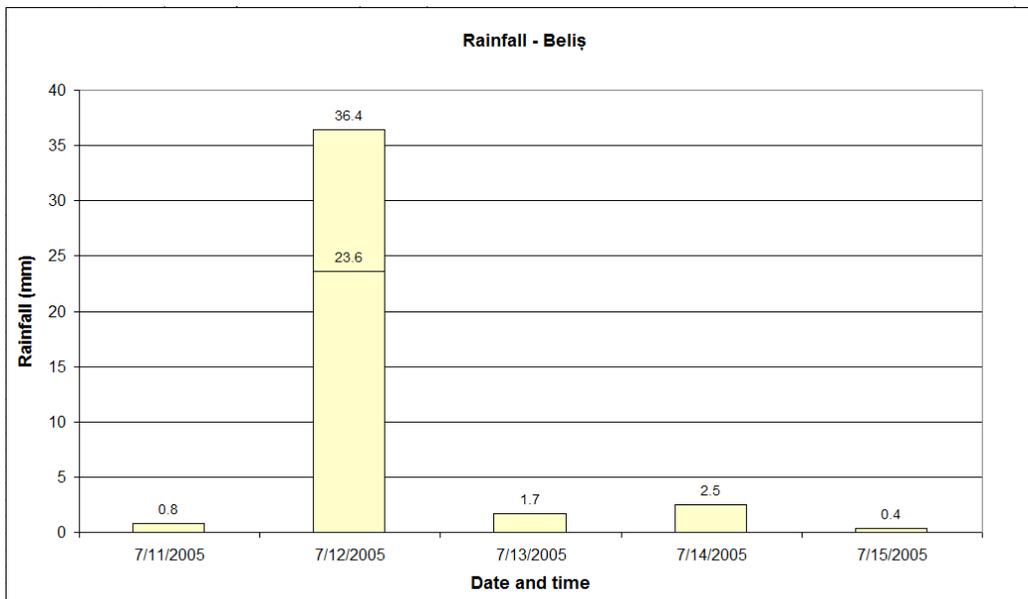
Fig. 8.5 – Rainfall values in the Beliș catchment (m³/s)

In the case of the Someșul Cald river a fast increase of the discharge from 4.73 m³/s at the start of July 11 up to over 30 m³/s in the afternoon (Fig 8.6) was observed. The peak is reached in the evening of 12 July and then the discharge starts decreasing, reaching 6 m³/s in the 16th of July because of less rainfall after the 12th of July.

The rise time was about 36 hours and the decrease time about 96 hours. The exact chape of the hydrograph is not well defined due to the lack of measurements. The points corresponding to the actual measurements are marked on the hydrograph.
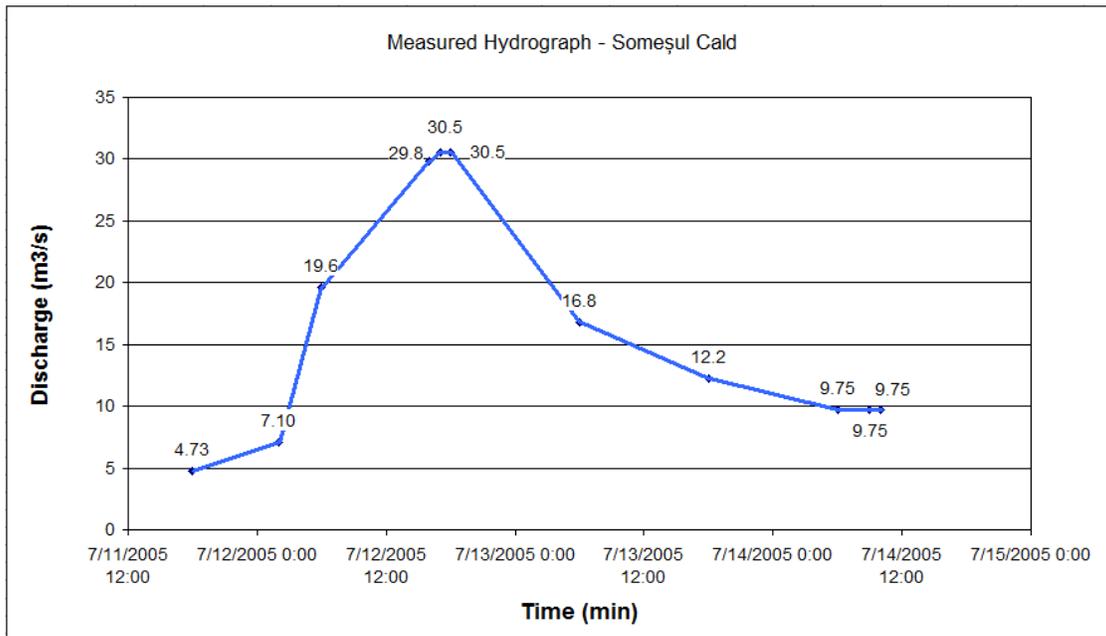
Fig. 8.6 Measured hydrograph for the Someșul Cald catchment

204
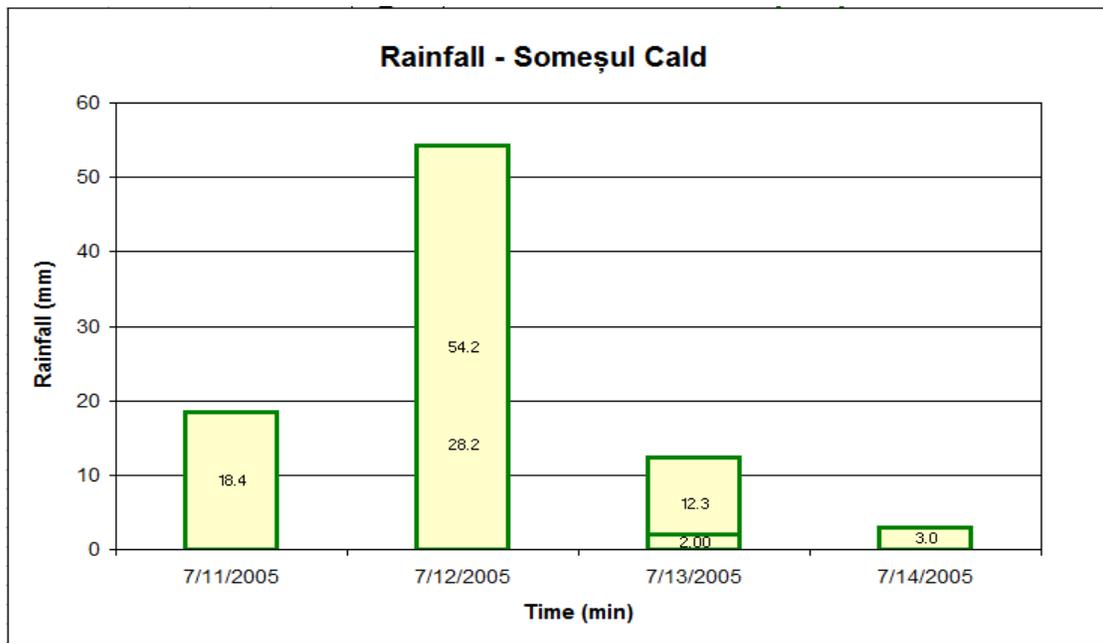
Fig. 8.7 – Rainfall in the Someșul Cald catchment

The travel time was calculated with SAGA GIS for both of the catchments. The first catchment presented is the Beliș catchment, where the calculated time of concentration was 1210 min. The time-area diagram and travel time can be seen in Fig. 8.8
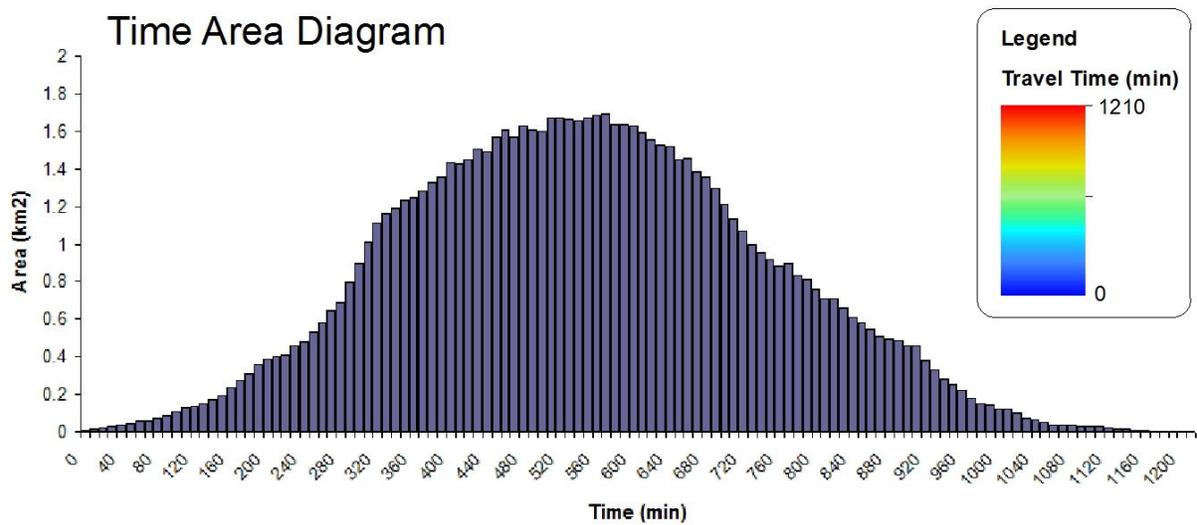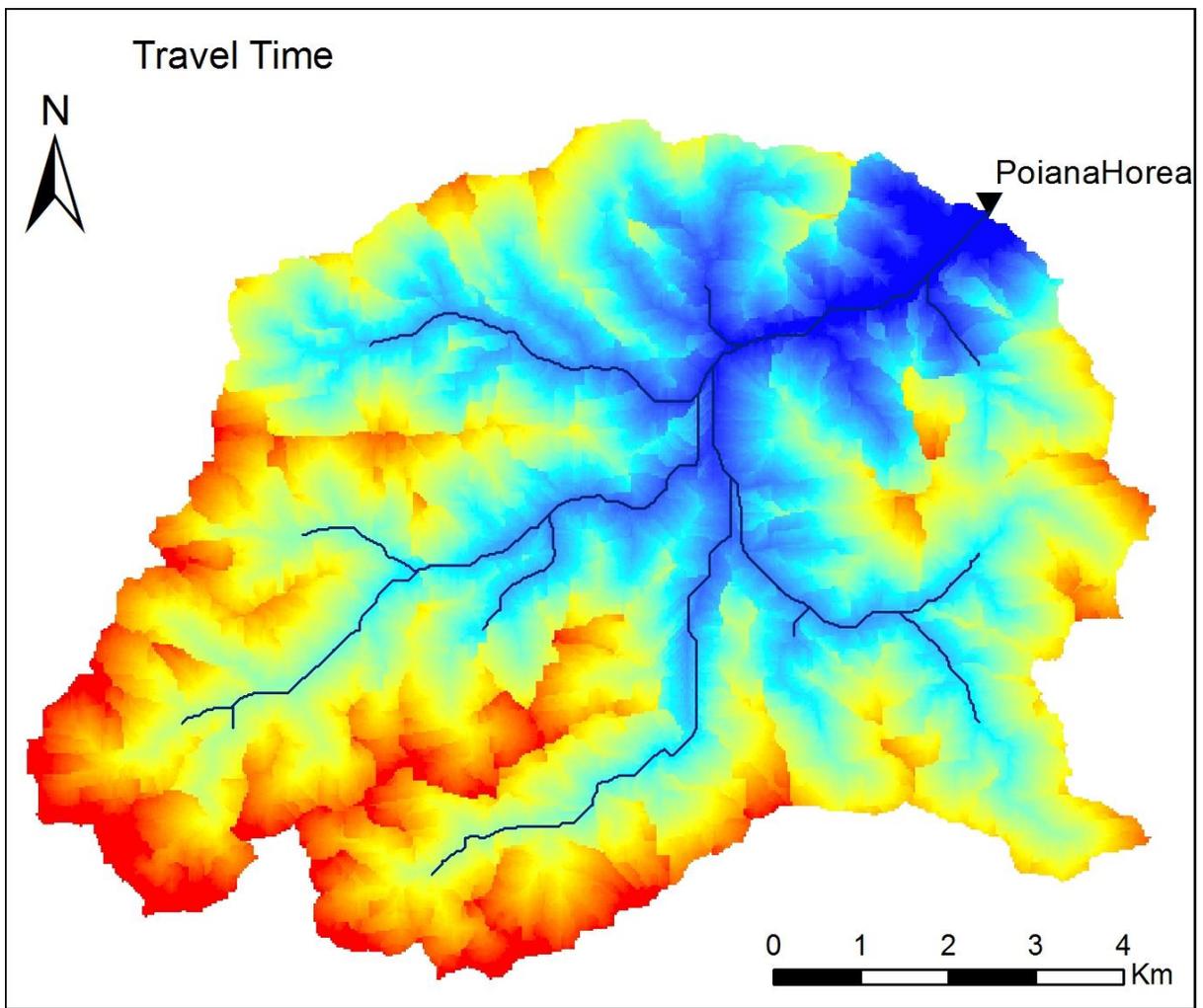
Fig. 8.8 Time Area Diagram and travel time for the Beliş basin

The hydrographs were calculated for each day during the studied periods (11-15 July 2005) and were cumulated using the algorithm that applies the hydrograph superposition technique in

MATLAB. The final discharge values for each rainfall were added up according to the starting times of the rainfalls to obtain the final hydrograph.

