UDC: 556.09:528.9]:004.6

DEM-Based GIS Algorithms for Automatic Creation of Hydrological Models Data

D. Prodanović[1]*, M. Stanić[2], V. Milivojević[3], Z. Simić[4], M. Arsić[5]

Faculty of Civil Engineering, University of Belgrade, 73 Bulevar Kralja Aleksandra St., 11000 Belgrade, Serbia; e-mail: [1]eprodano@hikom.grf.bg.ac.rs, [2]mstanic@grf.bg.ac.rs
Institute for Development of Water Resources "Jaroslav Černi", 80 Jaroslava Černog St., 11226 Beli Potok, Serbia; e-mail: [3]vladimir.milivojevic@gmail.com, [4]zoran.simic@jcerni.co.rs, [5]miomir.arsic@jcerni.co.rs
Corresponding author

Abstract

In this paper the authors discuss the Digital Elevation Model (DEM) based GIS algorithms, applicable for automatic creation of input data needed by hydrological models. The expected source of DEM data are standard digital data sets that a modeler can purchase from the official data sources. The DEM has to be cleared from errors and prepared for standard hydrology usage as the depressionless version, or a version of DEM where only small depressions are filled-in, while larger depressions are extracted as separate objects. The slopes and aspects are calculated from such DEM. The flow accumulation image is created using slopes and aspects and using information about large ponds. From flow accumulation image the surface drainage network is extracted with user-specified density. Based on slopes and aspects, the catchment is delineated into subcatchments, the areas drained by each river segment. Finally, the model-specific parameters like length of flow segment along the terrain and along the river to the subcatchment outlet, slope of those segments, subcatchment shape factor, or subcatchments mean or weighted slope are calculated. Paper explains the used algorithms and emphasizes the problems one can encounter during the usage of DEM data using examples from River Drina catchment. Finally, the authors comment on overall usability of presented GIS algorithms, especially if purchased data are of low quality.

Keywords: Hydrologic models, data preparation, DEM

1. Introduction

Hydrologic models are simplified, conceptual representations of a part of the hydrologic cycle. They are primarily used for hydrologic prediction and for understanding of hydrologic processes. Two major types of hydrologic models can be distinguished: a) stochastic (or black box) models, which use mathematical and statistical concepts to link a certain input (for instance rainfall) to the model output (for instance runoff), and b) process-based (or physically based or deterministic) models, which try to describe by equations the physical processes involved, such as snow melting, evapotranspiration, surface runoff, subsurface flow and channel flow. Physically based hydrologic models can be subdivided into single-event models and continuous simulation models. Over the past several decades, the physically based modeling has become a standard method for investigation and study of almost all kinds of catchment-related hydrological problems: the effect of catchment changes, prediction of ungauged catchment behavior, spatial variability in catchment inputs and outputs, movement of pollutants and sediment etc. By using hydrologic modeling, the researcher tries to develop a more global approach to the understanding of the behavior of hydrologic systems, in order to make better predictions and to face the major challenges in water resources management.

The essential component of each physically based model is the availability of data about the system itself. For hydrology models, this means that the detailed knowledge of land cover and vegetation types, land topology, underground soil types and water levels is needed. For distributed models, the spatial and temporal variability of all those data is needed. The possibility to rapidly combine large data volumes of different types in a GIS (Geographic Information System) has led to a significant increase in its use in hydrological applications. It also provides the opportunity to combine data from different sources and of different types (Elgy et al. 1993). Some of the typical applications of this approach is the usage of remotely sensed images for extraction of the terrain canopy data or the use of a digital terrain model (DTM), or digital elevation model (DEM) for extraction of hydrologic catchment properties, such as elevation matrix, flow direction matrix, ranked elevation matrix and flow accumulation matrix (Jenson and Domingue (1988), Prodanović et al. 1998).

There are several widely accepted physically based hydrological models on the market today. Some of them are: HYSIM (HYdrological SImulation Model) by WRA (2009), HSPF (Hydrologic Simulation Program Fortran) by U.S. EPA (2009), MIKE SHE by DHI (2009), SWAT (Soil and Water Assessment Tool) by USDA-ARS (2009) and WMS (Watershed Modeling System) by Aquaveo (2009).

In this paper, the usage of GIS routines, tailored for use with the SWAT hydrological model, is presented. The SWAT model delineates the catchment area into the set of the HRUs (Hydrology Response Units) with homogenous land, soil and vegetation parameters. Detailed description of the SWAT model itself (Arnold et al. 2002) and specific upgrades of the SWAT model are given in the paper (Simić et al. 2009).

The purpose of routines described in this paper is to automatically prepare the DEM-related parameters for each HRU and link HRUs to the downstream HP (Hydrology Profile, or catchment), where water balance is made. Firstly, in this paper the availability of DEMs will be analyzed, along with the necessary resolution and accuracy. From DEM, a number of space-related derivatives are created and used: slopes, aspects, ponds, surface flow paths etc. The paper will describe in detail those algorithms and will present some results based on the River Drina catchment area. Also, the methods of HRU creation and spatial parameter averaging within HRU areas will be discussed.

For the sake of brevity, the other types of data needed by SWAT, which could be also prepared using GIS routines, are not covered by this paper. This includes the GIS operations with the linear features, such as rivers, and point features (meteorological stations mostly) with the everlasting problem of optimal spatial extrapolation. Of course, both features strongly interact with the DEM: the rivers can be extracted from the DEM and since the rainfall intensity depends on the altitude, the spatial interpolation has to be DEM driven.

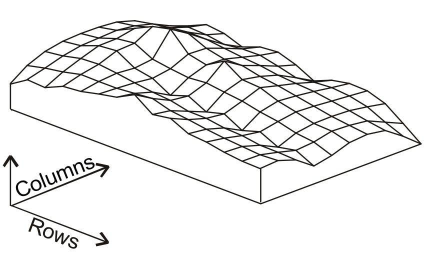
Throughout the paper, the data accuracy and usability will be assessed, especially if data are purchased in digital format, with the intention to “use them as they are”, i.e. without further processing (Prodanović et al. 2006). Different data users need different kinds of terrain elevation interpretation in digital form (or digital terrain representation). For hydrology, for example, the absolute elevation is not so important, but the slope and the direction of slope are very important. Also, river bifurcation in flat areas is important for cartographers, but for simple hydraulic river flow models, as the ones used in hydrology models, parallel river flow paths have to be divided into the major, connected, flow path and the minor, disconnected one.

2. Spatial data - DEM

Performance and reliability of hydrology models are highly dependent on the quality of terrain elevation representation, in terms of accuracy and resolution. Most physical processes (such as surface flow and surface retention, temperature and precipitation distribution and land cover distribution) depend on terrain slopes, sun exposure and absolute terrain height. The methodology of DEM creation is too complex and is out of the scope of this paper. It will be assumed that DEM of sufficient resolution and accuracy is available in raster format, with regular grid size.

2.1. DEM presentation

For representation of the terrain heights in digital form, a number of techniques can be used (Burrough 1993): isoheight lines, spot heights, regular grid and TIN – Triangular Irregular Network. The pros and cons of different data formats usage in hydrology are discussed in (Prodanović 1999). The regular grid (Figure 1), the matrix of equally spaced Z values, can be considered as an optimal solution, since, due to the fact that simple data storage scheme for the spatial analyses is easy to define, most contemporary data capturing techniques use that format and the fact that the amount of the stored data is acceptable. However, one of the main drawbacks of this approach is that spatial resolution has to be defined in advance because no sub-grid height data can be obtained (Garbrecht and Martz 1996).



Rows \ Columns	0	1	2	3	4	5	6	7	8	9
0	50.0	62.5	70.7	74.3	75.0	75.1	73.3	67.7	55.7	45.0
1	62.7	67.4	73.5	77.0	77.7	78.6	76.6	71.0	62.0	55.0
2	60.9	70.9	77.5	84.3	84.5	88.2	82.0	74.7	67.9	63.4
3	70.0	75.3	78.9	86.1	86.9	97.5	86.6	76.7	67.5	61.5
4	71.6	75.4	76.8	75.4	77.3	99.9	89.8	76.6	64.7	55.0
5	76.3	75.0	73.7	71.8	74.5	83.2	81.8	74.9	67.2	61.6
6	79.8	76.8	71.2	61.0	72.5	79.0	77.5	72.9	67.6	63.8
7	90.3	79.1	70.1	67.3	74.4	90.0	78.0	70.9	63.5	57.3
8	76.5	73.4	65.8	58.2	71.7	77.5	74.3	68.6	58.8	48.0
9	80.0	71.5	65.6	63.2	69.7	76.2	71.1	67.1	61.2	56.0
10	70.6	68.2	62.8	53.0	65.0	69.6	67.7	65.4	62.6	60.4
11	69.3	66.6	61.9	58.6	62.2	65.6	64.5	62.9	60.9	59.1
12	65.7	64.8	59.3	50.0	57.9	60.4	60.8	59.7	56.5	53.3
13	63.3	63.1	58.9	53.8	55.5	59.7	58.5	56.2	48.0	39.3
14	62.7	64.7	58.2	44.3	48.7	55.0	56.5	53.7	40.2	25.0

Fig. 1. DEM representation using regular grid

As an example of raster DEM, Figure 2 presents the DEM of the Drina catchment. On the left-hand side of the figure, the cartographic representation of River Drina and its tributaries is given, while on the right-hand side of figure the River Drina catchment's DEM is plotted by using colour coded elevations. DEM originated from the Serbian National Military Geodesy Department (NMGD): contour lines and spot heights were digitized from 1:300000 maps and processed using TIN and regular grid generated with the 100x100 m grid size.

