Progressive increase of inputs in floodplain delineation based on the DEM: application and evaluation of the model in the catchment of the Opava River

FILIP HARTVICH[1] A JAN JEDLIČKA[2]

[1] Dept. of engineering geology, Institute of Rock Structure and Mechanics,
 Academy of Sciences of the Czech Republic
[2] Dept. of cartography and geoinformatics, Faculty of Science, Charles University in Prague

Abstract

This contribution concerns the current stage of development of a practically applicable model for automatized delineation of the floodplain. The chief principle, which was added in the present model development stage, was application of the progressive change of model input parameters. The motivation of embedding of this principle was to converge the modelled floodplain to the situation and real properties of the floodplain. In particular, the progressively changing parameters were downstream increasing maximal relative height of the floodplain above the river level and downstream widening buffer zones of maximal floodplain width. The increase of the parameters is defined by a mathematical function. Partial automatization of the process in the ModelBuilder extension in the ArcGIS environment allows its application on the whole catchment at once, therefore it is suitable for floodplain delineation in vast areas, where the complete field mapping would be expensive or time-demanding, or where only partial field data are available. The model was applied in the catchment of the Opava River, and its results were compared to those of the previous stage model and also field-controlled on several localities. The conclusion from the verification is that the progressive change of parameters is a suitable technique that improves the model performance in previously problematic areas. There are, however, still some problems, particularly in setting of the model parameters.

Key words: floodplain delineation – DEM – ArcGIS – model builder – progressive change of inputs

1. Introduction: demands and possibilities

This article develops the methods of automated floodplain delineation, based on the digital data, which have been designed by the authors during last few years. This attitude represents a cheaper, less complicated alternative to the remote sensing methods, currently widely used (Townsend and Walsh 1998, Hudson and Colditz 2003). The current stage of the project is aiming to achieve several goals:

- to improve the performance of the delineation algorithm by implementation of new techniques
- to apply and test the new algorithm in the catchment of the Opava River
- to compare the results of the new algorithm with the performance of the previous version (see Hartvich 2007a)
- to design a compact set of tools for ArcGIS, which can perform the whole delineation process automatically after setting several input parameters

The most improvement step, which was embedded into the delineation process in order to simulate the real situation, was the implementation of the progressive change of the parameters (further PCP, see Hartvich and Jedlička 2007). It means that the width and relative height of the floodplain increases gradually along the course of the river according to a mathematical formula (Fig. 1). This leads to more realistic floodplain delineation, rather narrower in the upper parts of the courses and wider along the lower course, which rule is basically valid for most of the rivers, with possible exception of strongly structurally or anthropogenically influenced streams (Křížek et al. 2006).

The new version of the model (PCP) was applied together with the original version (SP) in catchment of the Opava River, as far as Opava town. In this area, the performance of the two models was compared.
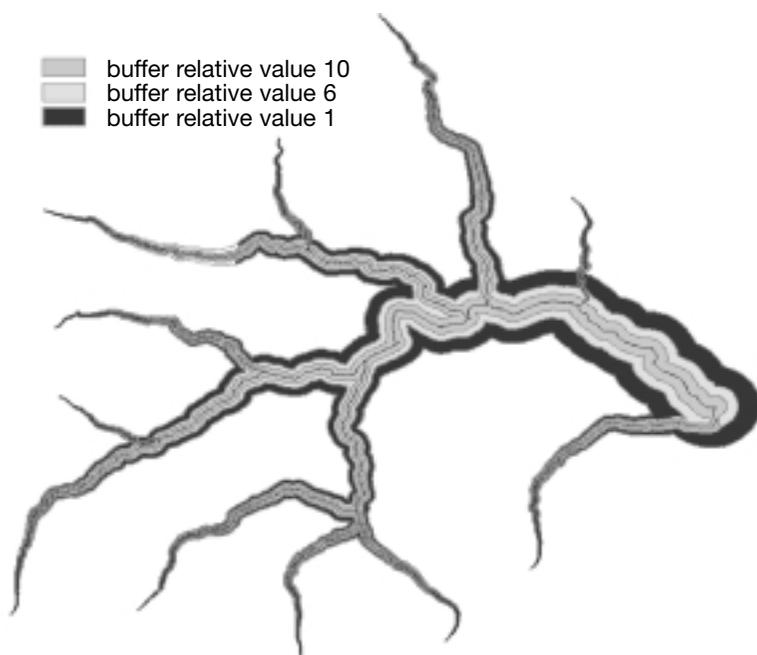
Fig. 1

1.1 Input data properties and pre-processing

An important factor, influencing the reliability of the final result, is the type, quality and processing of the input data. This concerns particularly the main inputs: layer of river polylines and the DEM. The more detailed and accurate is the DEM, the better can be the results of the delineation. However, in contrast stand the time and computational demands of the process, therefore it is necessary to choose suitable raster accuracy for the model (Andrysiak and Maidment 2000).