The starting point of the first rainfall was considered to be h 18:00 on the 11th of July. The other rainfalls were considered to start at the start of the day, which is not the case most of the time. Therefore, the discharge variation in time can be affected by the starting point of each rainfall. The information on rainfall is very scarce and the exact starting time, duration and intensity are not known. Therefore these parameters had to be chosen according to my opinion. In this case, I considered the rainfall intensity of 0.8 mm/min and the rainfall duration was calculated from the total rainfall and the intensity. This is not necessary the case and may lead to results higher than expected when the calculations are made.

The comparison between the discharge hydrograph obtained from the model application and the measured hydrograph can be seen in Fig. 8.9

Another model created in the Faculty of Gepgraphy in Cluj Napoca was applied on the catchment in the same conditions. The results of this model can be seen in fig. 8.10:

From the comparison it can be seen that the measured hydrograph peak is about 3 m^3/s higher than the modeled hydrograph peak. This can be explained from the fact that the model only accounts for surface runoff. The base discharge and groundwater flow are not taken into account. As it can be seen from the graph, the base flow in the river was about 3 m^3/s so if the model does also take that into account the result will be closer to the measured one.

The slower decrease of the discharge in the case of the measured hydrograph can also be explained by the slower flow processes in the catchment that are not modeled by the model. Groundwater flow and hypodermic flow can cause a longer period with higher discharge through the river. Another fact can also be considered. Due to the interval between the two consecutive measurements (12 hours) the discharge may have dropped faster than it is represented on the graph, as the graph only interpolates the measurements with straight lines and there are not enough measurements for an interpolation of values that would be closer to reality. The discharge hydrograph measured in the field can have a shape that is different from the one seen in the plot if more measurements are made.
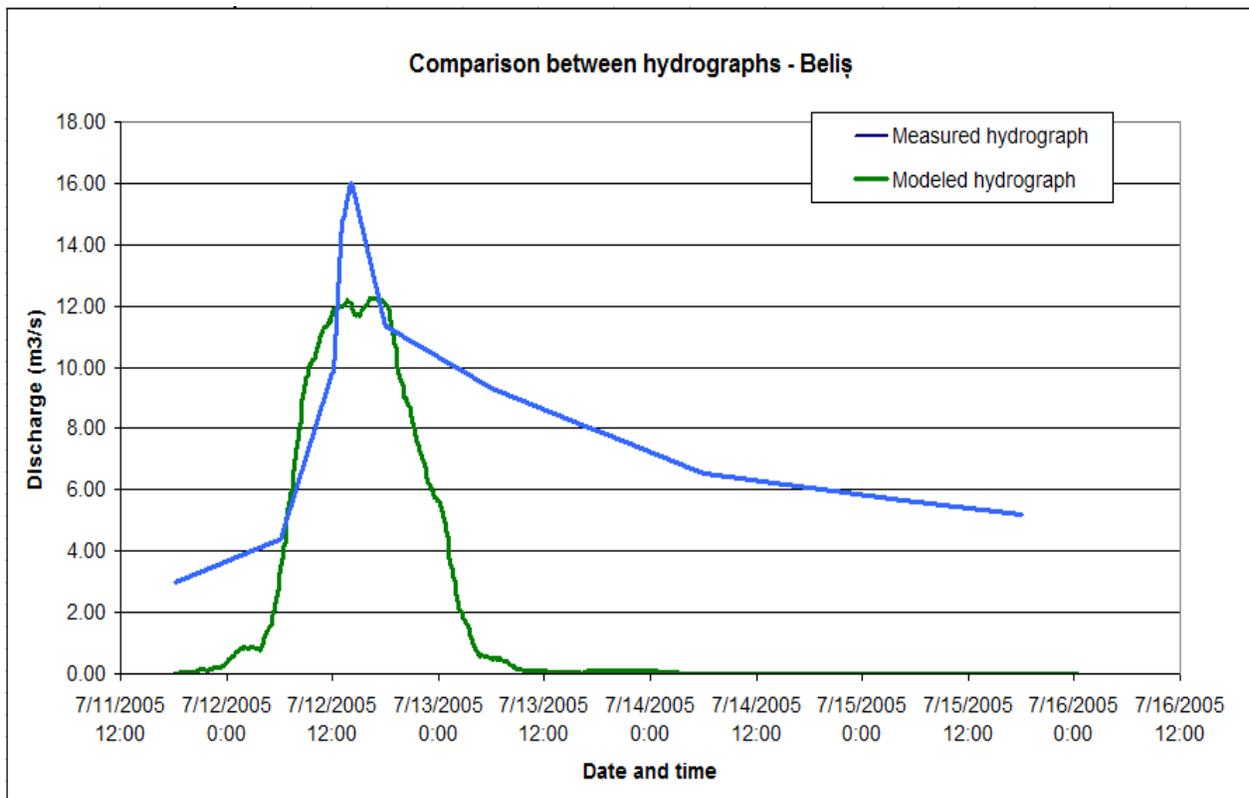
Fig 8.9 Comparison between measured and modeled hydrograph in the Beliș catchment
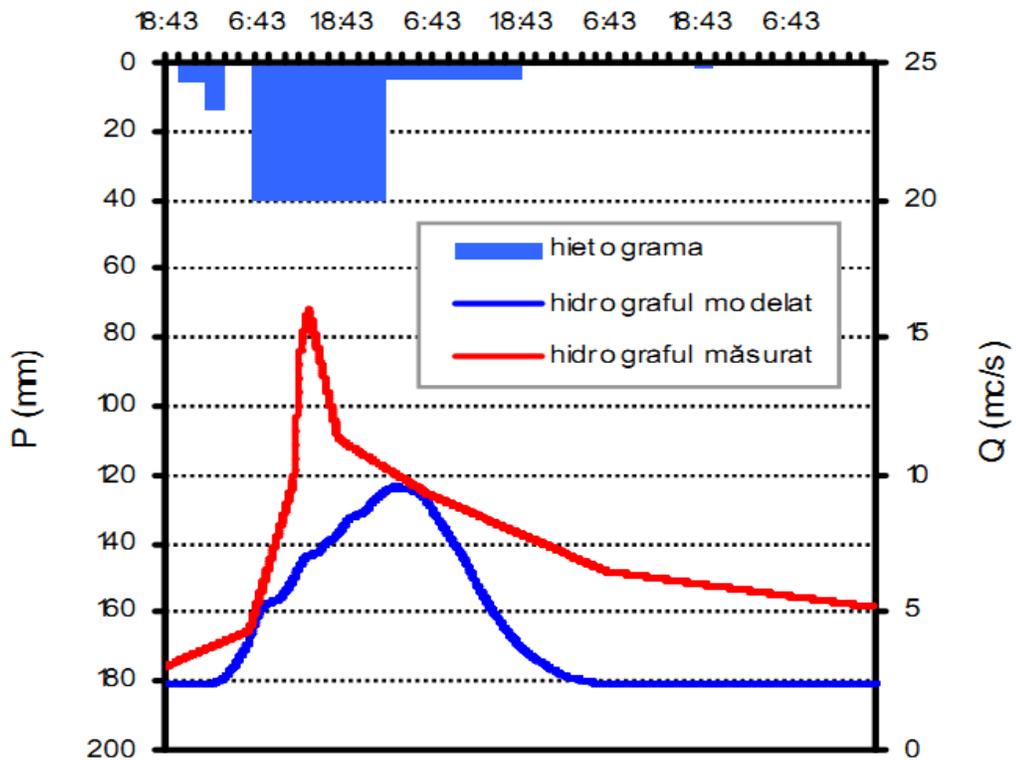
Fig 8.9 Comparison between measured and modeled hydrograph in the Beliș catchment (Crăciun 2010)
The model used by Crăciun (2010) follows some similar concepts (the time-area concept,

the SCS Curve Number method and flow speed calculation using SAGA GIS) but differs in some areas. The infiltration is taken into account separately in the Initial Abstraction parameter from the SCS method. The author considers that the interception from the vegetation can also contribute to the discharge later in the process of runoff.

Due to these reasons the water intercepted by vegetation lowers the discharge calculated by the second model to about 10 3 m^3/s. The variation of infiltration through time was also taken into account so the hydrograph rising curve is gradual rather then steep. The other characteristics of the hydrograph obtained from the second model can also be seen in my model: the steep decrease of the hydrograph through time and the underestimation of the discharge.

The time-area diagram shows the reason for the fast increase of the discharge in the modeled hydrograph. The fast accumulation of the flow due to the circular shape of the catchment causes this result.

An explanation for the fast decrease of the discharge is the speed calculation in the model. The speed calculation algorithm calculates the speed of flow in case of maximum flow for the rainfall intensity given as a parameter. This speed is then considered constant through time and does not account for flow variation through the area of each cell. Therefore, the hydrograph shape is very close to the shape of the time-area diagram and the rise and fall of the hydrograph are very steep. A routing procedure accounting for storage in the catchment and attenuation of flow would give a more accurate result. However, such a routing procedure is not possible to apply with the available data and needs an iterative execution with discharge calculation in each time step.

Other applications of the model will be made to compare the results and raise some clear conclusions on the usability of these results.

Someşul Cald basin

In the case of the Someşul Cald catchment, the Time-Area diagram and travel time can be seen in Fig. 8.11 The calculated time of concentration was 1812 min because of the larger area of the catchment compared to the Beliş catchment and because of the fact that the streams are more