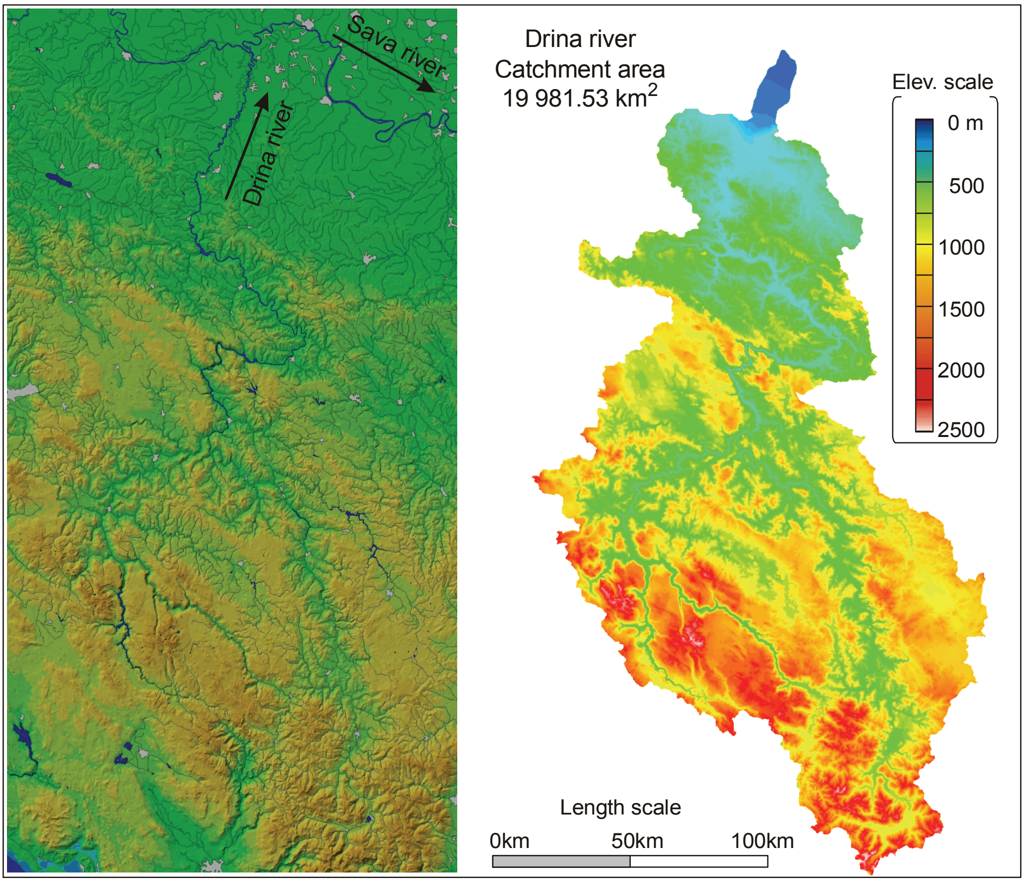


Fig. 2. River Drina and the DEM of its catchment, pixel size 100x100 m (Jaroslav Černi Institute 2005)

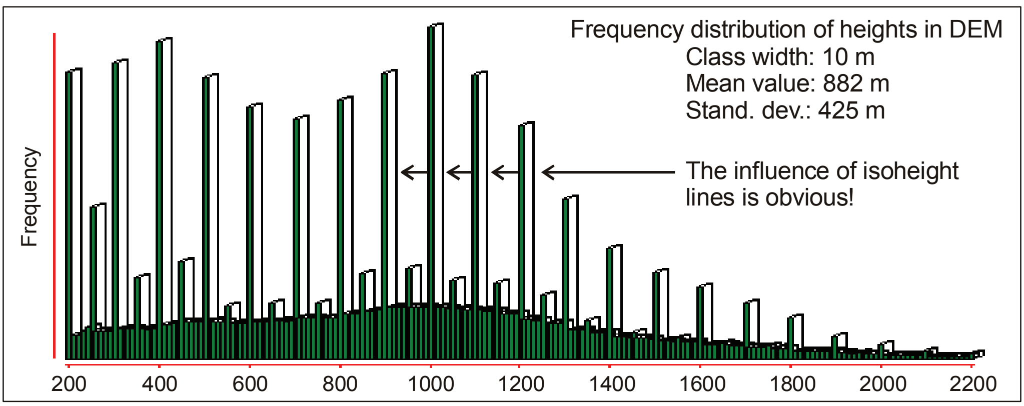


Fig. 3. Frequency distribution of heights in Drina catchment DEM

For hydrological application the DEM has two main usages: as a source of height data for modelling of processes that are strongly altitude-correlated, and for simulation of the surface water flow and computation of subcatchments, flow paths, ponds etc. For the second DEM usage, data imperfections contained in the DEM may directly compromise the results of such analysis, compromising the hydrology model. So it is important to have accurate DEM, with as little noise as possible (or, to have “hydrologicaly correct” DEM).

Figure 3 presents a common problem that can be found in commercial DEMs, created from paper maps. Since isoheight lines were used for DEM creation, they will largely influence the final product and will dominate in altitude (height) frequency distribution curve, i.e. the DEM will have a large number of pixels with constant Z values, and therefore a large number of flat regions. Since the quality of DEM mostly depends on the source of data, detailed DEM analysis and pre-processing are often required. The best approach is to produce a custom-tailored DEM with a pre-specified resolution (Garbrecht and Martz 2000). However, in the majority of cases, this solution is cost prohibitive, and a custom off-the-shelf DEM data set that is already available has to be used. Therefore, the usual procedure is to correct the existing DEM, as described below.

2.2. Slopes and aspects from DEM

One of the main parameters of all physically-based hydrological models is the slope of terrain and direction of the steepest slope. In general, with DEM in the raster format, it is quite simple to compute the first derivatives and determine the slope and the aspect (direction of maximal slope) of the terrain.

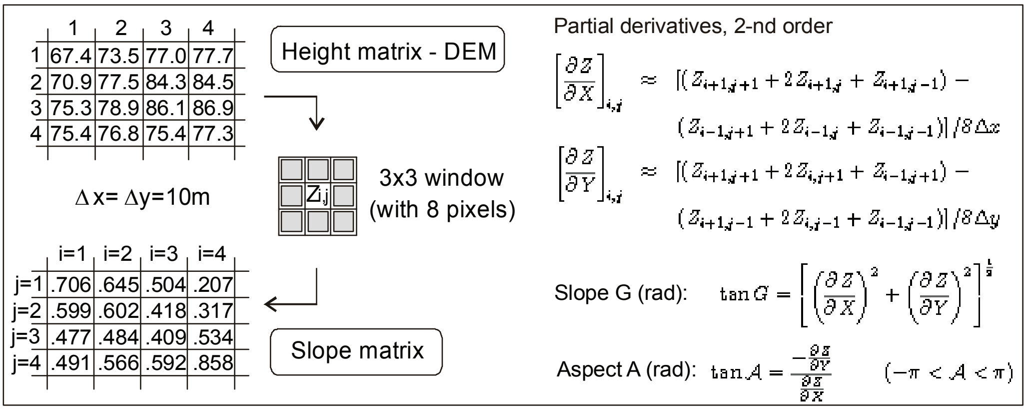


Fig. 4. Slope and aspect calculation from DEM – an example of second order derivative

There are numerous algorithms for slope and aspect calculation from raster DEM. All algorithms are “window-based”, i.e. the slope and aspect of the current cell are based upon the elevations of the surrounding cells. The simplest algorithm is D8, introduced by O'Callaghan and Mark (1984), where flow is transferred toward the lowest surrounding pixel, so that aspect values have a fixed 45^0 resolution. Somewhat better approach is D∞ (or DInf) used by Tarboton D.G. (1999) where aspect is a continuous value, but is not calculated by using all surrounding pixels. Burrough (1993) suggest calculation of slope and aspect by the process of fitting the flat surface through either 4 neighboring pixels in a 3x3 window (1-st order derivative), or 8 pixels (2-nd order derivative, as presented in Figure 4). As a general rule, higher accuracy slope calculation methods are more susceptible to errors and noise in DEM.

After the slopes are calculated, the accuracy and quality of used DEM can be better assessed. All pixels equal to zero can be plotted by simple reclassification of slope image, i.e.

all horizontal areas can be easily visualized. Within DEM it is assumed that horizontal areas are found only inside some lakes. However, in most cases, the result will be similar to the one obtained from DEM of River Drina catchment, Figure 5. Black areas of the central image represent the horizontal areas of the catchment. A closer look at the selected part of the catchment (left-hand image in Figure 5) presents the source of the problem: the slope around isoheight lines is not equal zero, and within the areas bounded by the same isoheight line, the height is constant, and the slope is zero.

The right-hand part of the Figure 5 gives the insight into the slope distribution. In almost 23% of the catchment area the slope is equal to zero. Also, there are some pixels with unusually high slopes, indicating the possible erroneous height values in DEM (often the case of mis-interpretation or mis-typing of height values during spot height digitalization).

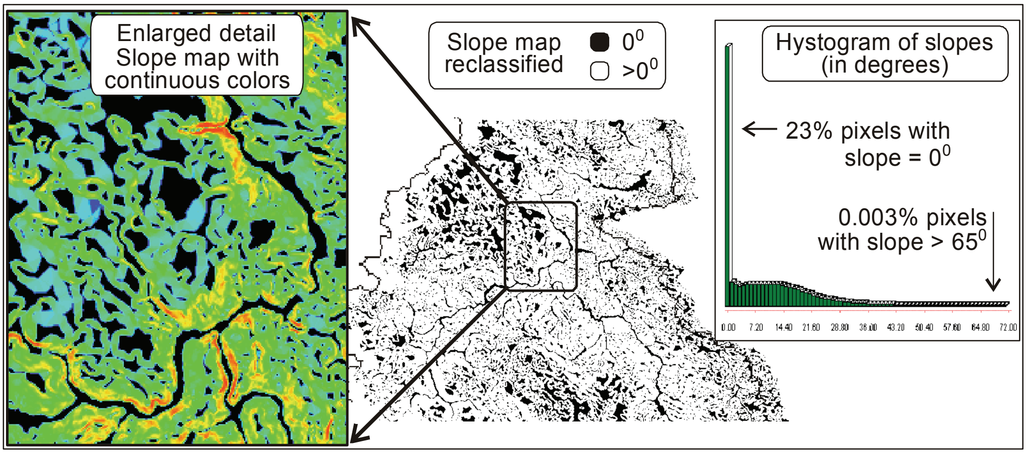


Fig. 5. Slope of DEM reclassified to flat (black areas) and non-flat areas (central part of figure) for River Drina catchment, detailed view of continuous slopes (left-hand part of figure) and frequency distribution of slopes (more than 23% of the catchment area has flat pixels)

2.3. Correction of horizontal areas in raster DEM

Large number of horizontal areas that can be found in raster DEM is directly related to the DEM creation phase. If raster DEM is produced from TIN, and TIN is created by digitizing the isolines (either from old paper maps or from orthophotos), the large number of points along one iso-line (known as “over-digitizing” practice) will tend to make triangles where all three points are within the same iso-line. This is presented on the left-hand side of Figure 6, where isoheight lines are densely digitized processed by TIN. Large number of triangles are created, with points laying on the same isoheight line (marked with circles on the figure), resulting in horizontal areas.