As the most widely used source data for creating the DEM are the contourlines (ZABAGED10 or DMU25), the function TopoToRaster was used. The advantage of this function is that the output DEM is hydrologically corrected, i.e. does not contain sinks or other flow-related errors and all the cells are part of some catchment, closed at the raster border. Same result can be reached by application of FillSinks function, included in the ArcHydro extension (Anderson 2000).

The other elementary data input is the layer of river polylines. Correct pre-processing of these data is also necessary before using the layer in the model. In particular, it is necessary that each stream has a unique ID, based on which it is possible to identify the whole stream. Secondly, the topology of the polylines must be clean so that the streams are connectible.

2. Methods: General model overview

Based on the experience from previous model construction, it was decided to use following three parameters as inputs for the delineating of the floodplain:
- Slope inclination (floodplain is generally flat)
- Distance from the riverbed (floodplain spreads around the stream course)
- Relative altitude above the water level (floodplain is created by the fluvial activity, therefore must be in the reach of the flooding)

☐	< 3 m
▨	3–5 m
▦	3–5 m
■	< 7 m

2000 m

S
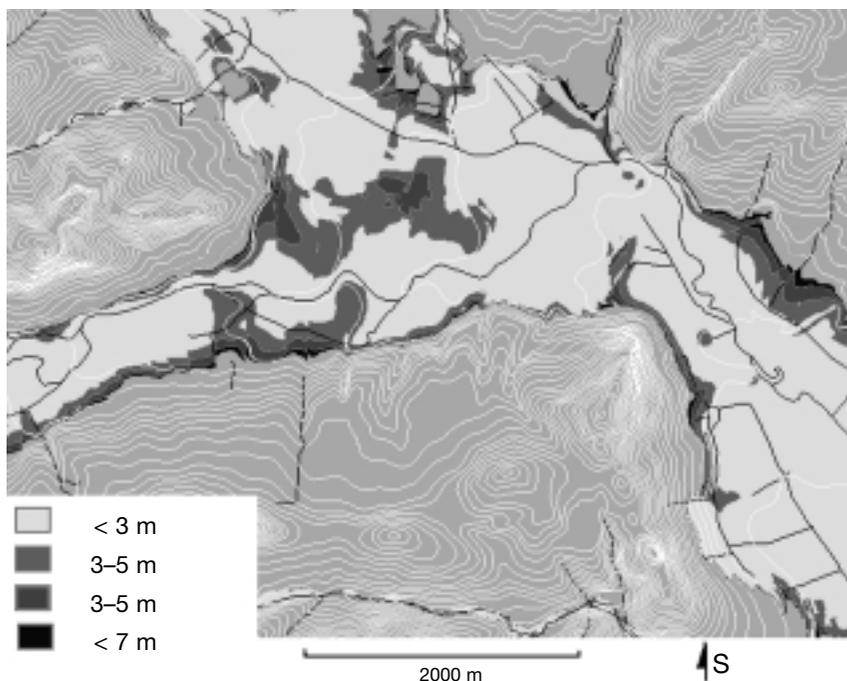
Fig. 2

As only limited amount of information was used for floodplain delineation in the model, it is clear that the floodplain can never be delineated precisely. There is always some level of uncertainty. To be able to work with this uncertainty, computed parameters were mapped to corresponding relative value. To each value of computed parameters, appropriate relative value was assigned, based on probability of floodplain occurrence. For example, the slopes with inclination $0-1°$ were labeled with a relative value of 10, while the slopes with inclination higher than $10°$ were assigned relative value 0, meaning no part of the floodplain – in the given raster resolution – can reach $10°$ (Fig. 2).

Similar approach was used for assignation of relative values to the other parameters, which describe distance of a particular raster cell from the riverbed, and the relative altitude of floodplain above level of river. In these case, the specific value wasn't assigned to every single cell (as in case of slopes), but to zones with specified distance from river (buffers, or to belts of certain height above the river level).

After relative values were assigned to all parameters used for floodplain delineation, these parameters were combined (summation of relative values) to find areas, where was high probability of flood plain occurrence. The assignment of relative values was the only subjective part in process of flood plain delineation and need some prior expert knowledge about the problem, and about area of interest.

The next step was to assure the technical functionality of the model, i.e. deriving the necessary information from the source data and its processing in the ArcGIS environment, particularly from the DEM and the river/stream network geometry.

For the last two parameters, a certain pre-processing of the inputs was necessary. Using adjustments of the linear referencing tools to calculate the distance along the river polyline was employed. As the complete description of the linear referencing problematics is beyond the scope of this paper, we present here only the general steps of the process.