spread through the catchment.
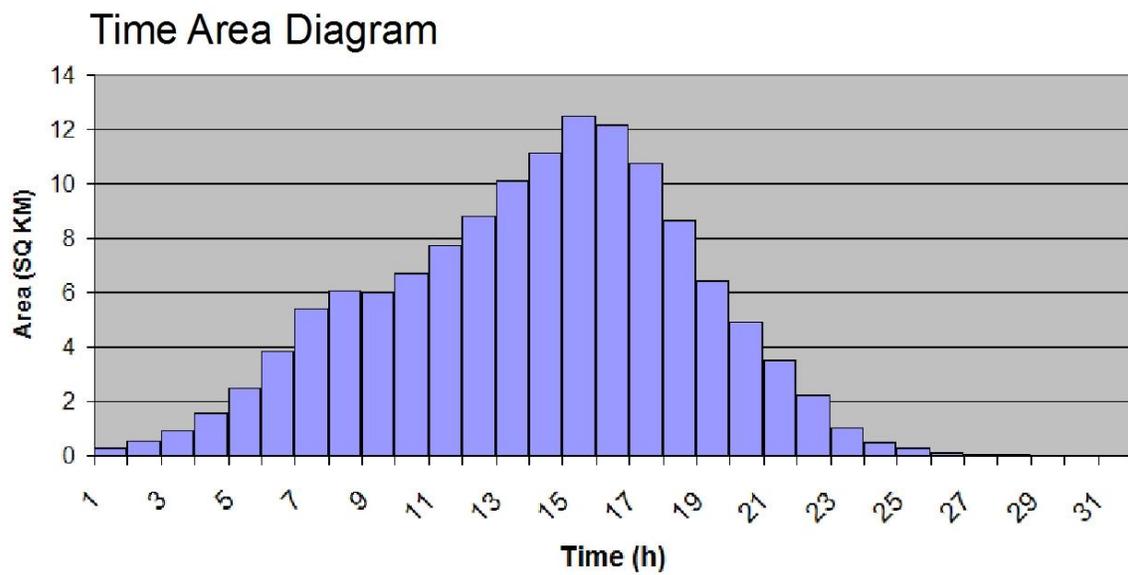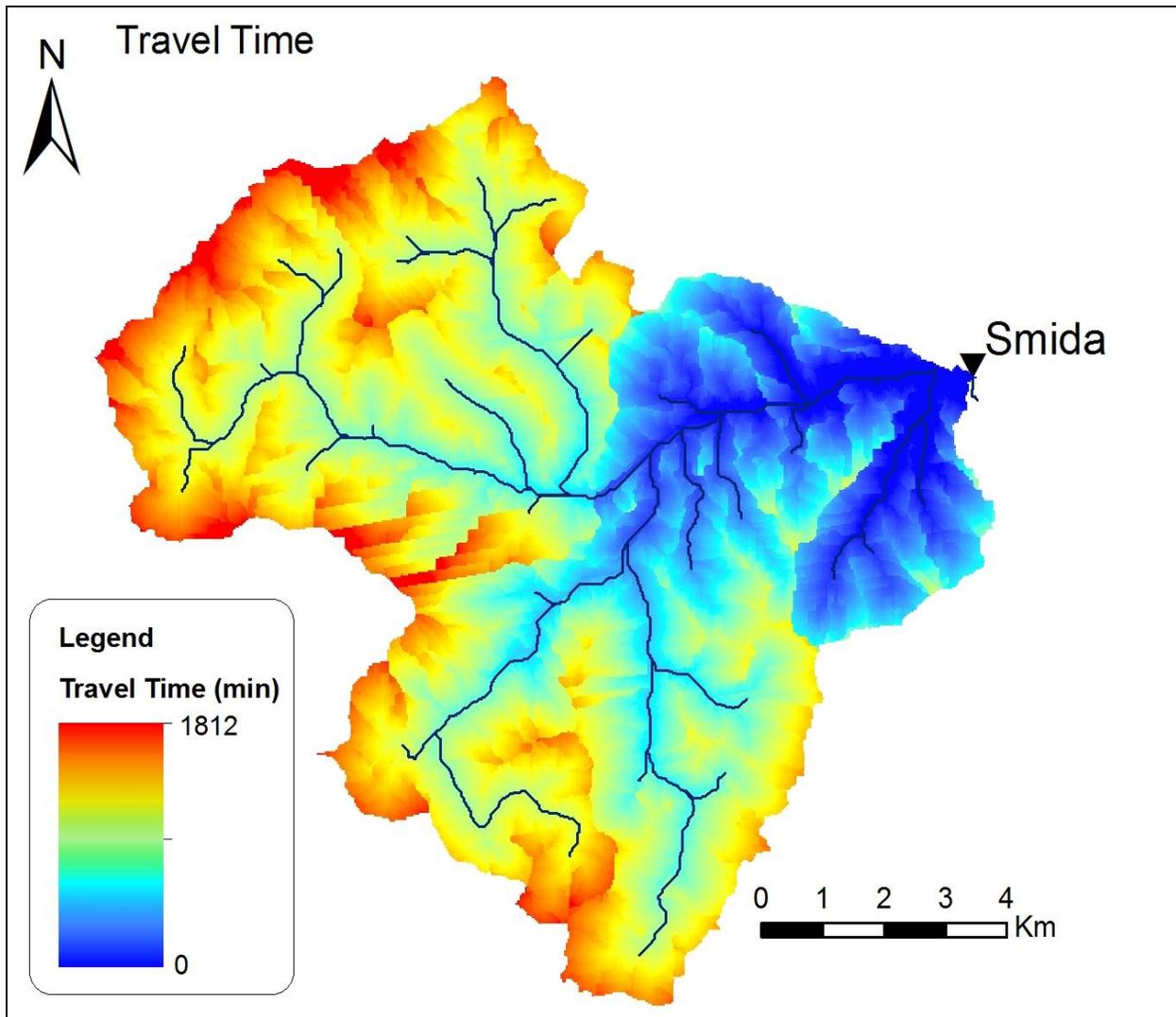
Fig. 8.11 Travel time and time-area diagram in the Someşul Cald basin

The hydrographs were calculated for each day during the studied period (11-15 July 2005) like in the Beliş case and were cumulated using the same MATLAB algorithm.

The comparison between the discharge hydrograph obtained from the model application and the measured hydrograph can be seen in Fig. 8.12.
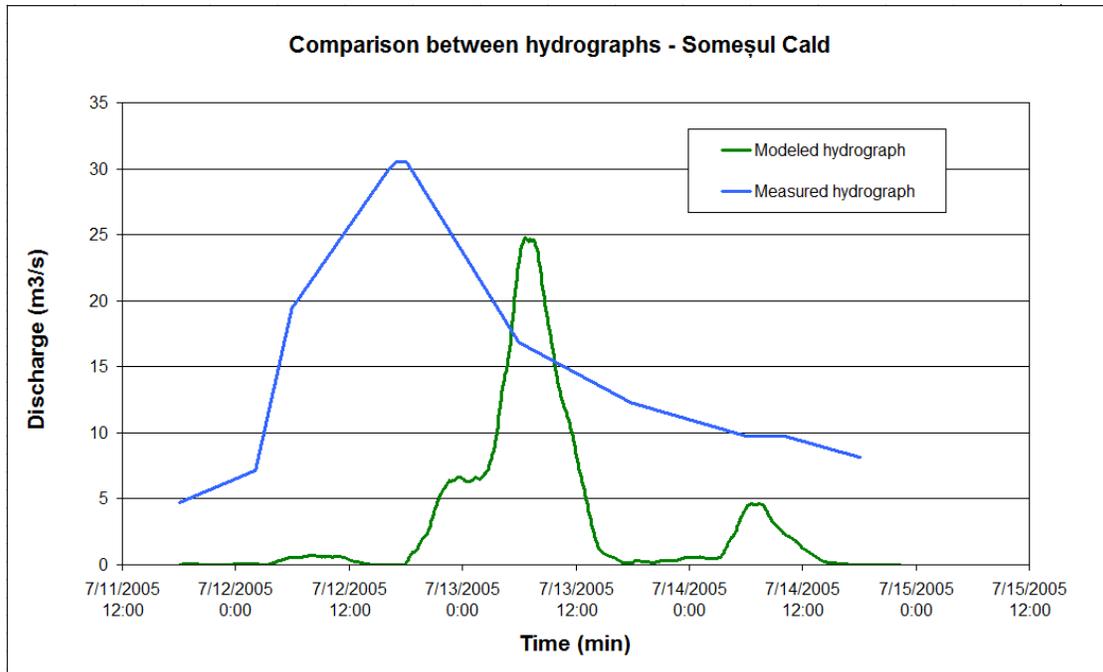

Fig 8.12 Comparison between measured and modeled hydrograph in the Someşul Cald catchment
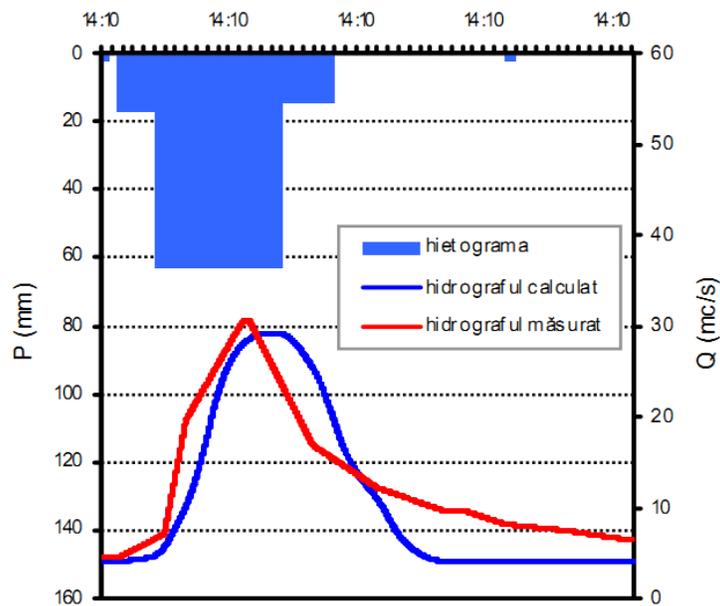
Fig 8.12 Comparison between measured and modeled hydrograph in the Someşul Cald catchment

The starting point of the first rainfall was considered to be h 18:00 on the 11th of July. The other rainfalls were placed continuously at the start of each day. The discharge variation would be modeled better of the starting points of each rainfall would be known.

The measured hydrograph peak is about 4 m^3/s higher than the modeled hydrograph peak like in the previous case. A model that would also account for hypodermic and groundwater flow would give results closer to the measured peak of the hydrograph. Another problem in the Someșul Cald catchment is the source of the river. The main spring of the Someșul Cald river comes from one of the largest karst areas in Romania and the groundwater processes taking place in these areas cannot be measured and accounted for. The model was not designed to model discharge generated from other sources than rainfall.

After the hydrograph peak, the discharge is quite steep in this case too. Due to the fact that the catchment is larger and has a longer total flow length for the streams, more storage areas from within the catchment are not taken into account. A storage coefficient could improve this situation and give more accurate results, but the coefficient would have to be obtained without surveys in the field, which are not possible most of the time.

The cumulated discharge values calculated for each catchment were calculated by adding the discharge generated by each rainfall event with the corresponding delay given by the rainfall start time. Different applications in areas that are not influenced by underground karst processes may lead to better results because of the little data available when running the model.

Validation conclusions

The case study on the flash floods from July 2005 from the two hydrographic basins has met the following difficulties:

- the reduced density of the places where the precipitations are measured (meteorological stations, pluviometric posts); the two pluviometric stations (Smida, Poiana Horea) would have also been helpful, together with the seven meteorological stations (Vladeasa, Huedin, Zalau, Cluj-Napoca, Dej, Baisoara, Turda), to obtain more specific interpolations.

-for the two pluviometric stations situated at the closing sections of the studied basins we only had data from 12 hours to 12 hours measurements of precipitations, without knowing the start point or the duration, or the intensity of the rain; this information would have been necessary to calculate the debits for each isochrone area and to generate the runoff hydrograph. The start point of the rain would allow a better positioning of the hydrograph obtained by modeling the time axis.

-the low number of debit measurements at the hydrometric stations (max 4-5 measurementrs/day), as compared to the modeled debit values (minute by minute); this inconsistency imposes differentiations especially on hydrograph forms.

To ensure a higher density (in time) of validation data and to avoid the problem of the underground flow the better solution would be to measure the debit in periods of time with abundant precipitations in small basins, uninfluenced or weakly influenced by the underground flow.

I do not exclude the imperfections of the algorithm used for modeling, but as long as the input data and the comparing data do pose a few problems it is very difficult to quantify the errors of the model.

Despite the low resolution of temporal and spatial data, the analyzed examples confirm that the small basins runoff estimation model is good and it stands for an alternate solution in anticipating flash floods.

8. Conclusions

This work presented the application of a methodology based on Geographical Information Systems for calculation the flash flood hydrograph in small catchments in mountainous areas. The model is automated, and only requires the input of some parameters related to the rainfall and the terrain.

This PhD thesis presented a methodology on Geographical Information Systems for indirect estimation of flows generated by torrential rainfalls that cause flash floods. This methodology can be very useful in the context of irrational changes in land use (mainly caused by massive deforestation) and a lack of flood protection infrastructure in small mountain basins.

The discharge estimation requires an indirect methodology because most small mountainous catchments lack flow measuring stations and difficulty moving through the area to make measurements is difficult. The created modules tried to make the estimates with a small number of parameters, most of which are available for free on the internet.

The GIS modeling methodology for flash floods according to land characteristics (soils, land use, elevation, slope) that I propose is based on known and tested methods from hydrology (SCS-CN method, rational method, time-area method) and uses different possibilities for modeling from GIS products to achieve the result. The modeling process is automated in the form of GIS modules that address different aspects of the problem and can be used together, independently or in other models as needed.

The model was created with ease of use and data availability in mind, so most of the data needed to apply it can be obtained for free from the internet. The study presented the methods of collecting and procesing the datasets needed for the model application and some workflows or scripts used to automate this processing.

For the implementation of the conceptual model presented I created Python scripts and ArcGIS models using the ModelBuilder interface. Every component used in the model is presented through the study in a graphical form (in the case of ModelBuilder models) or as commented

source code (for Python scripts).

The methodology can be used as a warning tool for flash flood if the predicted rainfall is used as an input. The model can also be used to determine the necessary information for building flood protection infrastructure or taking the necessary actions in preventing flood disasters.

Once the runoff hydrograph is estimated in the ungauged sections, another stage in the determination of the flash flood can be started, and that is the estimation of the water level caused by the flood and the areas that are affected by the flood.

The applications of the model have shown good results in some study catchments but the model needs further work on parameter calibration. More parameters have to be taken into account to obtain a better result and surface storage is one of the main parameters that needs to be considered.

The results are also affected by the flow velocity and concentration time calculated using SAGA GIS (Olaya 2003). This flow velocity is valid only for the maximum flows and therefore the velocity could be overestimated and the hydrograph peak may arrive faster than expected. It was not the case in the validation catchments but the rainfall starting time is unknown so this verification is not possible at the moment.

The runoff speed component can only be replaced by other methods that need more data and calibration. The Manningțs formula which is another empirical way of calculating flow speed is only usable in channel flow and the TR55 methodology cannot be applied without other data on soils amd time series with rainfall for several years.

An application with average speeds for different areas as described in books might give similar results with the one using SAGA GIS. The methodology for deriving average speeds for these areas is presented in the Urban Hydrology for small watersheds TR55 manual from USDA and could be implemented at a later time if the needed data become available.