Right-hand part of Figure 6 presents the simple solution to the problem of flat areas creation. The breakline has to be digitized first, as it lies along the ridge or streams. That breakline will prevent the creation of horizontal triangles. Also, number of digitized points has to be much smaller to keep the triangles as regular as possible.

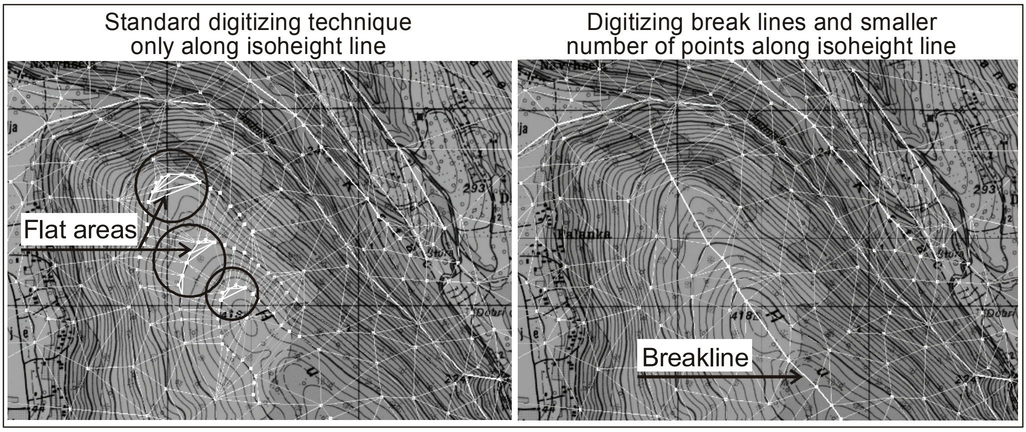


Fig. 6. DEM creation phase: Difference between terrain digitization without (image on the left) and with break lines (image on the right) that prevent formation of horizontal areas

In order to have DEM as accurate as possible, with good slope and aspect presentation, during the process of data capturing a care should be taken to have appropriate point density and all major breaklines. However, the explanation of the source of the problem (Figure 6) is useless since the user is faced with a purchased, ready-made DEM, in raster format. There is no possibility to go one step back, improve the TIN model using break lines and produce better DEM. Since automatic delineation of surface flow paths and subcatchments is largely influenced by aspect data, in such situations the correction of raster DEM, where no information about terrain features exist, is needed.

In order to correct a large number of horizontal areas in raster DEM, the standard procedure is to use the low-pass filter. Such filter can be applied with different kernel sizes, repetitively until all flat regions are filtered-out. However, this kind of filtering has one drawback – although it will “smooth” the surface and make it “curved”, it will also reduce the heights of terrain peaks and increase the heights of the valleys.

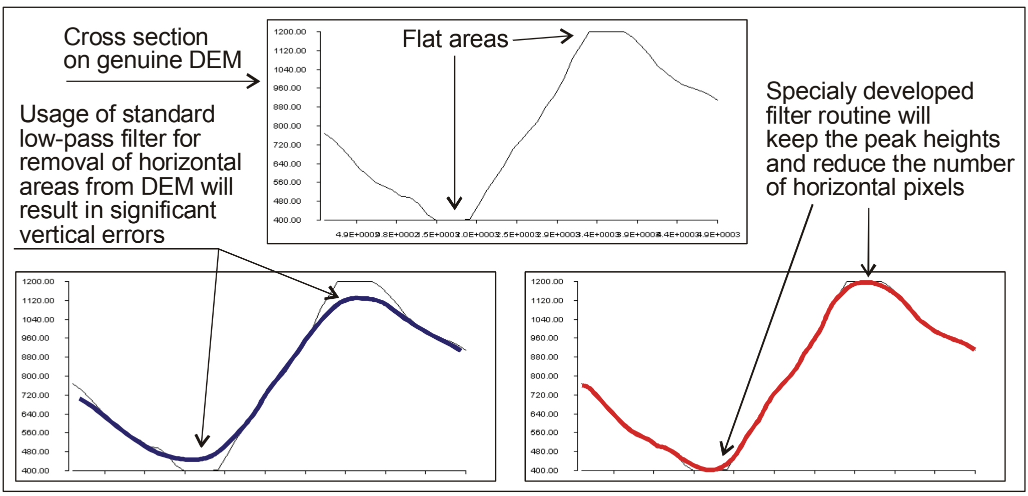


Fig. 7. Sample terrain cross section with errors - flat areas (upper image) after application of low-pass filter (lower left image) and modified conditional low-pass filter (lower right image)

Figure 7 illustrates the effect of heavy low-pass filtering of DEM: in the genuine DEM flat areas will be curved, but small variations in natural slope will also be smoothed-out and height of peaks and valleys will be changed. A much better approach is to apply the low-pass filter only on flat pixels which are located at the border of larger flat regions. Such routine was developed by Prodanović (Stanić et al. 2004) and successfully used in GIS based model for surface flooding (Boonya-aroonnet et al. 2007), (Leitão et al. 2008), (Maksimović et al. 2004) and (Maksimović et al. 2009) and naimed “conditional filtering”. Routine runs iteratively: during each iteration the flat areas are delineated and low-pass filter applied around the flat area’s border. In each iteration the edges of flat areas are curved in a way that changed pixels (or curved terrain) hold the genuine aspect of surrounding terrain. Lower right part of the Figure 7 presents the result of conditional filtering.

3. Raster DEM derivatives needed for hydrology model

3.1. Ponds and depressions from DEM

One of the most exciting uses of DEM in hydrology is the possibility to “roll a ball” over the terrain (using aspect and slope data) and to record the natural surface flow paths. This algorithm is often referred to as “surface flow tracing algorithm”. It will determine the natural flow paths represented through the branched structure of streams and rivers, thus it can delineate the catchment area (Đordjević et al. 1999). If the genuine DEM has depressions (or ponds) in it, they will prevent the flow tracing algorithm from reaching the end of catchment, flow path will be “trapped” by the lowest point in the depression, the “pit cell”.

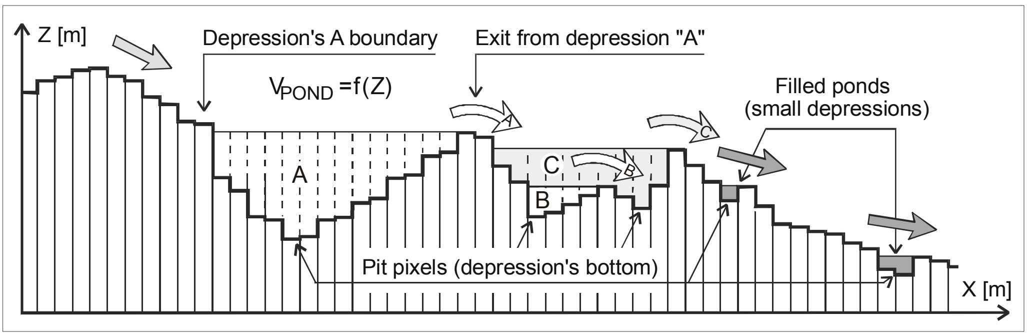


Fig. 8. Terrain cross-section with depressions – large, small and overlapping

Unfortunately, almost every DEM has pit cells in it (Figure 8). They can be a consequence of the existence of the natural depression which was correctly entered into the DEM (depressions A, B and C in Figure 8), or, more often, they can be the result of wrong DEM interpolation. In the latter case the area of the resulting pond will be in most cases small and should be filled (DEM pixels should be raised) in order to allow for uninterrupted surface flow.

However, if true depressions exist in the DEM, there is a possibility to fill those depressions also, thus creating the “depressionless DEM” (Band 1986). This approach is often used by commercial GIS packages, but if some depressions are left in the DEM on purpose, this method will mask them, although they represent an important hydrological feature. Another approach can be to delineate the boundary of the depression, find the exit point from the depression and calculate the stage (elevation)-volume curve and stage-area curve (Prodanović 1999). The data can then be used in surface flow modeling, in order to simulate the depression’s storage capability. The surface flow routine can then be used to create the link between the depressions, where the care should be taken about so-called “covered” depressions (depressions B and C in Figure 8).

The pond delineation algorithm that works with raster DEM image consists of several steps. First, a list of all pit pixels in DEM is created and sorted depending to their height (Figure 9, upper right-hand corner). The pixel represents a pit if it has the lowest Z value in its 3x3 window, i.e. within 8 surrounding neighbors. The list of sorted pit pixels is then used to delineate ponds, starting from the lowest one.