The main difficulty was that the standard ArcGIS routing feature tool can not differentiate the river source from the river mouth (or sink). This problem was worked out using the height information of these two points, derived from height raster. Naturally, the higher point was labelled as river source. When this information was available, it was possible to determinate correct way for distance accumulation. Correct distance accumulation is necessary for creation of the route feature, where for each river distance along river increase from source to sink.

2.1 Slope inclination

This indicator was derived from the hydrocorrected DEM, using tools available in ArcToolbox. After new slope raster was derived, its absolute values (in degree or percents) were reclassified to required relative values.

2.2 Distance from riverbed

The distance from riverbed could be calculated in ArcGIS using the Euclidian Distance or Multi Ring Buffer tools (creating zones, where all points are located
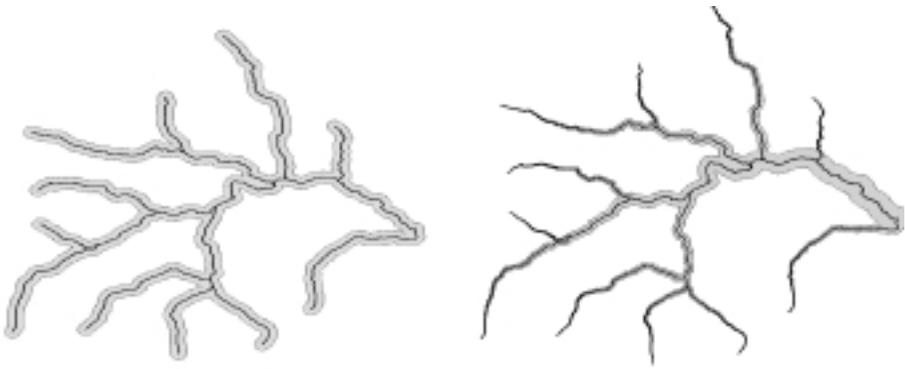
Fig. 3

within a given distance from the polyline riverbed). However, this approach does not meet our needs, because the buffer created using these tools maintains the same width along whole polyline length (Fig. 3).

Therefore, the above mentioned tools couldn't be used, because they compute distance from river without taking in account linear distance along river, so they do not comply with our model. To satisfy requirements of the proposed model, the general buffer tool had to be adjusted.

The common ArcGIS buffer tool enables to set the width of the buffer for each object of a layer based on a defined field from attribute table. This proved very useful, it did not, however, solve the whole problem. It was necessary to pre-process the input river polyline data so that first each river in the river network layer was split into short segments of equal length. Next, an average distance from river source (measured along river) for each river segment was computed, and stored in the attribute table. Based on this field, the width of buffer for each river segment was computed.

2.3 Relative height of floodplain above river level

The last indicator used for the floodplain delineation is the relative altitude of flood plain above the river level. For practical reasons, this surface was constructed in form

DEM
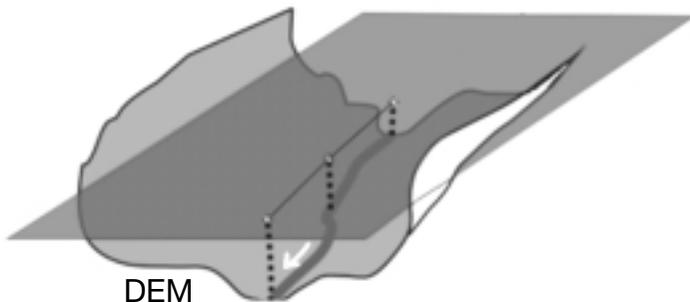
Fig. 4

of raster. Similarly as in the case of the distance from the riverbed, this part of the model also assumed that the relative expected floodplain height is minimal near the source and increases continuously towards river mouth. The main task was therefore to construct a surface (raster grid) that fulfils this condition. This surface should vary its relative height in direction along river, but the relative height above river should stay same in direction perpendicular to river geometry (Fig. 4).

From available tools, the IDW algorithm fitted best our requirements. Last step in our workflow was the creation of a point layer, which was then used as input to IDW tool. This point layer was generated based on the input river network layer so that each created point lies on the polyline, and the points are distributed at a constant distance. For each point, distance to the river source along the river polyline was computed and stored again in the attribute table. Based on this distance, the relative height of flood plain above level of river was estimated as a relevant fraction of a maximum value valid for the river mouth.

3 Results

3.1 Construction of the models

A methodological part of the results of this work focused on the creation of models or scripts, which would enable automatization of the floodplain delineation process. The three main scripts were created to automatise:
- creating linearly referenced river route layers
- creating a set of buffers with progressively increasing width
- creating a raster delineating maximum expected height of floodplain above level of river (progressively growing)

3.1.1 Linearly referenced river route

As mentioned in the previous chapter, creating of a correctly linearly referenced river network is essential for computation of all other parameters for floodplain delineation.