The research can also be continued by evaluating the flood risk in the settlements from the

basin area and estimation of the human and financial losses generated by a flood.

Results provided by the model are represented both as thematic layers that can be used in GIS (runoff depth, runoff coefficient, volume of water drained from each raster cell, time of water flow through the basin) as well as tabular results (flow–time series) or graphic results (runoff hydrographs). The model covers a major lack in hydrological modeling because most existing models are designed for large basins for urban areas. Also, the calibration of the model can be easily achieved due to the small number of parameters, making its application possible in areas where the available data are minimal.

Usability of the model includes:

- Warning tool based on weather forecasts

- Tool support in infrastructure projects for flood protection

- Planning for rural or urban development

- Directions for further development:

- The methodology can be implemented using only Open Source programs and libraries to build a public domain freely available model

- The results obtained with this methodology will be validated by direct measurements in the field, where possible

- Introducing the possibility of using other freely available datasets (for example NCDC climate data) or remote sensing data.

References

*** Engineering, Design – *Runoff from snowmelt*, U.S. Army Corps of Engineers, Washington, 1998.

*** *Rural Development in the European Union - Statistical and Economic Information* - Report 2010, http://ec.europa.eu/agriculture/agrista/rurdev2010/ , visited February 2011

*** United States Department of Agriculture, (1986), *Urban Hydrology For Small Watersheds*, Technical Release 55 (TR-55), Natural Resources Conservation Service, Conservation Engineering Division.

Abbott M.B., Bathurst J.C., Cunge J.A., O'Connell, Rasmussen J., (1986a), *An introduction to the European Hydrological System - Systeme Hydrologique Europeen, SHE. History, philosophy of a physically-based, distributed modeling system*, Joumal of Hydrology, 87, pag. 45-59.

Abbott M.B., Bathurst J.C., Cunge J.A., O'Connell, Rasmussen J., (1986b), *An introduction to the European Hydrological System-Systeme Hydrologique Europeen, SHE. Structure of a physically-based, distributed modeling system*, Journal of Hydrology, 87, pag. 61-77.

Abbott M.B., Refsgaard J.C., (1996), *Distributed Hydrological Modeling (Water Science, Technology Library),* Springer, USA.

Abdelkader M., Mohamed E., Abdelkader D., (2007), *Hydrologic model combined to a GIS for estimating hydrologic balance at watershed scale - Application to the watershed of Macta (Western Algerian)*, Proceedings of the 2nd IASME / WSEAS International Conference on Water Resources, Hydraulics & Hydrology, Portoroz, Slovenia, pp. 26-30.

Ackoff R.L., (1964), *General Systems Theory, Systems Research: Contrasting Conceptions Of Systems Science*, în: Views on general systems theory, ed. Mesarovic M.D., John Wiley & Sons, USA.

Ahmad B., Yang D., Koike T., Ishidaira Hiroshi, Li Chong, Graf T., (2005), *Remote sensing based snowmelt runoff model coupled with distributed hydrological model in Upper Yellow river basin*, Annual Journal of Hydraulic Engineering, JSCE, Vol. 49, pp. 331-336.

Allen D.W, (2011), *Getting to Know ArcGIS ModelBuilder*, ESRI Press ,USA.

Al-Smadi M., (1998), *Incorporating Spatial, Temporal Variation Of Watershed Response In A Gis-Based Hydrologic Model*, MsC Thesis, Faculty of the Virginia Polytechnic Institute, State University, Blacksburg http://scholar.lib.vt.edu/theses/available/etd-121698-112858/unrestricted/smadi.pdf , visited 12 February 2010

ANAR - *Diminuarea riscurilor in cazul producerii calamitatilor naturale si pregatirea pentru situatii de urgenta* - http://www.rowater.ro/Continut%20Site/Diminuarea%20riscurilor.aspx, visited 13 April 2011.

Anderson M.C., Burt T.P., (eds.), (1990), *Process Studies in Hillslope Hydrology.* John Wiley,

Chichester.

Andrews R.G., (1954), *The use of relative infiltration indices in computing runoff.* in: Rainfall-runoff relationship, Singh V.P., (ed.), Water Resources Publications, Littleton.

ANI-*Climate Change To Yield More Extreme Rainfall In Future,* http://www.thaindian.com/newsportal/health/climate-change-to-yield-more-extreme-rainfall-in-future_100234137.html, visited 18 August 2010.

Aniţan I., (1974), *Scurgerea maximă în bazinul hidrografic Someş-Crasna*, Teză de doctorat. Universitatea Babeş-Bolyai , Cluj-Napoca.

Arc Hydro Tools Overview v1.1 Final (handbook), (2005), ESRI Press ,USA

Arctur D., Zeiler M., (2004), *Designing Geodatabases: Case Studies In GIS Data Modeling*, ESRI Press, USA.

Arghiuş V., (2008), *Studiul viiturilor de pe cursurile de apă din estul Munţilor Apuseni şi riscurile asociate*, Casa Cărţii de Ştiinţă, Cluj-Napoca.

ASTER global digital elevation map http://asterweb.jpl.nasa.gov/gdem.asp

Bahremand A., De Smedt F., (2006), *Parameter Sensitivity, Uncertainty Analysis Of The Wetspa Model Using PEST*, Proceedings of the 2006 IASME/WSEAS Int. Conf. on Water Resources, Hydraulics & Hydrology, Chalkida, Greece, pp. 26-35.

Bárdossy A., Duckstein L., Bogárdi I., (1995), *Fuzzy Rule - Based Classification Of Atmospheric Circulation Patterns*, International Journal of Climatology, No.15, pp. 1087-1097.

Bathurst J.C., (1986a), *Physically-Based Distributed Modeling Of An Upland Catchment Using The Systeme Hydrofogique Europeen*, Journal of Hydrology, No.87, pp. 79-102.

Bathurst.C, Wicks J.M., O'Connell P.E., (1995), *The SHE/SHESED Basin Scale Water Flow, Sediment Transport Modeling System*, in: Singh, V.P., (ed.), Computer Models of Watershed Hydrology. Water Resource Publications, Highlands Ranch, CO, pp. 563-594.

Beasley D.B., Huggins L.Ft, Monke E.J., (1980), *Answers: A Model For Watershed Planning,* Transactions of the American Society of Agricultural Engineer, 23, pp. 938-944.

van den Besselaar Else, Tank A K., van der Schrier G., van Engelen A. (2011), *The European Climate Assessment & Dataset (ECA&D),* WCRP OSC Climate Research in Service to Society, 24-28 October 2011, Session C4, Poster TH257A

Beven J.K., (1997), *TOPMODEL User Notes. Windows Version 97.01*, Lancaster University, UK.

Beven J.K., (2001), *Rainfall-Runoff Modeling*, John Wiley & Sons Ltd, SUA.

Beven J.K., (2003), *Rainfall-Runoff Modeling. The Primer,* John Wiley & Sons, Ltd., Chichester, USA.

Beven J.K., Kirkby M.J., (1979), *A Physically Based Variable Contributing Area Model Of Basin Hydrology*, Hydrologic Science Bulletin, No. 24(1), pp. 43-69.

Beven J.K., Warren R., Zaoui, (1980), *SHE: Towards A Methodology For Physically-Based Distributed Forecasting In Hydrology*, IAHS Publication No. 129ᵣ Wallingford, UK, pp.133-137.

Beven J.K., Kirkby M.J., Schofield N., Tagg A.F., (1984), *Testing A Physically-Based Flood Forecasting Model (TOPMODEL) For Three U.K. Catchments*, Journal of Hydrology, No.69, pp.119- 143.

Beven K., Quinn P., Romanowicz R., Freer J., Fisher J., Lamb R., (1995). *TOPMODEL. A user guide to the distribution version for DOS (95.02).* Lancaster University, UK.

Bilaşco Şt., (2007), *Determinarea debitului maxim utilizând metoda hidrografului unitar*, Geographia Napocensis, Anul I, Nr. 1-2, pp. 84-90.

Bilaşco Şt., (2008), *Implemenarea S.I.G în modelarea viiturilor de versant*, Casa Cărţii de Ştiinţă, Cluj-Napoca.

Bilaşco Şt., (2008), *Model G.I.S de estimare a coeficientului de scurgere adaptat după Frevert*, Geographia Napocensis, Nr. 1, pp. 38-45.

Bilaşco Şt., Haidu I., (2006), *The Valuation of Maximum Runoff on Interbasinal Areas, Assisted by GIS*, Geographia Technica, No.2, pp.1-6, Cluj-Napoca.

Bossard M., Feranec J. and Otahel J., (2000), *CORINE land cover technical guide*, European Environment Agency, http://www.dmu.dk/fileadmin/Resources/DMU/Udgivelser/CLC2000/technical_guide_addenum.pdf, visited March 2011

Bourletsikas A., Baltas E., Mimikou Maria, (2006), *Rainfall-Runoff Modeling For An Experimental Watershed Of Western Greece Using Extended Time-Area Method, GIS*, Journal of Spatial Hydrology, Vol. 6, No. 1, pp. 93-104.

Brimicombe A., (2009), *GIS, Environmental Modeling, Engineering*, CRC Press I Llc.

Burt T.P., (1987), *Measuring Infiltration Capacity*, Geography Review, vol.1, pp.37–39.
Buta I., (1967), *Bazinul Someşului. Studiu hidrologic*, Teză de doctorat. Universitatea Babeș-Bolyai, Cluj-Napoca.

Butler D., Davies J.W., (2000), *Urban Drainage*, Spon Press, UK.

Butler H., (2005), *A guide to the Python Universe for ESRI Users*, ArcUser, Iowa State University, USA.

Caltech-*Modeling Shows How Climate Change Yields More Extreme Rainfall*
http://today.caltech.edu/today/story-display.tcl?story_id=38283,visited 26 Agust 2010.

Calver A., Wood W.L., (1995), *The Institute of Hydrology Distributed Model*, in: Singh V.P. (Ed.), Computer Models of Watershed Hydrology, Water Resource Publications, Colorado, pp.595-626.

Carolyn J. Merry, (2003), *Hydrologic Modeling in GIS CE808–SP03*, Ohio State University.

Changbin L., Feng Z.-D., Songbing Zou, Hui Han, Yongzhong Z., Youhao E., Min Feng, (2005), *GIS-Based Improvement Of the. Clark Method For Simulating Watershed-Scale Overland Flow In The Chinese Loess Plateau,* Geoscience, Remote Sensing Symposium, Proceedings,Vol.6, pp. 4399-4402.

Chanson H., (2004), *The Hydraulics of Open Channel Flow*, Butterworth-Heinemann, Oxford, UK.

Chen C.L., (1982), *Infiltration formulas by curve number procedure*, J.Hydr. Div., ASCE, vol. 108, no. HY7, pp. 823–829.

Chiradeep A., (2003), *A GIS - Remote Sensing Compatible Rainfall-Surface Runoff Model For Regional Level Planning*,
http://www.geospatialworld.net/index.php?view=article&catid=154%3Anatural-resource-management-water-resources&id=16927%3Aa-gis-remote-sensing-compatible-rainfall-surface-runoff-model-for-regional-level-planning&format=pdf&option=com_content&Itemid=41, visted 12 May 2011

Chorley R.J., Haggett P., (1967), *Models In Geography*, Methuen, London.

Chow V.T., Maidment D., Mays L., (1988), *Applied Hydrology*, McGraw-Hill Science/Engineering/Math, New York.

Chow V.T.,(1964), *Handbook Of Applied Hydrology. A Compendium Of Water-Resources Technology*, Mcgraw-HILL Book Company, New York.

Churches T., (2010), *A Demonstration of RPy*, http://rpy.sourceforge.net/rpy_demo.html, visited 26 April 2011.

Clark C.O., (1945), *Storage, The Unit Hydrograph,* Transactions of American Society of Civil Engineering, No.110, pp.1419-1446.

Coll E., Martinez-Llario J., Irigoyen J., Sanz J., (2005), *Publication Of Geographic Information Systems In Internet*, Proceedings: 6th WSEAS International Conference on Automation & Information (ICAI 05), Buenos Aires.

Comisia Europeană - *Răspunsul Uniunii Europene la inundaţiile din România, Ucraina şi Moldova*, 06/08/2008 - http://ec.europa.eu/romania/news/inundatii_ro.htm, visited 4 March 2010.

Congguo T., Congqiang L., (2007), *Simulation of Surface Runoff in Wujiang River Watershed Based on GIS.* Chinese Journal of Geochemistry, No.26(3), pp. 284-289.

Corine Land Cover main page - http://dataservice.eea.europa.eu/dataservice

Crăciun A.I., Haidu I., Bilaşco Şt., (2007), *The SCS-CN Model Assisted By G.I.S. - Alternative Estimation Of The Hydric Runoff In Real Time.* Geographia Technica Nr. 1/2007, pp. 1-7.

Crăciun A.I.., Haidu I., Magyari-Saska Zs., Imbroane Al., (2009), *Estimation Of Runoff Coefficient According To Soil Moisture Using GIS*, Geographia Technica, No. 2, pp. 1-10.