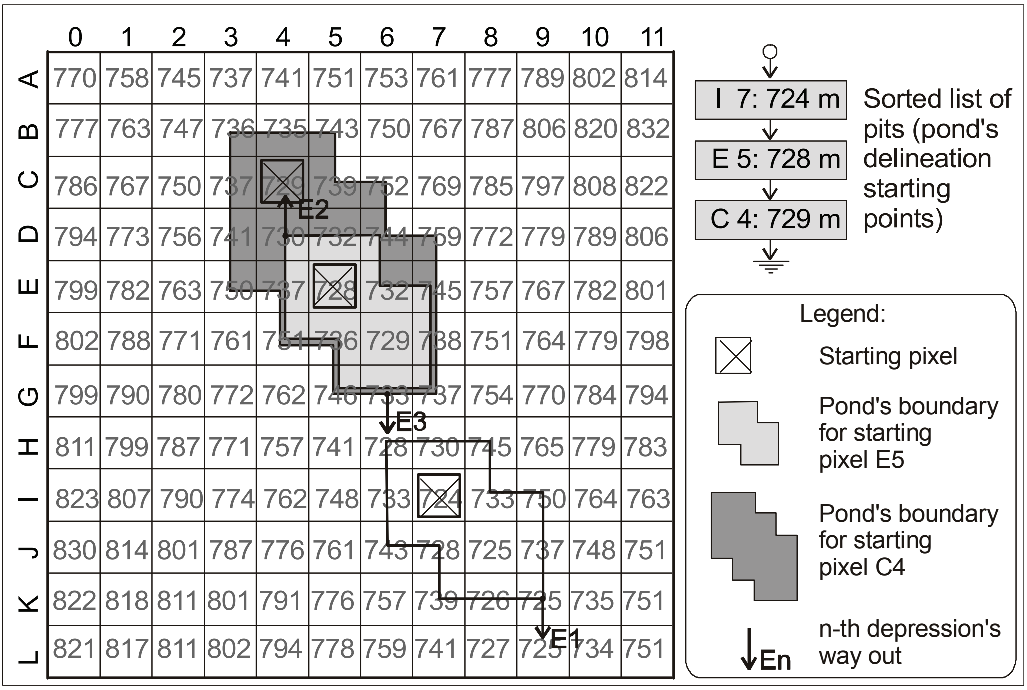


Fig. 9. Delineation of ponds – an example DEM with three ponds

For each pit from the pit list, iteratively two lists are created: the list of pond’s pixels and the list of boundary pixels (during the first iteration, the pond list holds the pit pixel only, and the boundary list holds the 8 neighboring pixels). Then, in each iteration:

a. The lowest pixel in the boundary list is checked for being the exit from the pond, either natural (i.e. if the aspect direction for this pixel points outwards), or through drainage pipe (i.e. if this pixel lies over the drainage pipe).

    - If this pixel represents a natural exit, the pond delineation process ends. The stage-volume and stage-area curves, the boundary polygon and data regarding the exit point and its direction are printed. DEM is changed in such a way that all pond pixels should have the same Z value, it being equal to the Z value of the exit cell. Next pit cell from the list is then selected and the process continues.
    - If this point is common to the drainage pipe, too, it is marked as the connecting point and the pond delineation process is continued.
    - Otherwise is the process of pond delineation is continued.

b. The lowest pixel in the boundary list is moved to the pond's list, the boundary list is updated, and the process is repeated from the step a) onward.

The Figure 9 explains the pond delineation process for the sample DEM and three ponds. Since ponds are delineated from the lowest pit cell, it is possible to have overlaying ponds, the ponds that are lying over the previously delineated pond (as pond No. 3 in Figure 9, which covers the pond No. 2). The developed algorithm takes care about covering ponds and shall print the list of existing links between ponds. So, the total volume of the pond No. 3, for example (Figure 9), can be easily determined from:

$$Vol(3) = Vol(2) + Vol(3) \tag{1}$$

Standard pond delineation procedure will terminate when the lowest exit point is found. However, the more advanced pond delineation algorithm is also developed, which can search for multiple exit points. The algorithm will not stop at the first exit point, but will continue to fill the pond until all exit points (or previously defined number of exit points) are found.

Figure 10 presents the result of multiple exit point algorithm: in the upper right corner, a sample DEM is given, with one pit cell and a pond with multiple exit points. This algorithm will find all exits, sort them and plot the cross-section profile (exit profile A and exit profile B, in Figure 9), which can be used as a channel profile for depression-outflow calculation.

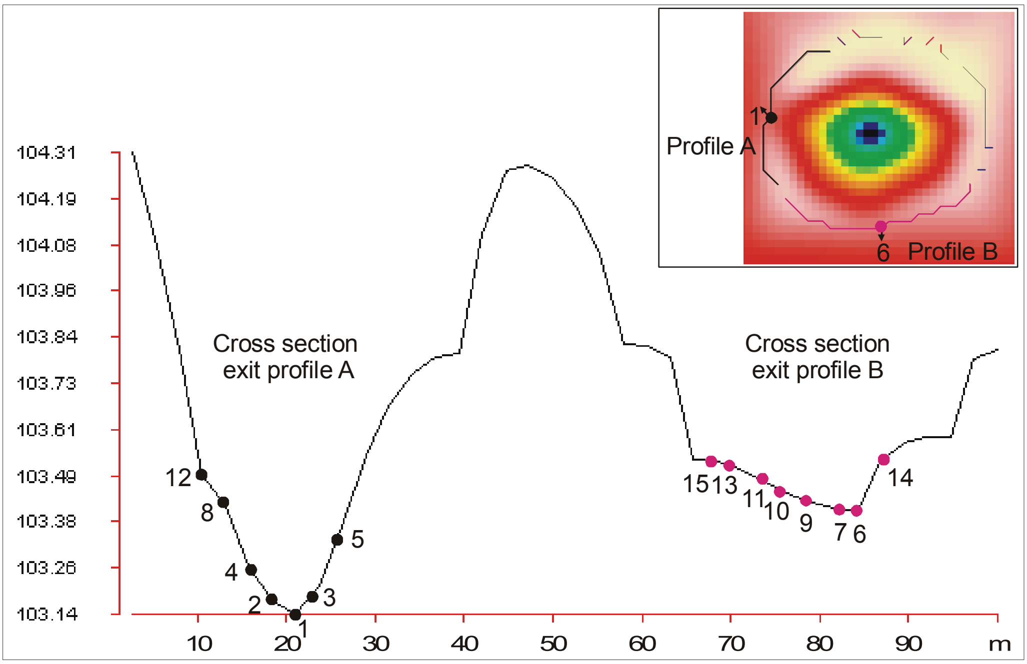


Fig. 10. Pond's exit cross section with multiple exits

The importance of DEM filling can be seen in Figure 11, where the left-hand side of the figure presents the nice 3D view along the River Drina valley. However, if a closer view is used (right-hand side of the figure) the consequences of a wrong TIN usage during DEM interpolation can be seen. Inadequately created triangle has formed the pond, which will prevent the natural propagation of surface flow path (for the sake of better visual presentation, the DEM was firstly corrected using rivers, averaged using low-pass filter and then overlaid on the top of genuine, wrong DEM).

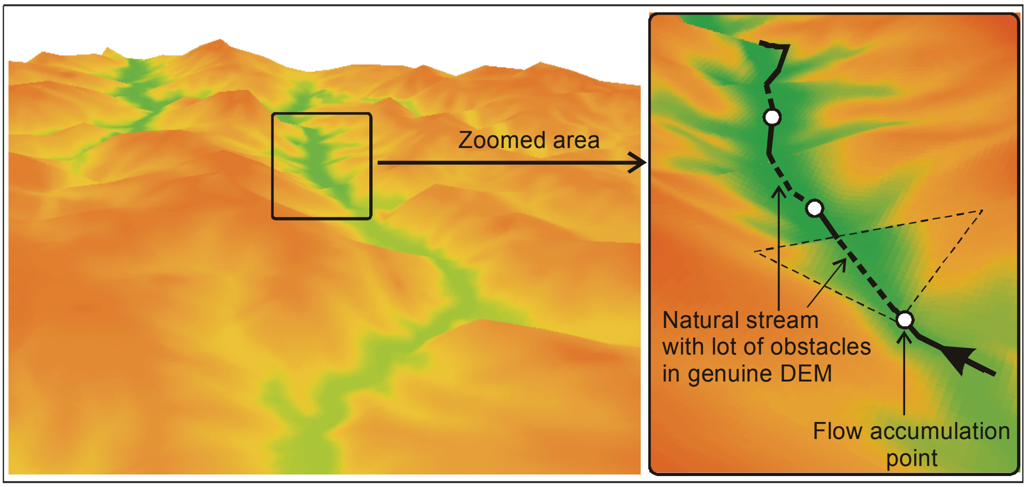


Fig. 11. An error in DEM prevents the successful flow path tracing

3.2. Surface flow path tracing

One of the main benefits of DEM usage in hydrology models is the possibility to automatically extract the surface flow network. If all morphological terrain characteristics are correctly represented by DEM, the obtained surface flow map should correspond to the existing map of rivers (Lhomme et al. 2004).

3.20	3.25	3.27	3.27	3.23	3.17
3.07	3.13	3.17	3.17	3.15	3.08
2.92	3.02	3.08	3.09	3.06	2.98
2.77	2.90	2.99	3.01	2.96	2.85
2.61	2.78	2.90	2.93	2.85	2.70

198	196	185	168	151	146
205	204	193	172	152	146
212	213	202	172	148	145
219	222	212	169	141	142
226	231	220	162	132	134

0	0	0	0	0	0
2	1	1	1	1	2
3	3	2	2	1	3
5	4	1	3	2	3
7	4	1	3	4	5

Fig. 12. Sample DEM (left), its aspect image (central) and flow accumulation image (right)

All surface flow path tracing algorithms are based on a calculated aspect image (Subsection 2.2, Figure 4), where each pixel holds the directional angle of its maximal slope. Figure 12 presents the sample case of DEM (left side of image, DEM heights are given for each pixel) and calculated aspects (in the central image, angle is calculated using 1st order derivative and the directional angle is oriented clockwise from the South). On the right-hand side of the Figure 12 the "flow accumulation image" is given: the value within each pixel equals the number of upstream pixels, i.e. for each pixel the total flow-contributing area (or the upstream area) is known.