In the first stage, we have tried to solve the task using the original tools from the ArcGIS. There is a Create Routes tool that enables automatic creation of a route feature in the ArcGIS environment. As was mentioned above, this tool cannot be used for all river polylines in a layer using the same settings. Therefore, prior to using the Create Routes tool, new field in the river network layer was created containing information on settings necessary for the particular object, based on the position of source and sink and minimal Bounding Box of river (more in Hartvich and Jedlička 2007, ESRI 2004, see Fig. 5).

Then, the rivers with the same code value were selected and on this selection, the Create Routes tool with appropriate parameters was run (parameters were set based on code value to one of possibilities UPPER_LEFT | LOWER_LEFT | UPPER_RIGHT | LOWER_RIGHT). Finally, all results were combined into one layer using the Append tool.

Direction of river flow
Badly determinated direction
Bounding Box

Fig. 5

Input height raster

Input river network (Polyline M)

Convert lines end points

Heigth from raster to points

Creating new empty feature

Copy items from input

Is first vertex of featur source

Accumulate distance in forward dir.

m = 0
m = 100
m = 200

Accumulate distance in backward dir.

m = 1600
m = 1500
m = 1400

Exist next feature

Fig. 6

It can be seen that this process of creation of the route feature described in previous paragraph is too clumsy and complicated, thus we focused on trying to perform this process more simply and also faster. The solution, however, required going deeper into the geometry properties of a shapefile and also some more advanced programming methods had to be employed. Programming was realized using the Python scripting language, which is widely supported by ESRI and is used for creation of geo-processing scripts. The overall scheme of the process is show on Fig. 6

3.1.2 Raster of buffers with progressively increasing width

For automatic creation of a buffer with progressively increasing width, the model for automatized workflow was created also using the Model Builder.

The first step in creating the progressive buffer was splitting of river (previously converted to route feature), to several line features with equal, pre-defined length. For this purpose, the tool Make Route Event Layer from Linear Referencing toolbox was used. Prior to application this tool, however, a table describing the distance from the source had to be created using a custom Python script.

The rest of the steps were done using standard ArcGIS tools. This included the computing of the average distance of each segment from to the river source. Next, the width of the buffer was computed based on that information and mathematical expression, which describes the dependency between buffer width and the distance from river source.
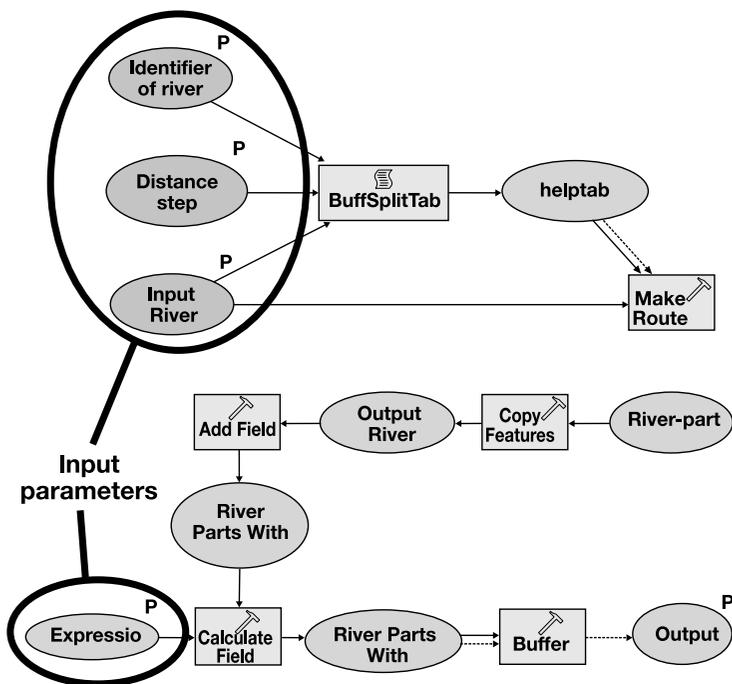
Fig. 7

Finally, the buffer was created using all segments of all input rivers. Process workflow is illustrated in Model Builder schema (Fig. 7). The model was also converted into a Python script, to enable better workflow control and better portability of this custom tool.