Crăciun A.I., (2011). *Estimarea indirectă, cu ajutorul GIS, a umezelii solului în scopul modelării viiturilor pluviale. Aplicaţii în Munţii Apuseni*, Teză de doctorat, Universitatea Babeş-Bolyai, Cluj-Napoca.

Crăciun A.I., (2007), *Use G.I.S To Establish Some Parameters Useful To Measure The Time Of Concentration, Runoff Coefficient*, Geographia Technica, No.2, pp. 12-19.

Crăciunescu V., (2004), *Site cartografic interactiv pentru Municipiul Suceava,* Lucrare de licenţă, Universitatea din Bucureşti, Facultatea de Geografie.

Crawford N.H., Burges S.J.,(1993), *History of the Stanford Watershed Model*, Water Resources Impact, Vol. 6, No. 2, pp. 141-148

Davie T., (2008), *Fundamentals of Hydrology,* Routledge Fundamentals of Physical Geography, UK.

Dawdy D.R., O'Donnell T., (1965), *Mathematical models of catchment behaviour*, Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers (HHY4), pp.123-127.

Delaunay B., (1934), *Sur la sphère vide,* Bulletin of Academy of Sciences of the USSR, No. 6., pp. 793-800

Dey B., Sharma V.K., Rango A., (1989), *A Test of Snowmlet-Runoff Model for Major River Basin in Western Himalayas*, Nordic Hydrology, No. 20, pp. 167-178.
DHI Water & Environment - MIKE SHE - integrated catchment modeling http://www.mikebydhi.com/Products/WaterResources/MIKESHE.aspx visited10 March 2011

DHI Water & Environment - *Process models details*, DHI Group, ttp://www.dhigroup.com/Software/WaterResources/MIKESHE/Details/ProcessModels.aspx, visited 11 March 2010.

DHI Water & Environment, (2007), *Mike SHE User Manual*, vol. 2: Reference Guide, DHI, Denmark.

D'Hont R.L., (2007), *Un site pour les Systèmes de Références Spatials*, ttp://3liz.com/blog/rldhont/index.php/2007/07/26/71-un-site-pour-les-systemes-de-references-spatials, visited 17 martie 2011.

Diaconu, C., Şerban P., (1994), *Sinteze si regionalizari hidrologice*, Editura Tehnica, Bucureşti.

Domniţa M., Crăciun A.I., Haidu I., Magyari-Sáska Zs., (2010), *Geographical Information System Module For Deriving The Flash Flood Hydrograph In Mountainous Areas*, Proceedings of the 4th European Computing Conference. Published by WSEAS Press, pp.260-265.

Domniţa M., Crăciun A.I., Haidu I., (2009), *GIS In Determination Of The Discharge Hydrograph Generated By Surface Runoff For Small Basins*, Geographia Technica No. 2, pp. 11-22.

Dooge J-C.I., (1957), *The Rational Method For Estimating Flood Peaks*, Engineering, No.184, pp. 311-313, 374-377.

Douglas, D.H., (1986), *Experiments to Locate Ridges, Channels to Create a New Type of Digital Elevation Models*, Cartographica, No.23(4), pp.29-61.

Downey A., (2008), *Think Python - How to Think Like a Computer Scientist,* Green Tea Press, Needham, Massachusetts, USA.

Dozier J., Painter H.T., Rittger K., Frew E.J., (2008), *Time-Space Continuity Of Daily Maps Of Fractional Snow Cover, Albedo From MODIS*, Advances in Water Resorces, No.31, pp. 1515-1526.

Dubois P.F., Hinsen K., Hugunin J., (1996), *Numerical Python*, Computers in Physics, Vol. 10, Nr. 3., pp. 262-267.

Eastman J.R., (2003), *IDRISI Kilimanjaro Guide to GIS, Image Processing,* Clark Labs, Clark University, Worcester, MA, USA, http://www.gis.unbc.ca/help/software/idrisi/kilimanjaro_manual.pdf, visited 24 April 2011.

Easton M.Z., Fuka R.D., Walter M.T., Cowan M.D., Schneiderman M.E., Steenhuis S.T., (2008), *Re-Conceptualizing The Soi Land Water Tool (SWAT) Model To Predict Runoff From Variable Source Areas*, Journal of Hydrology , 347, pp. 279-291.
ESRI, (2001), *ArcView Version 3.2a Software, User Manual,* Environmental Systems Research Institute, Redlands, CA.

Fairfield J., Leymarie P., (1991), *Drainage Networks from Grid Digital Elevation Models,* Water Resources Research, No.30(6), pp.1681-1692.

Federal Climate Complex, (2010), *Global Surface Summary Of Day Data Version 7*, updated 09/08/2010
ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt, visited 16 February 2011

Fennessey L.A., Hawkins R.H., (2001), *The NRCS Curve Number, a New Look at an Old Tool*. Proceedings of the 2001 Pennsylvania Stormwater Management Symposium. Villanova University.

Fennessey L.A., Hamlett J.M., Miller A.C., (2001b), *Accuracy, Precision of NRCS Models for Small Watersheds*, Journal of AWRA, no. 37(4).

Feuerbacher B., Stoewer H., (2005), *Utilization of Space: Today, Tomorrow*, Springer, USA.

Florea N., Balaceanu V., Munteanu I., (1994), *Soil map of Romania at 1:200000 scale*, Proceedings from Conferinta Nationala pentru Stiinta Solului, Tulcea (Romania), 29 Aug - 3 Sep.

Freeze R.A., Harlan R.L., (1969), *Blueprint For A Physically-Based Digitally Simulated Hydrologic Response Model*, Journal of Hydrology, no.9, pp. 237-258.

Gao J., (2009), *Digital Analysis of Remotely Sensed Imagery*, The McGraw Hill,USA.

Garbrecht J., Martz L.W., (1995), *TOPAZ: An Automated Digital Landscape Analysis Tool For Topographic Evaluation, Drainage Identification, WatershedSegmentation, Subcatchment Parameterization: TOPAZ User Manual*, USDA-ARS Publication No. 95(3).

Georgakakos K.P., (1986a), *On The Design Of National, Real-Time Warning Systems With Capability For Site-Specific, Flash-Flood Forecasts*, Bulletin of the American Meteorological Society, No.67, pp.1233-1239.

GRASS GIS - Simulation Models http://grass.itc.it/intro/modelintegration.html, visited 22 April 2011.

Grayson R.B., Blőschl G., Moore I.D., (1995), *Distributed parameter hydrologic modeling using vector elevation data: THALES, TAPES-c*, in: Singh V.P., (ed.), Computer Models of Watershed Hydrology, Water Resource Publications, Highlands Ranch, CO. pag. 669-696.

Green J.I., Nelson E.J., (2002), *Calculation Of Time Of Concentration For Hydrologic Design, Analysis Using Geographic Information System Vector Objects*, Journal of Hydroinformatics 04.2, pp. 75-81.

Gruntfest Eve, Huber C.J,.(1991), *Toward A Comprehensive National Assessment Of Flash Flooding In The United States,* Episodes*, No.***14**, pp.26-35.

Gruntfest Eve, Handmer J., (2001), *Coping with Flash Floods*, Kluwer Academic Publishers, Netherlands.

Halley Mary C., White Suzanne O., Watkins Edwin W., *ArcView GIS Extension for Estimating Curve Numbers* http://proceedings.esri.com/library/userconf/proc00/professional/papers/pap657/p657.htm, visited 4 March 2011.

Haidu I., (2006), *Elemente de Hidrologie*, ed. Academic Press, Cluj-Napoca.

Hengl T., Gruber S., Shrestha D.P., (2003), *Digital Terrain Analysis in ILWIS. Lecture notes*.

ITC - International Institute for Geo-Information Science, Earth Observation, http://www.itc.nl/library/Papers_2003/misca/hengl_digital.pdf., visited 2 April 2011.

Hengl T., Reuter H., (2008), *Geomorphometry*, vol. 33: Concepts, Software, Applications (Developments in Soil Science), Elsevier B.V., Hungary, pp. 387-410.
Hirschboeck K.K., (1987), *Catastrophic Flooding, Atmospheric Circulation Anomalies*, in: Mayer L., Nash D., (Eds.), Catastrophic Flooding, Allen, Unwin, Boston, pp. 23-56.

Hock R., (2003), *Temperature Index Melt Modeling In Mountain Areas*, Journal of Hydrology, No.282, pp. 104-115.
Hofierka J., Mitášová Helena, Neteler M., (2009), *Geomorphometry in GRASS GIS,* in:,

Hornberger G.M., Beven K.J., Cosby B.J., Sappington D.E., (1985), *Shenandoah Watershed Study: Calibration Of A Topography-Based, Variable Contributing Area Hydrological Model To A Small Forested Catchment*, Water Resources Research, No.21, pp.1841-1850.

Horton R.E., (1945), *Erosional Development Of Streams, Their Drainage Basins: Hydro-Physical Approach To Quantitative Morphology*, Geological Society of America Bulletin, No. 56 (3), pp. 275–370.

Huang B., Jiang B., (2002), *AVTOP: A Full Integration Of TOPMODEL Into GIS,* Environmental Modeling, Software, Vol. 17, No. 3, pp. 261-268.

Hutchinson M.F., (1996), *A Locally Adaptive Approach To The Interpolation Of Digital Elevation Models,* Third International Conference/Workshop on Integrating GIS, Environmental Modeling, NCGIA, University of California, Santa Barbara http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/papers/local.html, visited 3 April 2010

Hutchinson M.F., (2000), *Optimising the degree of data smoothing for locally adaptive finite element bivariate smoothing splines,* ANZIAM No.42, pp.774-796.

Jain S.K., Storm B., Bathurst J.C., Refsgaard J.C., Singh R.D., (1992), *Application Of The SHE To Catchments In India: Part 2 - Field Experiments, Simulation Studies With The SHE On The Kolar Subcatchment Of The Narmada River*, Journal of Hydrology, No.140, pp. 25-47.

Jintao Liu, Xi Chen, Jiabao Z., Flury M., (2009), *Coupling The Xinanjiang Model To A Kinematic Flow Model Based On Digital Drainage Networks For Flood Forecasting*, Hydrol. Process., No.23, pp. 1337–1348.

Johnson D., (1991), *Unit Hydrograph Theory*, Presented at Hydromet 99, http://www.comet.ucar.edu/class/hydromet/07_Jan19_1999/html/unit_hydro_theory/index.htm, visited 12 January 2011.

Jolma A., Ames D.P., Horning N., Mitasova Helena, Neteler M., (2008), *Environmental Modeling Using Open Source Tools*, in: Shekar S., Xiong H., (Eds.), Encyclopedia of GIS,

Springer, USA.

Kalechman, M, (2009), *Practical MATLAB basics for engineers,* Taylor & Francis Group, LLC

Kelsch M., (2001), *Hydrometeorological Characteristics Of Flash Floods,* in: Gruntfest E., Handmer J., (Eds.), Coping with Flash Floods, Environmental Security, Vol. 77, Kluwer Academic Publishers, Boston, pp. 181-193.

Kimball J.S., Running S.W., Nemani R.R., (1997), *An improved method for estimating surface humidity from daily minimum temperature*, Agriculture, Forest Meteorology, vol.85, pp. 87-98.

Kirkby M., (1975), *Hydrograph Modeling Strategies*, in: Peel R., Chisholm M., Haggetl P., (Eds.), Processes in Physical, Human Geography, Heinemann, London, pp. 69-90.

Kirkby M., (1978), *Implications For Sediment Transport*, in: Kirkby M.(ed.), Hillslope Hydrology. John Wiley, Chichester.

Kite C.W., (1995b), *The SLURP Model*, in: Singh V.P., (ed.), Computer Models of Watershed Hydrology, Water Resource Publications, Highlands Ranch, CO, pp. 521-561.

Kite G.W., (1994), *Hydrological Modeling Using Remotely Sensed Data, Geographic Information Systems*, in: Menon J., (ed.), Trends in Hydrology, CSRI, Trivandrum, India, pp. 191-208.

Kniivila, Matleena, (2004), Land Degradation and Land Use/Cover Data Sources, http://unstats.un.org/unsd/environment/envpdf/landdatafinal.pdf, visited January 2011

Kopp S., Noman N., (2008), *ArcGIS Spatial Analyst - Hydrologic Modeling*, ESRI User Conference Technical Workshop, http://www.scdhec.gov/gis/presentations/ESRI_Conference_08/tws/workshops/tw_37.pdf, visited 25 April 2011.

Kull D.W., Feldman A.D., (1998), *Evolution Of Clark's Unit Graph Method To Spatially Distributed Runoff*, Journal of Hydrological Engineering, ASCE, No.3, pp. 9-19.