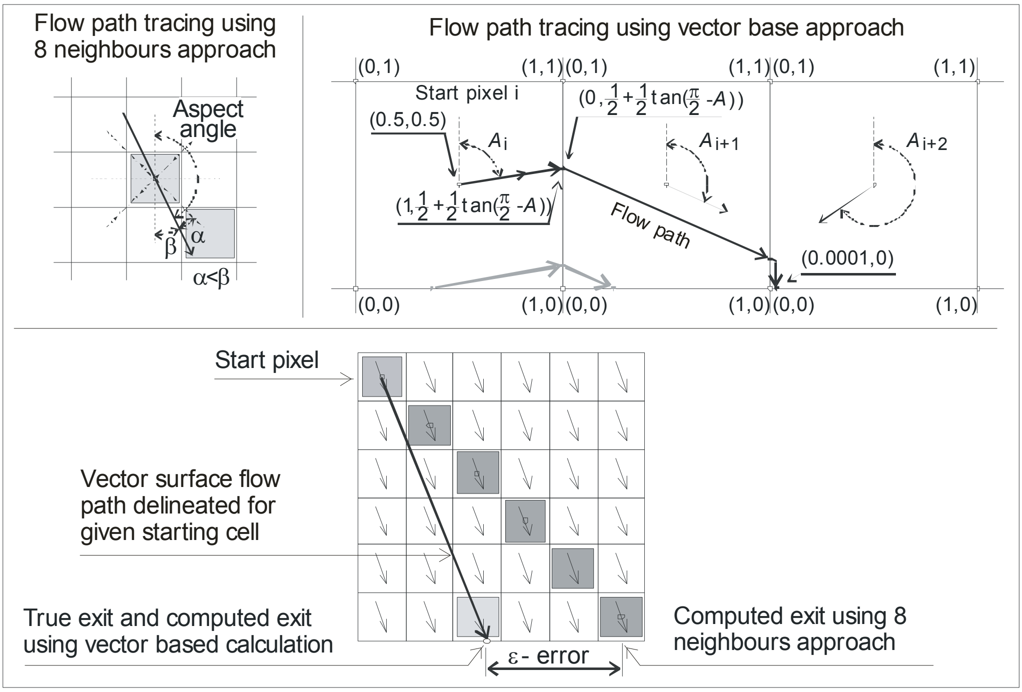


Fig. 13. Raster-based and raster/vector-based flow path tracing approach

The algorithm for flow accumulation generation in a raster system can be quite simple: for each pixel the downstream pixel (pixel along the aspect angle) among 8 surrounding pixels is marked and used for next iteration. The algorithm is presented in the upper left corner of Figure 13: starting with the central pixel, water will flow towards the marked down-right pixel. If such an algorithm is used for terrain presented in the lower half of the Figure 13, the computed exit point will significantly differ from the actual exit point.

Much better and more complex routine was developed and used in automatic preparation of data needed for SWAT hydrological models. The flow path tracing algorithm simulates the “ball rolling” over the terrain. This algorithm is the combination of raster and vector approaches: for each pixel in a DEM a water flow is traced using the vector representation of the flow (as given in the upper right-hand part of the Figure 13) until the end of the DEM is reached (starting from the pixel’s center and using the aspect angle as direction, the exit and entrance points for each pixel are calculated). Whenever the calculated flow path enters into the new pixel, its count of total upstream pixels is increased by one. By the use of such an algorithm on the test DEM presented in the lower part of the Figure 13, the correct exit point will be found.

The result of the used algorithm can be either global, in raster format, where flow accumulation image is created (Figure 12, image on the right-hand side), or local, in vector format, where vector representation of surface flow path originating from certain point within DEM is computed (Figure 13, the lower part). In hydrological modeling, the flow accumulation image is very useful, since it can be used for extraction of the river network with different flow contribution (upstream) areas. Figure 14 presents the result of such analysis performed for River Drina catchment: by increasing the upstream area from 300 ha to 10000 ha, the surface network can be easily simplified, leaving only main rivers.

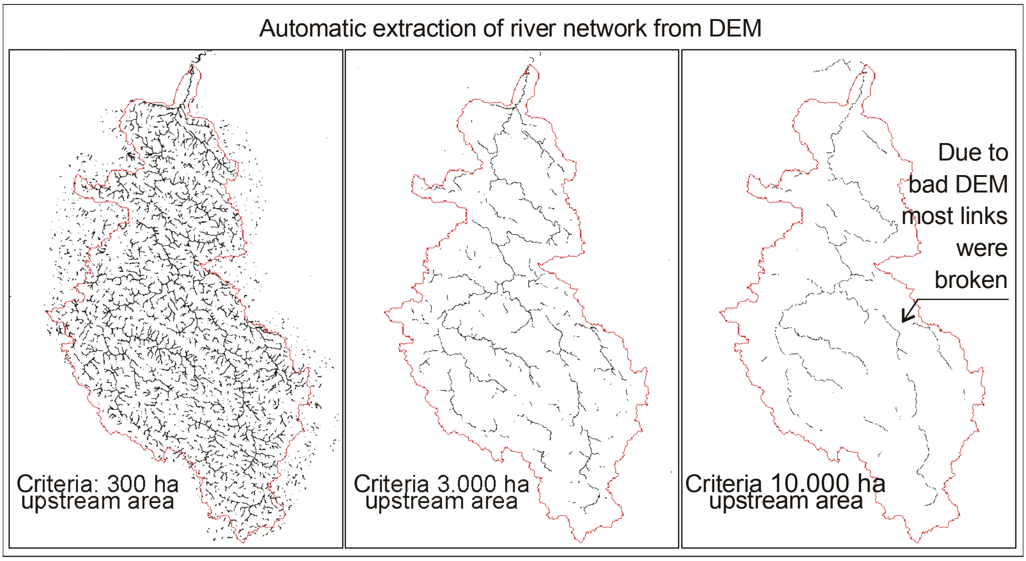


Fig. 14. Influence of different upstream area thresholds on river network density

However, the same figure clearly presents the influence of the DEM with large number of false ponds on the result of surface flow analysis. The surface flow tracing algorithm will in such situations terminate its operation as those false ponds are encountered (Figure 11 is an example of such behavior), leading to a large number of "broken", or disconnected river segments. To overcome those problems, the filling (Figure 8) and breaching (Maksimović et al. 2009) algorithms are necessary, or delienation of depressions and linkage of flow path and depressions.

3.3. Subcatchment delineation

Subcatchment is the area drained by certain sink point (manhole, for example, or small depression) or a line (river or channel). Subcatchment area and its topographic attributes (average slope, shape, length etc.) are needed for all physically based hydrological models. Once when drainage network is known, the catchment area partitioning (or delineation) should be done according to the elevation map, following the natural surface flow paths. Historically, it was performed manually, following ridges and watersheds from paper maps, or on-screen digitizing. However, having the DEM, the delineation for certain river network could be done automatically (Freeman 1991).

In general, there are two possibilities for automatically delineation of subcatchments. The first one would be to use the "downstream" approach: each pixel in grid system is used as the source pixel for flow path tracing, until some sink point or line is reached. The ID of the source pixel is then changed to reflect the ID of the found sink pixel. This approach would be based upon the previously described flow path tracing algorithm (Section 2.4) and would be consistent with the flow accumulation image and its derivative, the surface flow network. However, such an algorithm has also few major drawbacks: firstly it is highly sensitive to DEM errors and small depressions, and secondly, if the point sinks should be used, large portions of catchments would be un-drained since most flow paths would miss the terminal sink points.

Another possibility for delineation of the subcatchments is to use the "upstream" approach: the algorithm will "climb" upstream from given sink pixel(s) until all pixels that can supply the water to the sink pixel(s) are collected. Such approach is more robust regarding the small errors in DEM and can produce more realistic boundaries of subcatchments than the previous one.

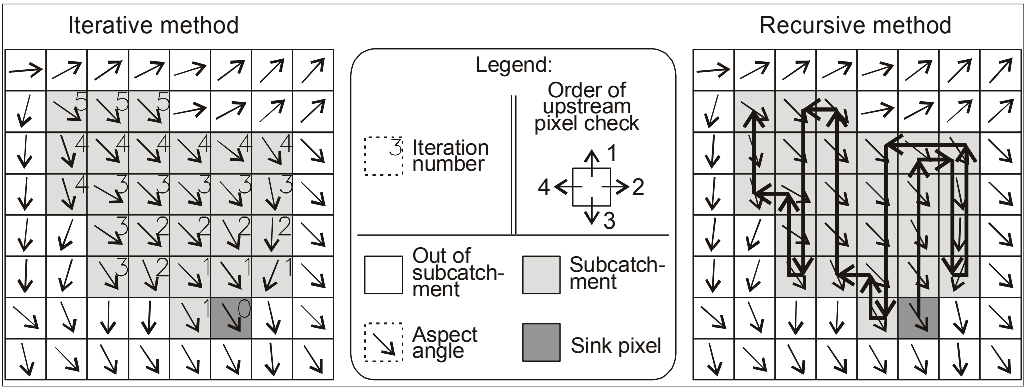


Fig. 15. Iterative and recursive method of subcatchment delineation

The “upstream” approach can be implemented either iteratively or recursively (Figure 15). The iterative approach will scan the whole DEM for pixels already assigned to certain subcatchment, which have unresolved upstream neighboring pixels. Such pixels will receive the ID’s of their downstream pixels. The process will reiterate until no new pixels can be resolved. The process is slow, especially for large DEMs that contain millions of pixels.

The recursive method of subcatchment delineation will, on the other side, “walk upstream” from the source pixel, until all pixels of the subcatchments are resolved (Figure 15, right-hand side). Both methods will produce the same result, but the recursive method is much faster for large DEM images. For both methods it is essential to use sorted list of source pixels.