3.1.3 Raster describing the maximal relative height of floodplain

In the case of this model, the general approach was quite similar to the one used in the creation of the progressive buffer model. The whole model is shown on Fig. 8. First input, the river network layer (already m enabled), was converted into a set of points, which were separated by same predefined distance. A table, describing position of
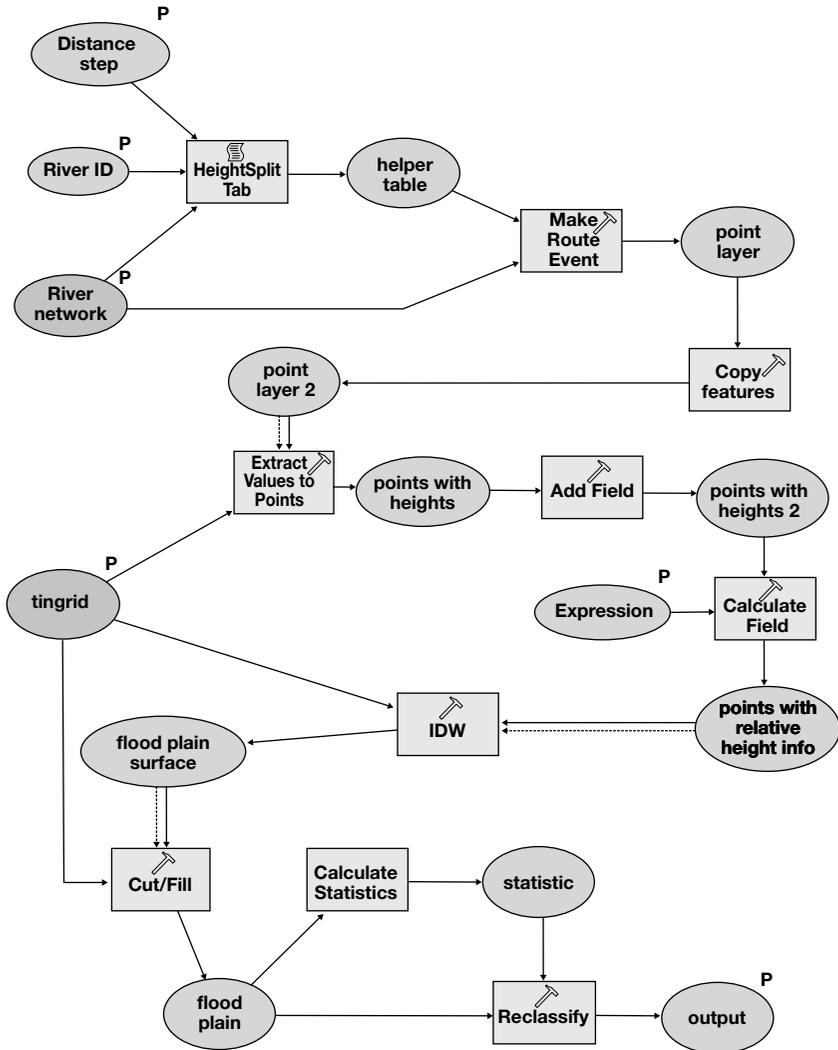
Fig. 8

these points on river in sense of linear referencing was created using a Python script. Then, a tool "Make Event Layer" was used to convert river network to points. At each point, the height information was derived from input height raster. After that expected height of flood plain was computed, using distance from river source, input math formula and height of that point derived from raster. This information was used to create a surface, describing maximal expected height of floodplain along river (using IDW tool). This surface and height raster were used as input to Cut/Fill tool (see Cut/Fill documentation). Based on output of this tool, area which is lower than surface describing maximal expected height of flood plain was defined. Also this Model Builder model was also converted into a Python script for better compatibility.

3.2 Application of the model in the Opava R. catchment

The above described methodics was applied on the streams in the catchment of the Opava River (above Opava town) in NW part of Moravia. Opava R. and its tributaries spring on the eastern slopes of the Jeseníky Mts., and flows generally to the southeast with a segment between Nové Heřmínovy and Krnov (Fig. 9), where the course turns suddenly to the NE, likely due to the morphostructural reasons.