L'hotellier R., Patriche C.V., (2009), *Using Low-Pass Filters To Improve Regression Models For Climate Variables. A Study Case Of Winter Mean Minimum Temperatures In The French Alps*, Geographia Technica, No. 1, pp. 50-56.

Lindfield G., Penny J., (2001), *Numerical Methods Using MATLAB*, Prentice Hall, USA.

Linsley R.K., Kohler M.A., Paulhus J.L.H., (1982), *Hydrology for engineers*. 2nd ed. McGraw-Hill, New York.

Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W., (2005), *Geographic Information Systems, Science*, John Wiley & Sons, USA.

Lopez-Moreno I.J., Stähli M.*,*(2008), *Statistical Anaysis Of The Snow Cover Variability In A*

Subalpine Watershed: Assessing The Role Of Topography, Forest Interactions, Journal of Hydrology, No. 348, pp. 379-394.

Lutz M., (2010), *Python Pocket Reference*, O'Reilly Media, Canada.

Magyari-Saska Zs., (2008), *Dezvoltarea algoritmilor S.I.G pentru calculul riscurilor geografice naturale. Aplicaţie la Bazinul Superior al Mureşului*, Teză de doctorat, Universitatea Babes-Balyai, Cluj-Napoca.

Maidment D.R, (1996), *GIS, Hydrologic Modeling - an Assessment of Progress* Presented at The Third International Conference on GIS, Environmental Modeling, http://www.ce.utexas.edu/prof/maidment/gishydro/meetings/santafe/santafe.htm, visited 18 October 2010.

Maidment D.R., (2002), *ArcHydro, GIS for Water Resources*, ESRI Press, Redlands, California.

Maidment D.R., (2004), *Arc Hydro Data Model: A Hydrologic Data Integration Tool*, Center for Research in Water Resources, GIS Applications in Hydrology, Vol. 3, No. 3, University of Texas, pp. 18-19.

Maidment, D.R., (1993), *Developing A Spatially Distributed Unit Hydrograph By Using GIS*, in: Kovar K., Nachtnebel H.fP., (Eds.), Application of Geographic Information Systems in Hydrology, Water Resources Management, IASH Publ.**,** pp. 181-192.

Marek M.A., (2009), *Hydraulic Design Manual,* Texas Department of Transportation (TxDOT), Design Division (DES).

Martinec J., Rango A., Roberts R., (2008), *Snowmelt Runoff Model (SRM), Users manual vers1.11*, New Mexico College of Agricultural land Home Economics.

Martínez V., Dal-Ré R., García A.I., Ayuga F., (2000),: *Modelación distribuida de la escorrentía superficial en pequeñas cuencas mediante SIG. Evaluación experimental,* in: Ingeniería Civil, No. 117, pp. 49-58.

Martinoni Daria. 2002. *Models and Experiments for Quality Handling in Digital Terrain Modeling*, PhD Thesis, University of Zurich.

Matziaris V., Sakellariou M., (2005), *Modeling The Effect Of Rainfall In Unsaturated Slopes*, IASME Transactions, International Conference On Energy, Environment, Ecosystems, Sustainable Development, Issue 3, Vol. 2, pp. 442-447.

McCuen R.H., (1982), *A Guide to Hydrologic Analysis Using SCS Methods Prentice Hall*, Inc. Englewood Cliffs, New Jersey.

Merry Carolyn J.,(2003), *Hydrologic Modeling in GIS CE808–SP03,* Ohio State University, USA.

Merwade V., (2010), *Tutorial On Using HEC-GeoRAS with ArcGIS 9.3*, School of Civil

Engineering, Purdue University, USA.

Mic Rodica, Corbuş C., (1999), *Estimarea debitelor maxime în cazul viiturilor rapide în Bazinul Hidrografic Crişuri pe baza precipitaţiilor şi a stării de umiditate a solului*, Institutul Naţional de Hidrologie şi Gospodărire a apelor, Bucureşti.

Michaelides Katerina, Wainwright J., (2002), *Modeling The Effects Of Hillslope-Channel Coupling On Catchment Hydrological Response*, Earth Surface Processes, Landforms, No.27, pp. 1441-1457.

Mihăilescu V., (1963), *Carpaţii sud-estici de pe teritoriul R. P. Române. Studiu de geografie fizică cu privire specială la relief*, Edit. Științifică, București.

Mihalik Elizabeth N., Levine S.N., Amatya M. Devendra, (2008), *Rainfall-Runoff Modeling of Chapel Branch Creek Watershed using GIS-based Rational, SCS-CN Methods*, ASABE, Rhode Island, SUA.

Miller S.N., Semmens D.J., Miller R.C., Hernandez M., Miller W.P., Goodrich D.C., Kepner W.G., Ebert D., (2002), *GIS-Based Hydrologic Modeling: The Automated Geospatial Watershed Assessment Tool*, Proc. 2nd Federal Interagency Conf. on Hydrologic, Las Vegas, NV.

Miller N., Cronshey R., (1989), *Runoff curve numbers, the nex Step,* Proc., Int. Conf. on Channel Flow, Catchment Runoff, Uni-versity of Virginia, Charlottesville.

Mishra S.K., Singh V.P., (2003), *Soil Conservation Service Curve Number (SCS-CN) Methodology*, Kluver Academic Publishers, Dordrecht, Netherland.

Mishra S.K., Singh V.P., (1999), *Another Look at SCS-CN Method*, J. Hydrol. Eng. vol.4, no.3. pp. 257-265

Mishra, S.K., Jain M.K., Singh, V.P., (2004), *Evaluation of the SCS-CN-based model incorporating antecedent moisture*, J. Water Resour. Management, no. 18, pp. 67–589.

Mishra S.K., Jain M.K., Pandey R.P., Singh V.P., (2003), *Evaluation of AMC-dependant SCS-CN-based models using large data of small watersheds*, Water, Energy Int. vol. 60, no. 3, pp.13–23.

Mitas L., Mitasova Helena, (1998), *Distributed Soil Erosion Simulation For Effective Erosion Prevention*, Water Resources Research, No.34(3), pp.505-516.

Mitasova Helena, Thaxton C., Hofierka J., McLaughlin R., Moore A., Mitas L., (2004), *Path Sampling Method For Modeling Overland Water Flow, Sediment Transport, Short Term Terrain Evolution In Open Source GIS*, In: Miller C.T., Farthing M.W., Gray V.G., Pinder G.F., (eds.), Computational Methods in Water Resources, Elsevier, UK.

Mockus, V., (1949), *Estimation of total (peak rates of) surface runoff for individual storms,* Exhibit A of Appendix B, Interim Survey Rep. Grand (Neosho) River Watershed, USDA, Washington, D.C.

Mogil H.M., Monro J.C., Groper H.S., (1978), *NWS's Flash Flood Warning, Disaster Preparedness Program,* Bulletin of the American Meteorological Society, No.59, pp.690-699.

Montgomery D.R., Efi Foufoula-Georgiou, (1993), *Channel Network Source Representation Using Digital Elevation Models,* Water Resources Research, Vol. 29, No. 12, pp. 3925-3934.

Monti Alberto, (2009), *Climate Change, Weather-Related Disasters: What Role for Insurance, Reinsurance, Financial Sectors*, Hastings West-Northwest Journal of Environmental Law & Policy, vol.15, no.1, 151.

Moore I.D., Grayson R.B., Ladson A.R ., (1991), *Digital Terrain Modeling: A Review Of Hydrological, Geomorphological, Biological Applications*, Hydrological Processes, Special Issue: Digital Terrain Modeling in Hydrology, vol.5, no. 1, pp. 3-30.

Moore R.J., Bell V.A., Austin R.M., Harding R.J., (1999), *Methods For Snowmelt Forecasting In Upland Britain*, Hydrology, Earth System Sciences, No. 3(2), pp. 233-246.

Moore R.J., Clarke R.T., (1981), *A Distribution Function Approach To Rainfall-Runoff Modeling*, Water Resources Research, No.17, pp. 1367-1382.

Morid S., Gosain A.K., Keshari A.K., (2004), *Response Of Different Snowmelt Algorithm To Synthesize Climatic Data For Runoff Simulation*, Journal of Earth, Space Physics, Vol. 30, No. 1, pp. 1-9.

Mulvaney T.J., (1851), *On The Use Of Self-Registering Rain, Flood Gauges In Making Observations Of The Relations Of Rainfall, Flood Discharges In A Given Catchment*, Proceedings of the Institution ofCivil Engineers of Ireland, no. 4, pp. 19-31.

Munteanu I., Dumitru M., Florea N., Canarache A., Lacatusu R., Vlad V., Simota C., Ciobanu C., Rosu C., (2005), *Status of Soil Mapping, Monitoring, Database Compilation in Romania at the beginning of the 21st century,* In: R.J.A. Jones, B. Houšková, P. Bullock, L. Montanarella (eds)., *Soil Resources of Europe*, European Soil Bureau Research Report No.9, Office for Official Publications of the European Communities, Luxembourg.

National Meteorological Administration, *Radar -*
http://vremea.meteoromania.ro/taxonomy/term/10 , visited 12 April 2011

National Meteorological Administration, *Available data*
http://www.meteoromania.ro/index.php?id=175 , visited 17 November 2011

NCDC - *What is NCDC?*
http://www.ncdc.noaa.gov/oa/about/whatisncdc.html, accesat 18 iulie 2010.

Neteler M., Mitasova Helena, (2008), *Open Source GIS: A GRASS GIS Approach*, The International Series in Engineering, Computer Science, Vol. 773, Springer, New York.

Neteler M., Helena Mitasova, (2004), *Open Source GIS: A GRASS GIS Approach*. The Kluwer

international series in Engineering, Computer Science (SECS): vol 773, Kluwer Academic Publishers/Springer, Boston.

Niţu C., (1992), *Contribuţii privind realizarea unui pachet de programe pentru construcţia automatizată a hărţilor*, Teza de doctorat, Academia Tehnica Militară, Bucureşti.

O'Callaghan J.F., Mark D.M., (1984), *The Extraction of Drainage Networks From Digital Elevation Data,* Comp. Vision, Graphics Image Process., Vol. 28, pp. 328-344.

Oancea D. I., *Geografia României, Volumul 3: Carpaţii româneşti şi Depresiunea Transilvaniei*, Editura Academiei Republicii Socialiste România, 1987

Obled Ch., Wendling J., Beven K.J., (1994), *The Sensitivity Of Hydrological Models To Spatial Rainfall Patterns: An Evaluation Using Observed Data,* Journal of Hydrology, No.159, pp. 305-333.

Official NASA SRTM site -http://www2.jpl.nasa.gov/srtm/.

Olaya V., (2004), *Hidrologia Computacional y Modelos Digitales del Terreno - Teoria, practica y filosofia de una nueva forma de analisis hidrologico,* Alqua, http://www.gabrielortiz.com/descargas/Hidrologia_Computacional_MDT_SIG.pdf, visited 15 April 2011.

Olaya V., (2010), *Introduction to SEXTANTE*, Edition 1.0, Rev. July 5, http://forge.osor.eu/docman/view.php/13/123/IntroductionToSEXTANTE.pdf, visited 12 April 2011.

Olaya V.A, Gimenez J.C., (2011), *SEXTANTE, A Versatile Open-Source Library For Spatial Data Analysis*, http://sextante.forge.osor.eu/pdf/SextantePaper.pdf, visited 12 April 2011.

Oliphant T.E., (2006), *Guide to NumPy,* Trelgol Publishing, http://www.tramy.us/numpybook.pdf, visited 15 March 2011.

Olivera F., Maidment D.R., (1999), *Geographic Information Systems (GIS)-Based Spatially Distributed Model for Runoff Routing*, Water Resour. Res., No.35(4), pp. 1155–1164.

Open Channel Flow, Pressure Pipe Flow: http://www.lmnoeng.com/literature.htm

Păcurar V.D., (2005), *Utilizarea Sistemelor de Informaţii Geografice în modelarea şi simularea proceselor hidrologice*, Edit. Lux Libris, Braşov.

Paniconi C., Wood E.F., (1993), *A Detailed Model For Simulation Of Catchment Scale Subsurface Hydrologic Processes*, Water Resources Research, No.29, pp. 1601-1620.

Parajka J., Blöschl G., (2008), *The Value Of MODIS Snow Cover Data In Validating, Calibrating Conceptual Models*, Journal of Hydrology, No. 358, pp. 240-258.

Parcul Natural Apuseni - http://www.parcapuseni.ro

Patko R., (2007), *Valea Crişului Alb – Studiu de ecologie si hidrologie*, Teză de doctorat, Universitatea din Bucuresti.

Patriche C.V., (2007), *About The Influence Of Space Scale On The Spatialization Of Meteo-Climatic Variables*, Geographia Technica, No. 1, pp. 68-75.

Pilgrim D.H., Cordery I., (1993), *Flood Runoff*, Handbook of Hydrology, Ch. 9, 9.1-9.42, McGraw-Hill, New York.