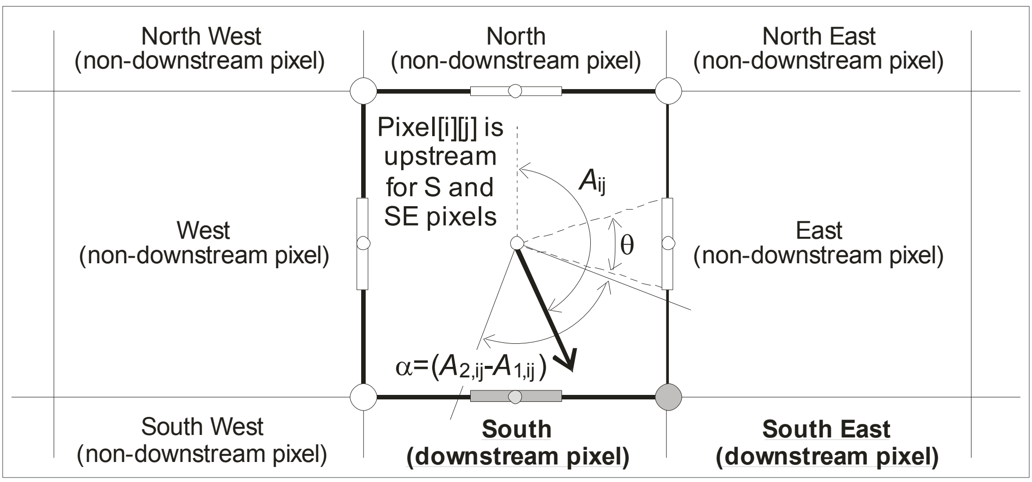


Fig. 16. Criteria for selection of a downstream pixel

The “upstream” approach of subcatchment delineation needs certain additional criteria when it has to decide whether certain corner pixel is upstream (or downstream) or not. Figure 16 presents the problem: the aspect value of [i][j] pixel Aij is such that water will flow towards its South-oriented pixel. But the south-oriented east pixel can be also seen as the downstream pixel, since certain amount of water can also flow towards that corner. To facilitate the "upstream" approach, two new variables are introduced, as presented in Figure 16:

1. the outflow angle α, which defines how wide can water spread from certain pixel, and
2. the acceptance angle θ, which defines the part of certain boundary that has to be covered by the outflow angle in order to consider that pixel as downstream pixel.

Both variables can be constant for the whole catchment, or can be a function of a DEM image (for example, the outflow angle can be large for flat areas and small for steep areas) or coverage image (for example, the influence of man-made objects can be simulated using those variables). The default value for outflow angle is 90^0 and for acceptance angle is 5.7^0. It is also possible to use iterative adjustment of acceptance angle: in the first iteration the delineation can be conducted with the default value, and if some pixels of the catchment are un-delineated, in the next iteration the value of the acceptance angle can be increased.

Figure 17 presents the result of automatic subcatchment delineation for given stream of surface flows – rivers. First, the rivers are rasterized by using the river ID as the pixel ID. Then, the pixels along the rivers are sorted depending to their altitude and used as source pixels in the upstream recursive method (Figure 15). The default values are used for outflow and acceptance angles. However, after delineation of the whole catchment, if some pixels should still remain unresolved, the user can re-run the procedure only for unresolved pixels, by using somewhat wider outflow and/or acceptance angles.

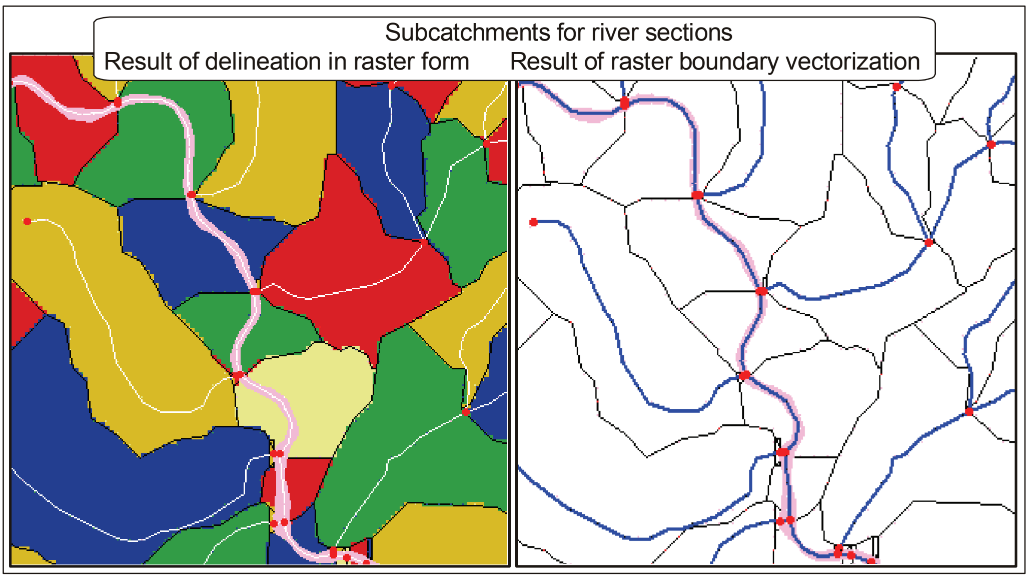


Fig. 17. Result of subcatchment delineation in raster in vector format

The left-hand side of the Figure 17 presents the output of this procedure, in raster format, where each pixel has the ID of the corresponding downstream receiving river. For most GIS systems it is better to have the boundary in the vector format, as well as lines and connecting points, therefore a specialized vectorization routine was developed. The routine will firstly create the common points at the intersection of the three subcatchments (Figure 18) and then will run the line generalization routine (Douglas and Peucker 1973), (Hershberger and Snoeyink

1992) for boundary lines between those common points. Depending on the used filtering parameter, the number of line vertexes could be dramatically reduced.

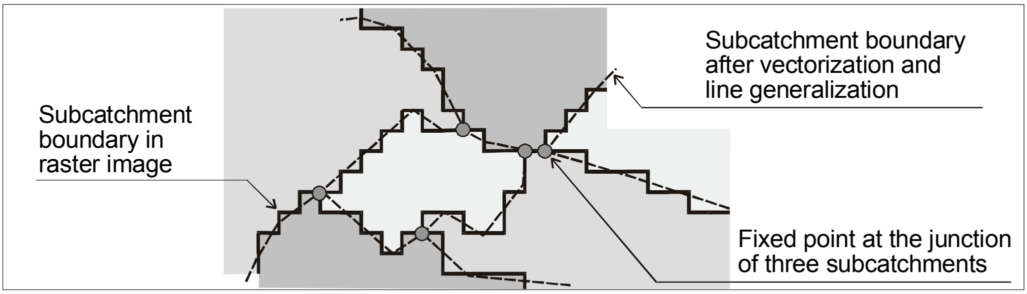


Fig. 18. Vectorization routine has to be customized, to keep the junction points in resulting vector image

The subcatchment delineation routine will delineate the whole catchment if the DEM is prepared as "depressionless DEM" (Band 1986), i.e. if all depressions are filled-in (Section 3.1). If larger depressions (or ponds) were not corrected in DEM, after subcatchment delineation the ponds would behave as shades – they would prevent the upstream water to reach the recipient, the river, since they would trap the water.

One of the possible solutions in the case that DEM has ponds in it is to use separately the previously delineated ponds as a source for the new delineation run. The result of such delineation will be the boundary of pond's subcatchment, i.e. the area drained by each pond. Applying the basic hydrology analysis to convert the net rainfall in such an area into the outflow that runs into the pond and fills it with water, the water balance equation can be applied by using the stage-volume curve extracted for that pond. By using the surface flow path tracing (Section 3.2) with the origin at the pond's exit point, the pond's overflow can then be routed to the nearest downstream river.

4. DEM-related parameters used in SWAT

There are a number of parameters related to DEM that most physically based hydrology models need. Firstly, all physically based models need a drainage network. The network can be entered into the model either directly, by digitizing the rivers from paper maps (or simply purchasing the layer of river vector lines), or by using DEM in order to prepare the flow accumulation image (Figure 12) and true drainage network (Figure 14) and extracting the river network for specified upstream threshold area. The DEM-based approach is much more versatile of the two, since user can easily change the level of detail (or river rank) in the hydrology model.

Secondly, the model needs the "contributing" area for each river segment. Traditionally, this area was entered by manual drafting and looking at the paper isoheight maps. The developed algorithms can use the DEM and the rivers as a source points, and "delineates" the whole catchment into the subcatchments, or "contributing" areas. Once that subcatchment boundaries are known, it is easy to calculate the subcatchment area, average, minimum and maximum height, or average, minimumal and maximum slope (Figure 4). Certain hydrology models also require the subcatchment shape factor, mostly calculated as the ratio of its area and perimeter. If other catchment data is available in the digital form (land use data, soil type data, canopy data etc.) the subcathment can be used as a “window” to extract all needed statistical parameters (percentages of soil types, for example).

The usage of the surface flow components largely depends on the concept of the hydrological model used. There are models that simulate in detail the actual surface flow, possible flow accumulation in ponds, the overflow of the water over the pond’s boundary and links between several ponds. In such models, the ponds represent an integral part of the surface drainage network. More often, the surface flow is not simulated in detail, but instead of that only the surface flow parameters are used for calculation of the “time of concentration” parameter – the time that water needs to reach the exit of the catchment (Simić et al. 2009).

The SWAT hydrological model uses the time of concentration for calculation of the surface flow rate. For the calculation of the time of concentration, the SWAT model needs the following parameters: length of the surface flow path from each pixel up to the receiving river, the average slope along that path, the length along the river and the average slope along the river.