General pattern of the morphology of the valley and floodplain is typical for the rivers springing in the mid-mountains and flowing to the flat basins: at first, the narrow floodplain occupies the whole floor of a deep, steeply incised valley. At the middle
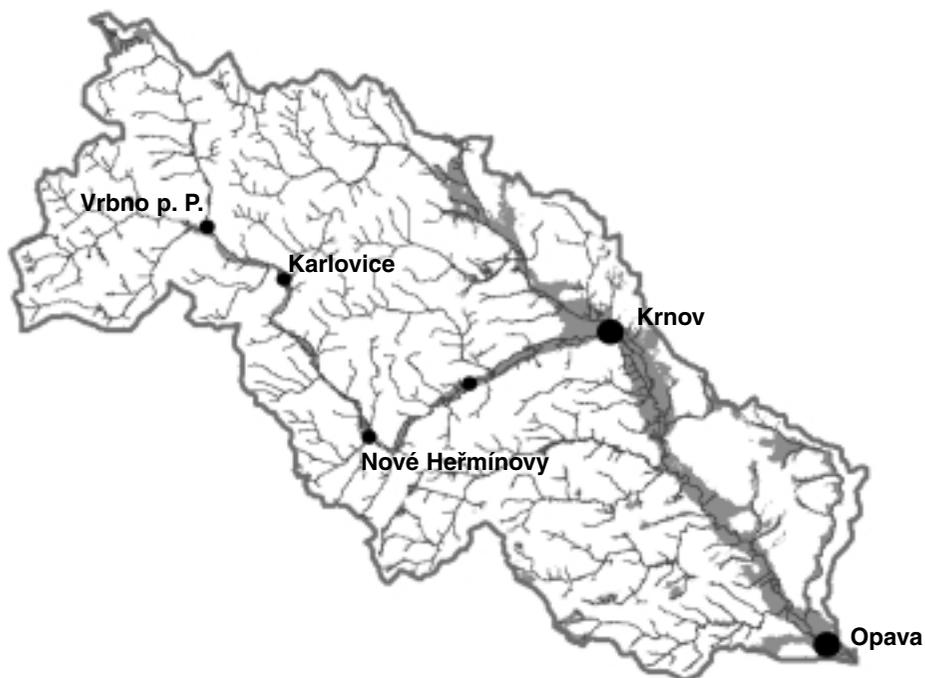
Fig. 9

course, approximately between Vrbno and Krnov, the valley opens and floodplain grows wider, up to 400–500 m, though locally the width can due to lithological or structural causes be as narrow as 100 m. Under Krnov, the floodplain widens up to its maximum values of almost 2.5 km, again locally limited to lower width, as in Krnov (Fig. 9). Here, the floodplain limit is also the most obscured by anthropogenic activities, such as the bodies of communication ramparts or industrial areals.
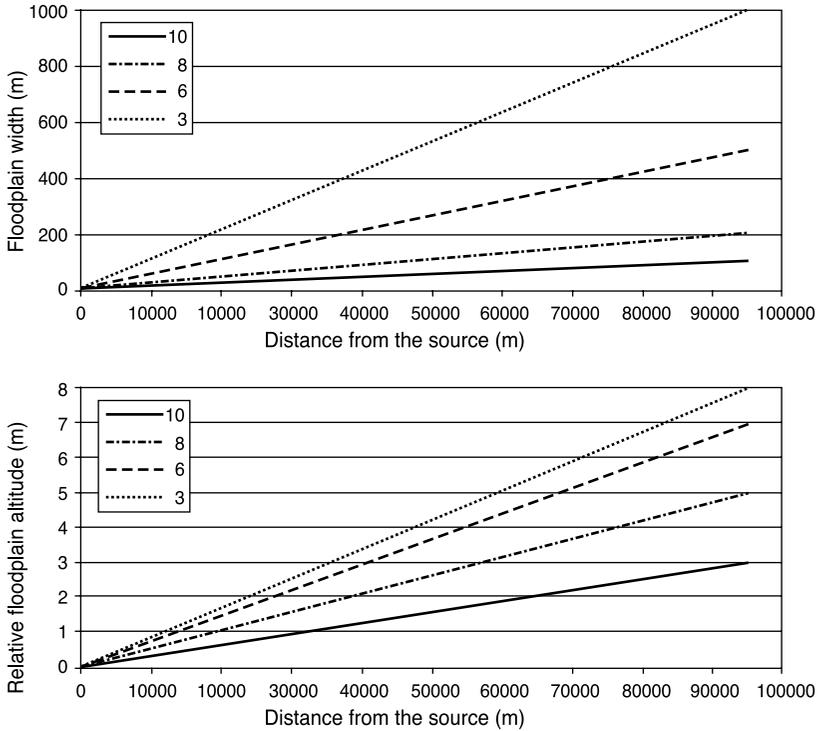
Fig. 10
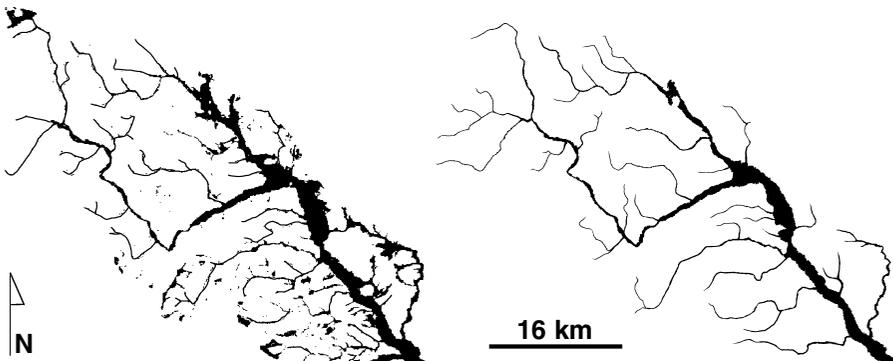
Fig. 11