Podobnikar T., *Suitable Digital Elevation Model*, Institute of Anthropological, Spatial Studies, Scientific Research Centre of the Slovenian Academy of Sciences, Arts, http://www.gisdevelopment.net/technology/tm/tm003.htm, visited 16 February 2011.

Ponce V.M., Hawkins R.H., (1996), *Runoff curve number: Has it reached maturity?* J. Hydrologic Engrg., ASCE, vol.1, no. 1, pp.11–19.

Quimpo R.G., Debarry P.A., (1999), , *GIS Modules, Distributed Models of the Watershed: Report,* ASCE Task Committee on GIS Modules, Distributed Models of the Watershed, American Society of Civil Engineers.

Quinn P.F., Beven K.J., Lamb R., (1995), *The Ln(A/Tan B) Index: How To Calculate It, How To Use It In The TOPMODEL Framework*, Hydrol. Processes, No.9, pp. 161-182.

Rallison R.E, (1980), *Origin, Evolution of the SCS Runoff Equation*. Watershed Management, ASCE, pp. 912-924.

Ramsey P., (2007), *The State* of *Open Source GIS*, Refractions Research, Victoria, Clifornia.

Rango A., Martinec J., (1987), *Large Scale Effects Of Seasonal Snow Cover, Temperature Increase On Runoff*, IAHS, No. 166, pp. 121-127.

Reginald W.T., Richard J.H., (1980), *Modeling in Geography: a Mathematical Approach*, Barnes, Noble, USA .

Register of Ecological Models: TOPMODEL
http://ecobas.org/www-server/rem/mdb/topmodel.html

Robson A.J., Beven K.J., Neal C., (1992), *Towards Identifying Sources Of Subsurface Flow: A Comparison Of Components Identified By A Physically Based Runoff Model, Those Determined By Mixing Techniques*, Hydrological Processes, No.6, pp. 199-214.

Robson A.J., Whitehead P.G., Johnson R.C., (1993*), An Application Of A Physically Based Semi-Distributed Model To The Balquhidder Catchments,* Journal of Hydrology, No.145, pp. 357-370.

Ross C.N., (1921), *The Calculation Of Flood Discharge By The Use Of Time Contour Plan Isochrones*, Transactions of the Institute of Engineers, No.2, Australia, pp. 85-92.

Rosson M.B., (2005), *The End Of Users*, in: Proceedings of the OOPSLA Conference (keynote presentation), San Diego, California.

Rovey E.W., Woolhiser D.A., Smith R.E., (1977), *A Distributed Kinematic Model For Upland Watersheds*, Hydrology Paper, No. 93, Colorado State university.

Rutter A.J., Kershaw K.A, Robins, P.C., Morton A.J., (1971) A predictive model of rainfall interception in forests, I. Derivation of the model from observations in a plantation of Corsican pine. Agnc. Met., Vol. 9, pp. 367-384.

Rutter A.J., (1975), The hydrological cycle in vegetation, Vol. 1, Vegetation, the Atmosphere, (Ed: JL Montheith), Academic Press, London: 111-154.

Samani Z., (2000), *Estimating Solar Radiation, Evapotranspiration Using Minimum Climatological Data*, Journal of Irrigation, Drainage Engineering, Vol. 126, No. 4, pp. 265-267.

Sanjay K.J., Naresh Kumar, Tanvear Ahmad, (1998), *SLURP Model, GIS For Estimation Of Runoff In A Part Of Satluj Catchment*, *India*, Hydrological Sciences, Journal des Sciences Hydrologiques, no.43(6), pp. 875-884.

Sarangi A., Madramootoo C.A., Enright P., Prasher S.O., (2006), *Evaluation Of Three Unit Hydrograph Models To Predict The Surface Runoff From A Canadian Watershed,* Water Resources Management, Vol. 21, No. 7, pp. 1127

Satheesh R., Sathikumar M., Madhu N., Kindersley Dorling, (2007), *Advanced Surveying: Total Station, GIS, Remote Sensing,* India.

Schwab G.O., Fangmeier D.D., Elliot W.J., Freveret R.K., (1993), *Soil, Water Conservation Engineering,* J. Wiley, Sons, New York.

Şerban G., (2004), *Lacurile de acumulare din Bazinul Superior al Someşului Mic*, Teză de doctorat, Universitatea Babeş-Bolyai, Cluj-Napoca.

Şerban P., (1995), *Modele Hidrologice Deterministe*, Editura Didactică şi Pedagogică, Bucureşti.

Şerban P., Stănescu A.V., Roman P., (1989), *Hidrologie dinamică*, Editura Tehnică Bucureşti.

Sherman, L.K., (1949). *The unit hydrograph method,* in: O. E. Meinzer, (ed.), *Physics of the Earth*, vol.IX, Hydrology, McGraw-Hill, New York.

Sherman, L.K., (1942), *Hydrograph of runoff,* in: O. E. Meinzer, (ed.), *Physics of the Earth*, vol.IX, Hydrology, McGraw-Hill, New York.

Silburn D.M., Connolly R.D., (1995), *Distributed Parameter Hydrology Model (Answers) Applied To A Range Of Catchment Scales Using Rainfall Simulator Data. I. Infiltration Modeling, Parameter Measurement*. Journal of Hydrology, No.172, pp. 87-104.

Singh P.V., Fiorentino M., (1996), *Geographical Information Systems in hydrology*, Kluwer Academic Publishers, USA

Smith Keith, Roy Ward, (1998), *Floods: Physical Processes, Human Impacts,* John Wiley & Sons Ltd, New York.

Smith R.E., Goodrich D.C, Woolhiser D.A., Unkrich C L., (1995), *KINEROS - A Kinematic Runoff, Erosion Model*, in: Singh V.P., (Ed.), Computer Models of Watershed Hydrology. Water Resource Publications, Highlands Ranch, CO., pp. 697-732.

Smith R.E, (1976), *Approximations for vertical infiltration*. Trans. ASAE., vol. 19, no. 3, pp. 505–509.

Snyder F.F., (1938), *Synthetic Unit Hydrographs,* Trans. Am. Geophys. Union, No.19, pp. 447-454.

Soil Conservation Service, (1986), *Urban Hydrology for Small Watersheds*, Technical Release 55 (TR-55).

Stein A., Riley J., Halberg N., (2001), *Issues Of Scale For Environmental Indicators. Agriculture*, Ecosystems & Environment,No. 87 (2), pp.215–232.

Stephenson G.R., Freeze R.A., (1974), *Mathematical Simulation Of Subsurface Flow Contributions To Snowmelt Runoff, Reynolds Creek, Idaho*, Water Resources Research, No.1012, pp. 284-298.

Strahler, A.N., (1957), *Quantitative Analysis Of Watershed Geomorphology*, Transactions of the American Geophysical Union, No.8 (6), pp.913–920.

Tarboton D.G., (1997), *New Method For The Determination Of Flow Directions, Upslope Areas In Grid Digital Elevation Models*, Water Resources Research, American Geophysical Union ,No.33(2), pp. 309-319.

Tarboton D.G., (2003), *Watershed Delineation from Digital Elevation Models*, Utah State University http://www.ce.utexas.edu/prof/maidment/giswr2003/ex4/Ex42003.htm, visited la 15 February 2010.

Tarboton D.G., Bras R.L., Rodriguez-Iturbe I., (1991), *On The Extraction Of Channel Networks From Digital Elevation Data*, Hydrological Processes ,No. 5 (1), pp.81–100.

Tatu G., (1995), *Sisteme hidraulice in regim tranzitoriu*, Universitatea Tehnică de Construcţii, București.

TCL/Tk History of Scripting, http://www.tcl.tk/doc/scriptHistory.html, visited 15 November 2010.

The LANDSAT program - http://landsat.gsfc.nasa.gov/

Thornton, P.E., Running S.W., (1999), *An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, precipitation*, Agriculture, Forest Meteorology, no.93,pp. 211-228.

Tosi S.**,** (2009), *Matplotlib for Python Developers*, Packt Publishing, Birmingham, UK.

TOPMODEL - Appendix A: Model Fact Sheets,
http://www.epa.gov/nrmrl/pubs/600r05149/600r05149topmodel.pdf
United States Department of Agriculture Natural Resources Conservation Service

Urbanek S., (2011), *proj4: A Simple Interface to the PROJ.4,* Cartographic Projections Library. R package version 1.0.7., http://www.rforge.net/proj4/, visited 4 March 2011.

USDA National Resources Conservation Service, (1972), *National Engineering Handbook, Part 630,* , Washington DC, http://www.mi.nrcs.usda.gov/technical/engineering/neh.html, visited 12 January 2011.

USDA-SCS, (1985), *National Engineering Handbook (NEH), Section 4 – Hydrology*, Washington, D.C., USDA-SCS.

Usul N., Yilmaz M., (2003), *Application of Clark Synthetic Unit Hydrograph Method with GIS for two Basins in Turkey*, International Conference of GIS, Remote Sensing in Hydrology, Water Resources, Environment (ICGRHWE), Three Gorges Dam, China.

Variable Infiltration Capacity (VIC) Macroscale Hydrologic Model documentation, http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Overview/ModelOverview.shtml, visited 18 May 2010.

Verbyla D., (2005), *Arcgis Geoprocessing Workshop*, Alaska Surveying Mapping Conference, GIS Jam, http://nrm.salrm.uaf.edu/~dverbyla/arcgis_geoprocessing/index.html, visited 26 August 2010

Vertessy R.A., Hatton T.J., O'Shaughnessy P., Jayasuriya M.D.A., (1993), *Predicting Water Yield From A Mountain Ash Forest Using A Terrain Analysis Based Catchment Model*, Journal of Hydrology, No.150**,** pp. 665-700.

Wade Tasha, Sommer Shelly (Eds.), (2006), *A To Z GIS: An Illustrated Dictionary Of Geographic Information Systems,* ESRI Press, USA.

Wise S.M., (2000), *Assessing The Quality For Hydrological Applications Of Digital Elevation Models Derived From Contours,* Hydrological Processes, No.14 (11–12), pp.1909–1929.

Wolock D.M., (1995), *Effects Of Subbasin Size On Topographic Characteristics, Simulated Flow Paths In Sleepers River Watershed, Vermont*, Water Resources Research, No.31(8), pp.1989-1997.

Wolock D.M., Hornberger G.M., Musgrove T.M., (1990), *Topographic Effects On Flow Path Length, Surface Water Chemistry Of The Llyn Brianne Catchments In Wales*, Journal of

Hydrology, No.115, pp. 243-259.

Woolhiser D.A., Hanson C.L., Kuhlman A.R., (1970), *Overland Flow On Rangeland Watersheds*, Journal of Hydrology (New Zealand), No.9(2), pp. 336-356.

Woolhiser D.A., Smith R.E., Goodrich D.C., (1990), *KINEROS, A Kinematic Runoff, Erosion Model: Documentation, User Manual,* U S. Department of Agriculture, Agricultural Research Service, ARS-77.

Xi Chen, Yongqin David Chen, Chong-yu Xu, (2007), *A distributed monthly hydrological model for integrating spatial variations of basin topography, rainfall,* Hydrol. Process. 21, pp. 242–252

Xiaoyong Zhan, Min-Lang, (2004), *Arccn-Runoff: An Arcgis Tool For Generating Curve Number, Runoff Maps*, Environmental Modeling & Software, Vol. 19, Issue 10, pp. 875-879.

Xing Fang, Thompson B.D., Cleveland G.T., Pratistha Pradhan, (2007), *Variations of Time of Concentration Estimates Using NRCS Velocity Method*, Journal of irrigation, drainage engineering, pp. 314-322.

Zăvoianu Ion, (1978), *Morfometria Bazinelor Hidrografice,* Editura Academiei Republicii Socialiste România, Bucureşti.

Zeiler M., (2001), *Modeling Our World: The ESRI Guide to Geodatabase Design,* ESRI Press, USA.

Zhang L., Dawes W.R., Hatton T.J., Reece P.H., Beale G.T.H., Packer Z., (1999), *Estimation of soil moisture, groundwater recharge using the TOPOG-IRM model*, Water Resources Research, no.35, pp. 149-161.

Zhao R., Zuang Y., Fang L., Liu X., Zhang Q., (1980), *The Xinanjiang Model, Hydrological Forecasting,* Proceedings of the Oxford Symposium, Actes du Colloque d'Oxford No.129.

Zoch R.T., (1934), *On The Relation Between Rainfall and Streamflow*, Monthly Weather Review, No. 62, (91), pp.315-322.

Glossary

Algorithm (fr. *algorithme*) - A mathematical procedure used to solve problems with a series of steps. Algorithms are usually encoded as a sequence of computer commands.

AMC - indicator for characterizing anterior conditions of terrain humidity on the basis of the last five days precipitations. It contains three categories: AMC I(dry), AMC II (normal), AMC III (saturated).

Cumulative infiltration (fr. *infiltration cumulative*) – total quantity of water infiltrated in soil after rainfall event.