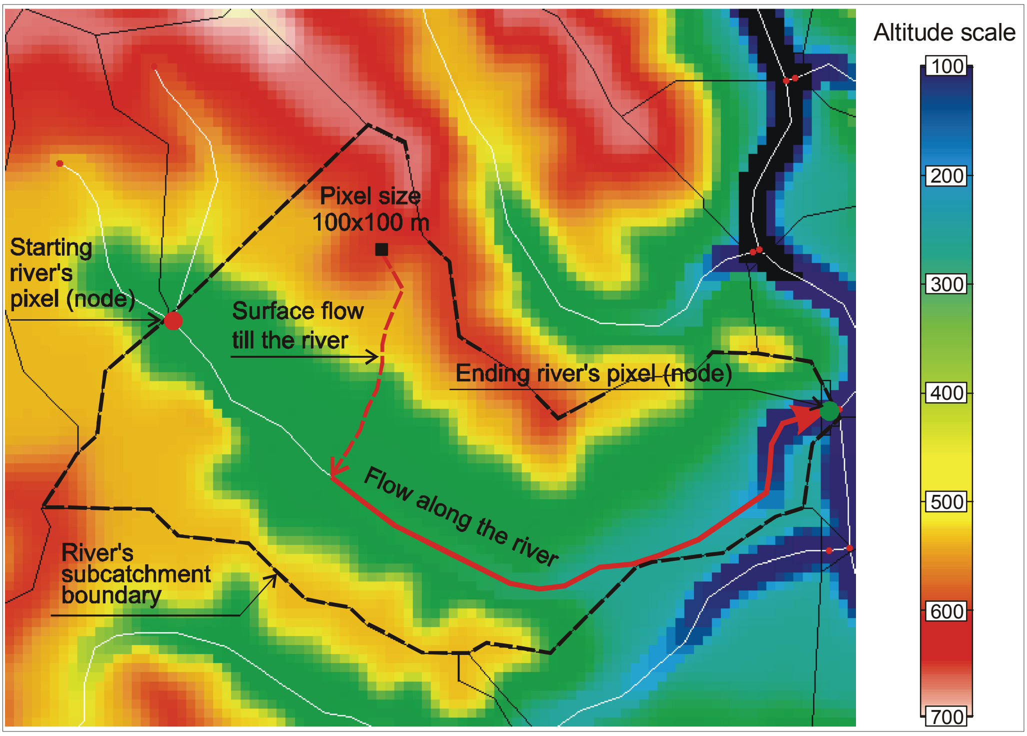


Fig. 19. DEM with river segment (white line), contributing catchment (subcatchment) and the flow paths along the surface and along the river

Figure 19 presents the sample DEM with the river network (white line), segmented into the ranks. The river network is presented as an oriented graph: each river’s segment is oriented from its upstream to its downstream node and all river segments are connected to form a branched network. Each river’s segment is used to delineate the subcatchments. The boundary of one selected subcatchment is presented in Figure 19 with the dashed black line. To calculate the parameters needed for the SWAT model, each pixel from within one subcatchment has to be used to trace the surface flow path, up to the crossing with the river segment (dashed red line).

Then, starting from the point of the surface flow path and the river intersection, the length of the river's segment up to its ending node (i.e. up to the exit from the subcatchment) is calculated (solid red line). For both surface and river flow path, the average (the arithmetic mean value) slope is calculated by using the slope image.

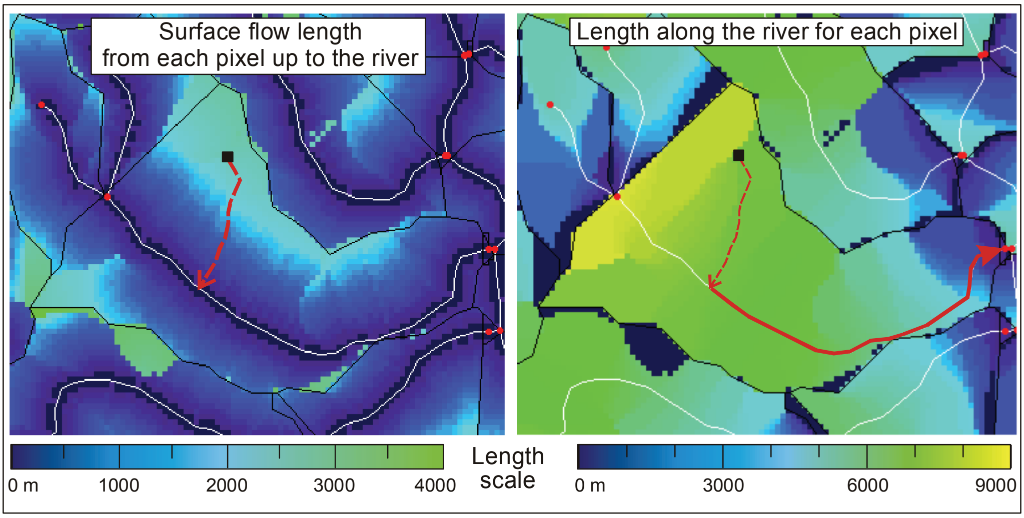


Fig. 20 (a) DEM related parameters needed by SWAT model (for subcatchment on Figure 19)

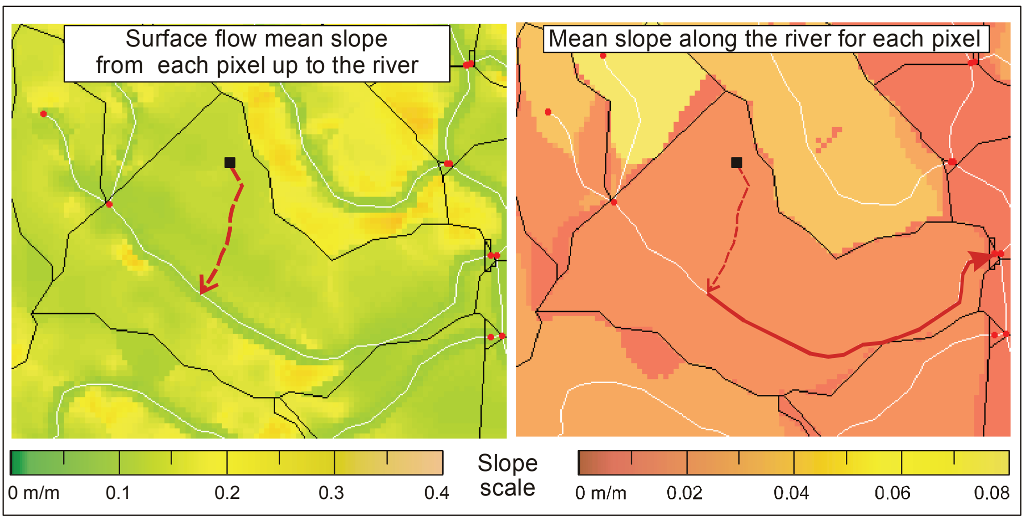


Fig. 20 (b) DEM related parameters needed by SWAT model (for subcatchment on Figure 19)

The algorithm explained in Figure 19 has to be invoked for each pixel in the subcatchment, in order to calculate the two lengths and two slopes. The result of the calculation are 4 new layers, stored in 4 separate images, as presented in Figures 20-a and 20-b (the starting pixel used during the calculation presented in Figure 19 is highlighted). The pixel size is 100x100 m.

Basic DEM resolution, or raster size, is always a compromise between the level of detail needed for computation, the accuracy of the data and the total computing power and time resources that one can allocate. On one hand it is important to keep the pixel size small enough to capture all important terrain features. On the other side, the concept of the most physically based hydrological models is such that a computational unit has to be large enough to allow for a proper spatial averaging.

As mentioned in the introductory section of this paper, the SWAT model delineates the space into the set of the HRUs (Hydrology Response Units). HRU can be of different shape, but it has to have the homogenous land, soil and vegetation parameters. For each HRU the SWAT model will use a set of 5 reservoirs to simulate the water transfer between rainfall, evaporation, surface runoff, and underground flow (Simić et al. 2009). To apply such concept, the HRU has to be larger then pixel size used in DEM processing.

Also, the catchment delineation concept, where a large number of small subcatchments, one per each river segment, is automatically created, is too detailed for the SWAT model. Although the water balance should be calculated per each HRU, the true outflow (or hydrograph) is generated only at the exit (outlet) of one larger subcatchment, which consists of several upstream small subcatchemnts. In the SWAT terminology, such outlet is referred to as “Hydrology Profile” (HP).

Figure 21 presents an example of HRU and HP aggregation. Within developed algorithm, all HRUs are rectangular and defined as an average of the n x n pixels (in the case of Drina catchment, n=10. and HRU size 1 x 1 km). At the selected point along the river the user has to decide where he wants the establish an HP, meaning that in that point a hydrograph will be available and that all upstream subcatchments would be combined into one larger subcatchment (the dashed line in Figure 21).

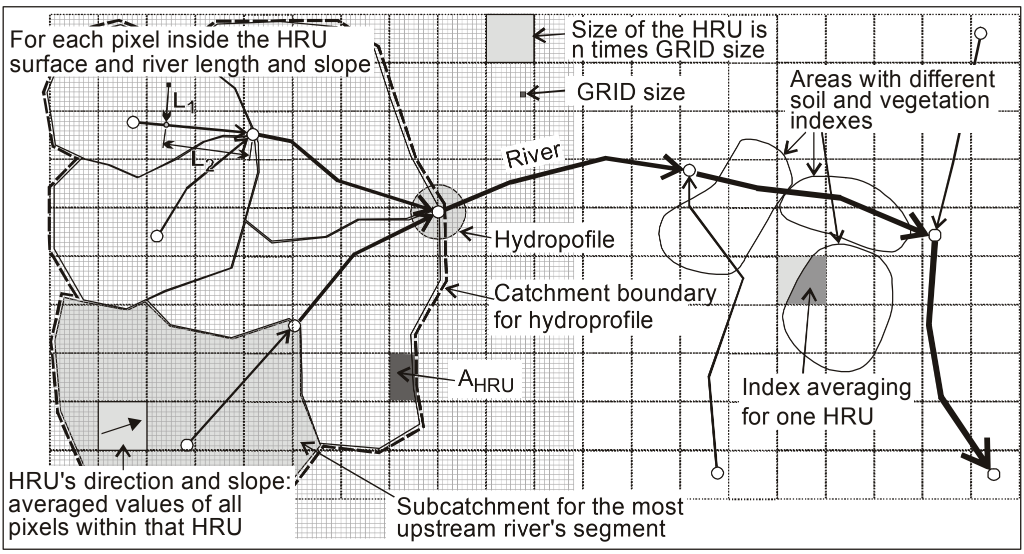


Fig. 21. Aggregation of subcatchments and pixels to form SWAT input data

All parameters (height, slope, distances from the river and lengths along the river, slopes toward the river and along the river etc.) are averaged within the HRU. The flow path length along the river (L2 in Figure 21), which was calculated till the end of the small subcatchment, is corrected by the total length of downstream rivers, up to the selected HP. For HRUs located at the subcatchment boundary, it is necessary to divide each HRU into smaller parts, one per subcatchment.