Table 1: Values assigned to the parameters for floodplain delineation, SP model (left) and PCP model (right)

Relative height above the river level (m)		Slope inclination (°)		Relative height above the river level (m)*		Slope inclination (°)		Buffer width (m)*	
interval	value	interval	value	interval	value	interval	value	interval	value
0–3	10	0–1	10	0–3	10	0–1	10	0–100	10
3–5	8	1–2	8	3–5	8	1–2	8	100–200	8
5–7	6	2–3	6	5–7	6	2–3	6	200–500	6
7–8	3	3–5	3	7–8	3	3–5	3	500–1000	3
8 +	0	5 +	0	8 +	0	5 +	0	1000 +	0

* In the case of the PCP model the interval extents are valid only for the closing profile

The settings of the input parameters used in the case of the Opava and its tributaries for both models (SP and PCP) are shown in Tab. 1. The progressive parameter change function applied was linear, the particular functions for each interval of relative height and buffer width are presented on the charts in Fig. 10. Using these settings, a raster of floodplain was calculated for each of the two models (Fig. 11).

4. Discussion and evaluation of the model

4.1 Verification of the PCP model performance

The standard procedure of the verification of the floodplain delineation should include comparison of the floodplain extent from the model with a field-mapped extent in several various and morphologically different control segments (Andrysiak and Maidment 2000).
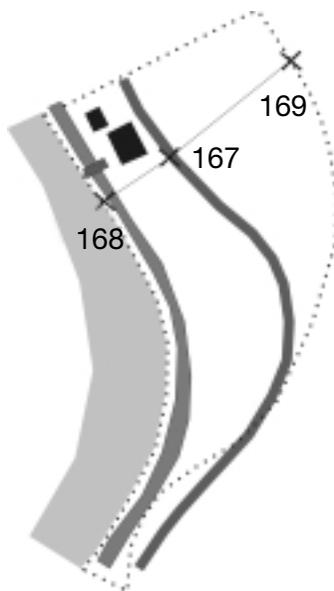
Fig. 12
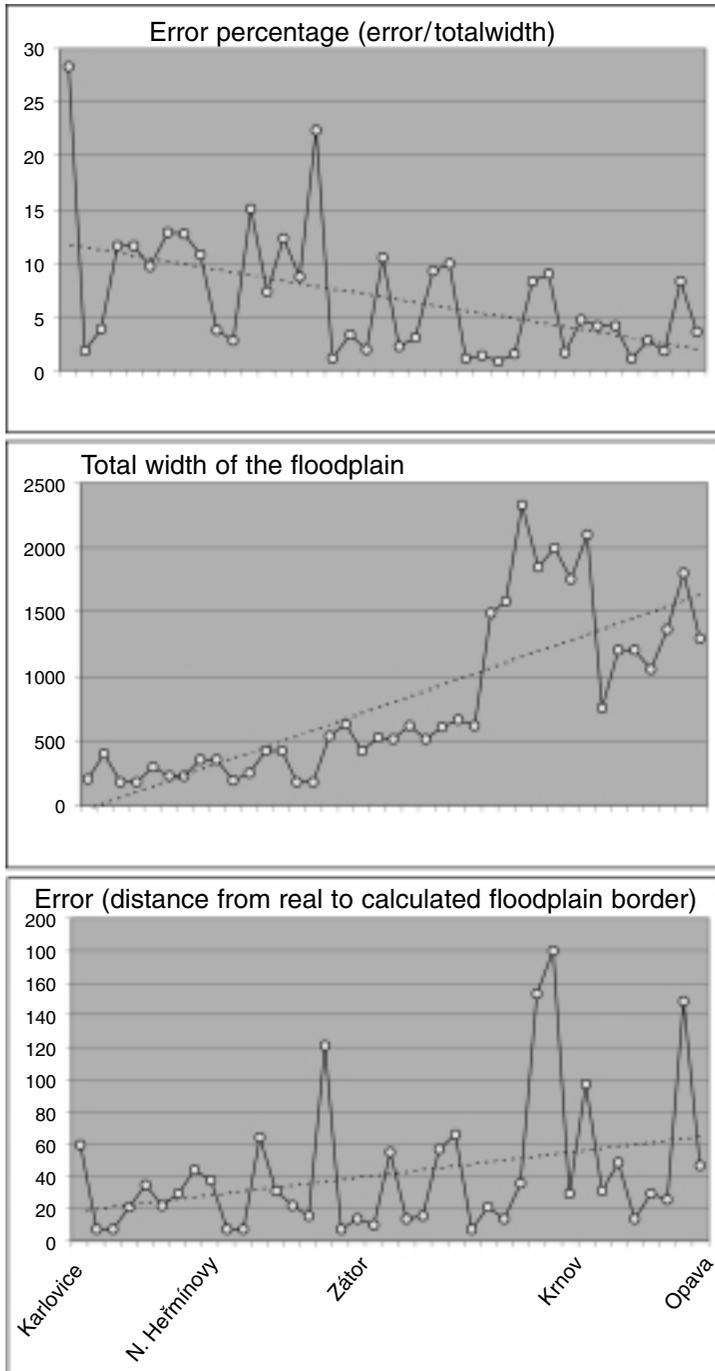
Fig. 13
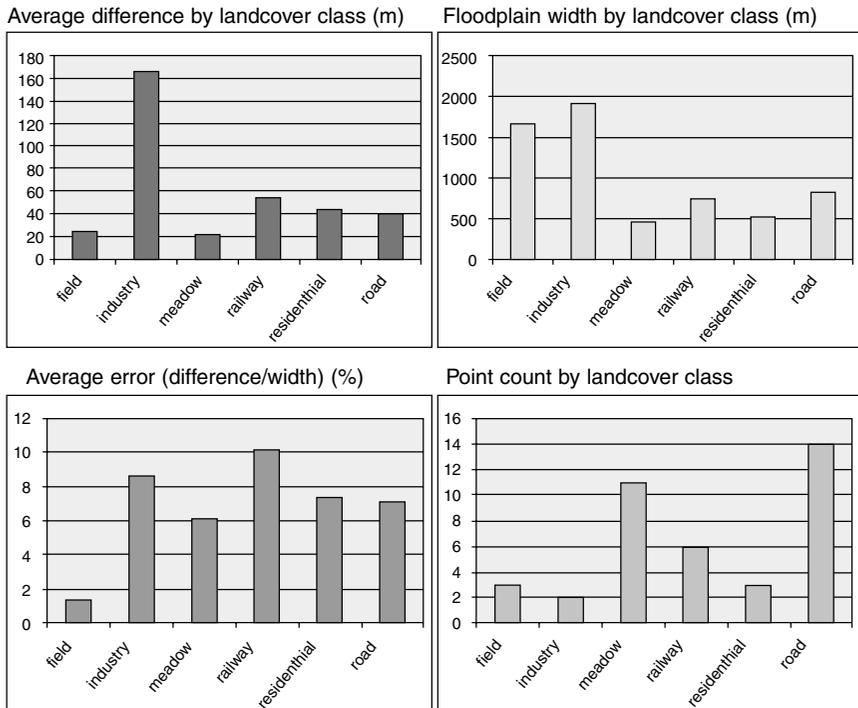
Fig. 14