Curve Number (fr. *Numéro de la Courbe*) – indicator used in SCS-CN model, for characterizing infiltration/ runoff capacity, according to soil hydrological group and the terrain usage.

DEM (Digital Elevation Model) - The representation of continuous elevation values over a topographic surface by a regular array of z-values, referenced to a common datum. DEMs are typically used to represent terrain relief.

Evapotranspiration (fr. *évapotranspiration*) – quantity of water resulted through evaporation on soil surface and through plan transpiration.

Excess rainfall (fr. *précipitation nete*) - the quantity of water available for runoff by excelling the infiltration capacity

Gradient (fr. *gradient*) – The ratio between vertical distance (rise) and horizontal distance (run), often expressed as a percentage. A 10-percent gradient rises 10 feet for every 100 feet of horizontal distance.

HSG: Hydrologic Soil Group (fr. *Grupe Hydrologique du Sol*) – classification indicator according to the infiltration capacity, which is deduced on the basis of soil texture.

Infiltration capacity (fr. *capacité d'infiltratio*) – a characteristic of the soil referring to its maximum potential of water penetration.

Infiltration rate (fr. *taux d'infiltration*) – quantity of water infiltrated in soil in time unit (mm/min, mm/h).

Initial abstractions (fr. *pertes au début du précipitation*) – parameter of SCS method referring to the quantity of water initially retained by vegetation, and lost through evaporation or other losses.

Interception capacity (fr. *capacité d'interception*) – the capacity of vegetation to catch and retain a part of the precipitations, a quantity that is afterwards subject to evapotranspiration and does not reach the soil.

Lag time (fr. *le temps montée de la crue*) – time between the center of the rain and the flash flood hydrograph peak.

Layer (fr. *couche*) –a reference to a data source, such as a shapefile, coverage, geodatabase feature class, or raster.

Map Algebra – A language that defines a syntax for combining map themes by applying mathematical operations and analytical functions to create new map themes. In a map algebra expression, the operators are a combination of mathematical, logical, or Boolean operators (+, >, AND, tan, and so on), and spatial analysis functions (slope, shortest path, spline, and so on), and the operands are spatial data and numbers.

ModelBuilder – The interface used to build and edit geoprocessing models in ArcGIS.

Overlay – A spatial operation in which two or more maps or layers registered to a common coordinate system are superimposed, either digitally or on a transparent material, for the purpose of showing the relationships between features that occupy the same geographic space.

Percolation (fr. *percolation*) – The movement of water through the openings in rock or soil.

Potential maximum retention (fr. *le potentiel maximale de retention*) – the maximum quantity of water that can be retained by soil until it reaches saturation.

Raster – A spatial data model that defines space as an array of equally sized cells arranged in rows and columns, and composed of single or multiple bands. Each cell contains an attribute value and location coordinates. Unlike a vector structure, which stores coordinates explicitly, raster coordinates are contained in the ordering of the matrix. Groups of cells that share the same value represent the same type of geographic feature.

Soil moisture, soil wetness, soil water content (fr. *humidité du sol*) – quantity of water from the soil that is above the groundwater level.

Spatial analysis (fr. *analyse spatiale*) – The process of examining the locations, attributes, and relationships of features in spatial data through overlay and other analytical techniques in order to address a question or gain useful knowledge. Spatial analysis extracts or creates new information from spatial data.

Spatial resolution (fr. *résolution spatiale*) – The detail with which a map depicts the location and shape of geographic features. The larger the map scale, the higher the possible resolution. As scale decreases, resolution diminishes and feature boundaries must be smoothed, simplified, or not shown at all; for example, small areas may have to be represented as points.

Subsurface flow, throughflow (fr. *ruisselement hypodermique*) – water runoff on the first centimeters of soil profile, according to terrain slope.

Thematic map (fr. *carte thématique*) – A map designed to convey information about a single topic or theme, such as population density or geology.

Time of concentration (fr. *temps de concentration*) – the necessary time for a water particle from the furthest part of the basin to reach the outlet.

TIN (Triangulated Irregular Network) –A vector data structure that partitions geographic space into contiguous, nonoverlapping triangles. The vertices of each triangle are sample data points with x-, y-, and z-values. These sample points are connected by lines to form Delaunay triangles. TINs are used to store and display surface models.

Travel time – the time a water particle needs to reach from a point of the basin to the outlet.

Vector (fr. *vecteur*) – A coordinate-based data model that represents geographic features as points, lines, and polygons. Each point feature is represented as a single coordinate pair, while line and polygon features are represented as ordered lists of vertices. Attributes are associated with each vector feature, as opposed to a raster data model, which associates attributes with grid cells.

Annexes

Annex 1: CORINE Land Cover land use, Hydrologic Soil Groups and Curve Number

HSG	CLCCODE	LANDUSE LABEL	CN
A	244	Agro-forestry areas	52
B	244	Agro-forestry areas	69
C	244	Agro-forestry areas	79
D	244	Agro-forestry areas	84
A	124	Airports	80
B	124	Airports	85
C	124	Airports	88
D	124	Airports	93
A	241	Annual crops associated with permanent crops	64
B	241	Annual crops associated with permanent crops	75
C	241	Annual crops associated with permanent crops	82
D	241	Annual crops associated with permanent crops	85
A	332	Bare rocks	77
B	332	Bare rocks	86
C	332	Bare rocks	91
D	332	Bare rocks	94
A	331	Beaches, dunes, sands	63
B	331	Beaches, dunes, sands	77
C	331	Beaches, dunes, sands	85
D	331	Beaches, dunes, sands	88
A	311	Broad-leaved forest	42
B	311	Broad-leaved forest	66
C	311	Broad-leaved forest	79
D	311	Broad-leaved forest	85
A	334	Burnt areas	77
B	334	Burnt areas	86
C	334	Burnt areas	91
D	334	Burnt areas	94
A	521	Coastal lagoons	0
B	521	Coastal lagoons	0
C	521	Coastal lagoons	0
D	521	Coastal lagoons	0
A	242	Complex cultivation patterns	67
B	242	Complex cultivation patterns	78
C	242	Complex cultivation patterns	85
D	242	Complex cultivation patterns	89
A	312	Coniferous forest	34
B	312	Coniferous forest	60
C	312	Coniferous forest	73
D	312	Coniferous forest	79
A	133	Construction sites	80
B	133	Construction sites	85
C	133	Construction sites	88
D	133	Construction sites	93
A	111	Continuous urban fabric	85
B	111	Continuous urban fabric	89

C	111	Continuous urban fabric	92
D	111	Continuous urban fabric	98
A	112	Discontinuous urban fabric	77
B	112	Discontinuous urban fabric	85
C	112	Discontinuous urban fabric	90
D	112	Discontinuous urban fabric	95
A	132	Dump sites	80
B	132	Dump sites	85
C	132	Dump sites	88
D	132	Dump sites	93
A	522	Estuaries	0
B	522	Estuaries	0
C	522	Estuaries	0
D	522	Estuaries	0
A	222	Fruit trees and berry plantations	43
B	222	Fruit trees and berry plantations	65
C	222	Fruit trees and berry plantations	76
D	222	Fruit trees and berry plantations	82
A	335	Glaciers and perpetual snow	0
B	335	Glaciers and perpetual snow	0
C	335	Glaciers and perpetual snow	0
D	335	Glaciers and perpetual snow	0
A	141	Green urban areas	48
B	141	Green urban areas	66
C	141	Green urban areas	76
D	141	Green urban areas	82
A	121	Industrial or commercial units	81
B	121	Industrial or commercial units	88
C	121	Industrial or commercial units	91
D	121	Industrial or commercial units	93
A	411	Inland marshes	0
B	411	Inland marshes	0
C	411	Inland marshes	0
D	411	Inland marshes	0
A	423	Intertidal flats	95
B	423	Intertidal flats	95
C	423	Intertidal flats	95
D	423	Intertidal flats	95
A	243	Land principally occupied by agriculture, with significant areas of natural vegetation	52
B	243	Land principally occupied by agriculture, with significant areas of natural vegetation	69
C	243	Land principally occupied by agriculture, with significant areas of natural vegetation	79
D	243	Land principally occupied by agriculture, with significant areas of natural vegetation	84
A	131	Mineral extraction sites	80
B	131	Mineral extraction sites	85
C	131	Mineral extraction sites	88
D	131	Mineral extraction sites	93
A	313	Mixed forest	38
B	313	Mixed forest	62

C	313	Mixed forest	75
D	313	Mixed forest	81
A	322	Moors and heathland	49
B	322	Moors and heathland	69
C	322	Moors and heathland	79
D	322	Moors and heathland	84
A	321	Natural grasslands	49
B	321	Natural grasslands	69
C	321	Natural grasslands	79
D	321	Natural grasslands	84
A	211	Non-irrigated arable land	67
B	211	Non-irrigated arable land	78
C	211	Non-irrigated arable land	85
D	211	Non-irrigated arable land	89
A	223	Olive grives	55
B	223	Olive grives	75
C	223	Olive grives	82
D	223	Olive grives	89
A	231	Pastures	49
B	231	Pastures	69
C	231	Pastures	79
D	231	Pastures	84
A	412	Peat bogs	30
B	412	Peat bogs	58
C	412	Peat bogs	71
D	412	Peat bogs	78
A	212	Permanently irrigated land	67
B	212	Permanently irrigated land	78
C	212	Permanently irrigated land	85
D	212	Permanently irrigated land	89
A	123	Port areas	0
B	123	Port areas	0
C	123	Port areas	0
D	123	Port areas	0
A	213	Rice fields	67
B	213	Rice fields	78
C	213	Rice fields	85
D	213	Rice fields	89
A	122	Road and rail networks and associated land	83
B	122	Road and rail networks and associated land	89
C	122	Road and rail networks and associated land	92
D	122	Road and rail networks and associated land	93
A	422	Salines	0
B	422	Salines	0
C	422	Salines	0
D	422	Salines	0
A	421	Salt marshes	74
B	421	Salt marshes	84
C	421	Salt marshes	90
D	421	Salt marshes	92
A	323	Sclerophyllous vegetation	45
B	323	Sclerophyllous vegetation	60

C	323	Sclerophyllous vegetation	73
D	323	Sclerophyllous vegetation	78
A	523	Sea and ocean	0
B	523	Sea and ocean	0
C	523	Sea and ocean	0
D	523	Sea and ocean	0
A	333	Sparsely vegetated areas	72
B	333	Sparsely vegetated areas	82
C	333	Sparsely vegetated areas	83
D	333	Sparsely vegetated areas	87
A	142	Sport and leisure facilities	51
B	142	Sport and leisure facilities	68
C	142	Sport and leisure facilities	79
D	142	Sport and leisure facilities	84
A	324	Transitional woodland-shrub	45
B	324	Transitional woodland-shrub	60
C	324	Transitional woodland-shrub	73
D	324	Transitional woodland-shrub	78
A	221	Vineyards	46
B	221	Vineyards	67
C	221	Vineyards	78
D	221	Vineyards	83
A	512	Water bodies	0
B	512	Water bodies	0
C	512	Water bodies	0
D	512	Water bodies	0
A	511	Water courses	100
B	511	Water courses	100
C	511	Water courses	100
D	511	Water courses	100

Annex 2 : Mannings n calculation table from CLC landuse layer

CLCCODE	LANDUSE LABEL	MANNING n
111	Continuous urban fabric	0.01
112	Discontinuous urban fabric	0.01
121	Industrial or commercial units	0.01
122	Road and rail networks and associated land	0.01
123	Port areas	0.01
124	Airports	0.01
131	Mineral extraction sites	0.01
132	Dump sites	0.01
133	Construction sites	0.01
141	Green urban areas	0.6
142	Sport and leisure facilities	0.6
211	Non-irrigated arable land	0.17
212	Permanently irrigated land	0.17
213	Rice fields	0.17
221	Vineyards	0.17
222	Fruit trees and berry plantations	0.17
223	Olive grives	0.17
231	Pastures	0.24
241	Annual crops associated with permanent crops	0.17
242	Complex cultivation patterns	0.17
243	Land principally occupied by agriculture, with significant areas of natural vegetation	0.17
244	Agro-forestry areas	0.17
311	Broad-leaved forest	0.15
312	Coniferous forest	0.15
313	Mixed forest	0.15
321	Natural grasslands	0.6
322	Moors and heathland	0.6
323	Sclerophyllous vegetation	0.6
324	Transitional woodland-shrub	0.6
331	Beaches, dunes, sands	0.01
332	Bare rocks	0.01
333	Sparsely vegetated areas	0.01
334	Burnt areas	0.01
335	Glaciers and perpetual snow	0.01
411	Inland marshes	0.01
412	Peat bogs	0.01
421	Salt marshes	0.01
422	Salines	0.01
423	Intertidal flats	0.01
511	Water courses	0.01
512	Water bodies	0.01
521	Coastal lagoons	0.01
522	Estuaries	0.01
523	Sea and ocean	0.01