The right-hand side of the Figure 21 presents the handling of the other spatial types of data, such as soil and vegetation. They are entered as polygons and their data is stored within the HRU as the percentage of different area types.

5. Conclusions

The life of simulation model users, especially hydrologists, would be much easier if they could have available measured data about observed inputs and outputs from the system, together with system parameters, and if they could use the contemporary data acquisition and handling techniques for automatic creation of input files. Although the first problem of data availability is more important, this paper assumes that user can obtain enough data for model calibration and verification. This paper address the second problem, how availability of massive and high precision input data, such as terrain elevation, surface (canopy) description, rainfall distribution etc., can speed up creation of simulation models and even provoke the changing of fundamental model concepts.

In this paper the usage of digital elevation models (DEM) is explored for the needs of automatic creation of input data necessary for the application of physically based hydrology models. For the computations in River Drina catchment the regular grid of DEM type was considered, with 100x100 m pixel size. The DEM containeda number of errors, such as flat areas, peaks and pits, so additional data processing (like “deflattening” or conditional filtering and small pit removal) was necessary in order to eliminate the imperfections in DEM.

From corrected DEM, a set of input data for the SWAT hydrology model was created. First, a slope and aspect images were computed. Then, the larger ponds were recognized and their stage-volume and stage-area curves extracted. By using the flow accumulation image, surface flow drainage network is created and checked against the existing river vector data. For all river segments contributing areas, or subcatchments, were delineated. Next, for each pixel within a certain subcatchment ,the length and slope of the surface flow section up to the river and the length and slope of the flow section along the river till the subcatchment’s outflow was computed. Finally, all data were resampled and aggregated into the network of HRUs, 1x1 km in size.

It is important to mention that presented DEM-based GIS algorithms and produced data are not the only data that hydrological models need. This paper has not covered the GIS operations (mostly overlay and reclassifications) on spatial coverage data (representing soil type, land use etc.) defined using polygons. Also, the number of GIS algorithms could be used for dealing with the linear-features data such as networks, which are defined by using points and connections in a form of oriented graph, in a way that an up-stream/down-stream propagation through the network is possible. Finally, the point data are also used in GIS in order to prepare the data for hydrology models, since points are used in order to keep the measurements from the meteorological stations (rainfall intensities, temperatures, radiation, humidity and other meteorological quantities). The problem of spatial extrapolation of such point-measured data is important and so is the correlation of point-measured data with other spatial data (DEM, coverage type, etc.).

Throughout the paper, the problem of DEM accuracy and usability is assessed, especially if data is purchased from a general data supplier in a digital raster format with fixed resolution.

Regarding a purchase of digital data for hydrological modeling, the authors can suggest the following data checks to be performed: a) check of the fraction of the area that has the slope of 0%, b) use of the surface flow routing tool (can be found in most GIS packages) in order to check whether water will flow across the whole study area if arbitrary source point is selected, c) check of the orientation and interconnection of the stream network, and d) use of streams as "cut-outs" across the DEM and check whether the slope of a longitudinal section toward the lowest, exit node, is continuous. The authors' experience is that in most cases (for instance in the case of River Drina catchment), a large amount of additional work is needed for data processing and attributing. It is, of course, up to the scientific community to warn the digital data producers that they should cope with this problem and be ready to prepare the user-specific (or problem specific) views on the same datasets.

References

Arnold J.G., Williams J.R., Srinivasan R., King K.W., Griggs R.H.. (2002). SWAT - Soil and Water Assessment Tool, USDA, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, 808 East Blackland Road, Temple, TWRI Report TR-191.

Aquaveo (2009). WMS (Watershed Modeling System). Aquaveo, Provo, Utah 84606, US, http://www.aquaveo.com/wms-details

Band L.E. (1986). Topographic partitioning of watersheds with digital elevation models. Water Resources Research 22(1), 15-24.

Boonya-aroonnet S., Maksimović Č., Prodanović D., Djordjević, S. (2007). Urban pluvial flooding: Development of GIS based pathways model for surface flooding and interface with surcharged sewer model. Proc. 6th Intl. Conf. NOVATECH, Lyon, France, 481-488.

Burrough P.A. (1993). Principles of Geographical Information Systems for Land Resources Assessment. Clarendon press, Oxford.

DHI (2009). MIKE SHE. DHI, Agern Allé 5, DK-2970 Hørsholm, Denmark, http://www.dhigroup.com/Software/WaterResources/MIKESHE.aspx

Douglas D., Peucker T. (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, The Canadian Cartographer 10(2), 112-122

Djordjević S., Prodanović D., Maksimović Č. (1999). An approach to simulation of dual drainage. Water Science & Technology 39(9), 95-103.

Elgy J., Maksimović Č., Prodanović D. (1993). Matching standard GIS packages with urban storm drainage simulation software. Proc. HydroGIS 93: Applications of Geographic Information Systems in Hydrology and Water Resources, IAHS Publication 211, 151-160.

EPA (2009). HSPF (Hydrological Simulation Program – FORTRAN, U.S. Environmental Protection Agency, http://www.epa.gov/ceampubl/swater/hspf/

Freeman T.G. (1991). Calculating catchment area with divergent flow based on a regular grid. Computers and Geosciences 17(3), 413-422.

Garbrecht J., Martz L.W. (1996). Comment on Digital elevation model GridSize, landscape representation, and hydrologic simulations, by W. Zhang and D.R. Montgomery. Water Resources Research 32(5), 1461-1462.

Garbrecht J., Martz L.W. (2000). Digital elevation model issues in water resources modeling. Hydrologic and hydraulic modeling support with geographic information systems. ESRI Press, Redlands CA, USA.

Hershberger J., Snoeyink J. (1992) Speeding Up the Douglas-Peucker Line-Simplification Algorithm, Proc 5th Symp on Data Handling, 134-143

Institute for Development of Water Resources "Jaroslav Černi" (2005). Hydroinformatic System of Drina Catchment, version 2. Publication in Serbian.

Jenson S.K., Domingue J.O. (1988). Extracting topographic structure from digital elevation data for geographical information systems analysis. Photogr. Engng. Remote Sensing, 54(11), 1593-1600.

Leitão, J. P., Boonya-aroonnet, S., Allitt R., Maksimović, Č., Prodanović, D., (2008). Modelling of flooding and analysis of pluvial flood risk - Demo case of UK catchment. In proceedings: FLOODRisk 2008 Conference, Oxford, UK.

Lhomme J., Bouvier C., Perrin J.L. (2004). Applying a GIS-based geomorphological routing model in urban catchments. Journal Hydrology 299(3/4), 203-216.

Maksimović Č., Prodanović D., Marković I. (2004). Adaptable Urban Drainage Addressing Climate Change in Intensity, Occurence and Uncertainty of Stormwater – AUDACIOUS. EPSRC Founded project. Internal publication.

Maksimović, Č., Prodanović, D., Boonya-aroonnet, S., Leitão, J. P., Djordjević, S. and Allitt, R. (2009). Overland flow and pathway analysis for modelling of urban pluvial flooding. Journal of Hydraulic Research 47(4), 512-523.

O'Callaghan J.F., Mark D.M. (1984). The extraction of drainage networks from digital elevation data. Computer Vision, Graphics, and Image Processing 28, 423-444.

Prodanović D., Djordević S., Maksimović Č. (1998). GIS assisted model for dual drainage simulation, V. Babovic, L.C. Larsen, eds. Proc. 3rd Intl. Conf. Hydroinformatics, Copenhagen, Denmark, 535-542. Balkema, Rotterdam NL.

Prodanović D. (1999). Improvements of Hydroinformatic Methods in Analysis of Flow From Urban Catchments. PhD Thesis, Faculty of Civil Engineering, University of Belgrade. Published in Serbian.

Prodanović D., Stanić M., Divac D. (2006): Automated Feeding of Numerical Models with Digital Input Data: Case Study of the Drina Catchment. 7th International Conference on Urban Drainage Modelling UDM '06, Melbourne, Australia.

Simić Z., Milivojević N., Prodanović D., Milivojević V., Perović N. (2009). New Rainfall-Runoff Model of Complex Catchments Based on SWAT Methodology – Theoretical Background and Numerical Procedures. Journal of the Serbian Society for Computational Mechanics, Vol. 3, No. 1.

Stanić M., Prodanović D., Komatina D. (2004) Software development for automatic data preparation for Drina Catchment model, Internal publication in Serbian.

Tarboton D.G. (1999) TARDEM, A suite of programs for the Analysis of Digital Elevation Data, http://www.crwr.utexas.edu/gis/gishydro99/uwrl/tardem.html

USDA-ARS (2009). SWAT (Soil and Water Assessment Tool). Grassland, Soil & Water Research Laboratory, USDA-ARS, 808 East Blackland Road, Temple, Texas 76502 USA http://www.brc.tamus.edu/swat/

WRA (2009). HYSIM (HYdrological SImulation Model). Water Resource Associates, PO Box 838, Wallingford, Oxon, OX10 9XA, United Kingdom, http://www.watres.com/software/soft-hysim.html