However, as the complete field mapping results from the Opava are not yet available, the results of the model were checked in the field on several localities. Altogether on 17 sites along the course from Karlovice to Opava town were investigated, sketched (example of the control sketch is on Fig. 12) and position of the floodplain limit was recorded using a GPS on 39 control points. The GPS points recorded on the field-detected limit of the floodplain were embedded into the ArcGIS environment.

For each control point (i.e. real floodplain limit), a distance to the floodplain limit delineated by the model was calculated, total width of the floodplain was measured and landcover noted. The results are shown in the overview charts (Fig. 13 and 14). The highest errors are connected to artifical structures in the landscape, particularly the high industrial and communicational earthworks.

As concerns the distribution of the performance, with wider floodplain along the river course slightly increase the absolute error values, while percentage decreases. The cause of this fact can lie in two factors: the border of the floodplain is less sharp in flatter landscape of lower segments, and the wide floodplain is more affected by human activity and therefore there are more objects, obscuring the floodplain borderline.

Overall performance of the PCP model is satisfactory, the average error in the floodplain extent is lower than 7%, which represents distance of 40 m, i.e. 4 raster pixels. As the mean slope in the floodplain reaches only slightly over 1°, the vertical error would in average be lower than 0.8 m, which, considering the accuracy of the DEM, is rather surprisingly low value.

The performance of the two models, the original (SP) and the new one using the progressive change of parameters (PCP), was compared in the catchment of the Opava R. Several areas, which were due to some characteristic relief parameters typically problematic in the previous model development stages, were selected and the result floodplain rasters from both models were analysed.

The problematic areas were of several types. In the source and upper segments of the rivers, the low watershed divides and generally flat relief do not tally with the narrow floodplain of the small streams. The SP model usually tends to highly overestimate the floodplain extent in such areas. As can be seen from the Fig. 15, the PCP model resolves the problem, as the flat relief is counterbalanced by the low values of buffer and relative height extent.
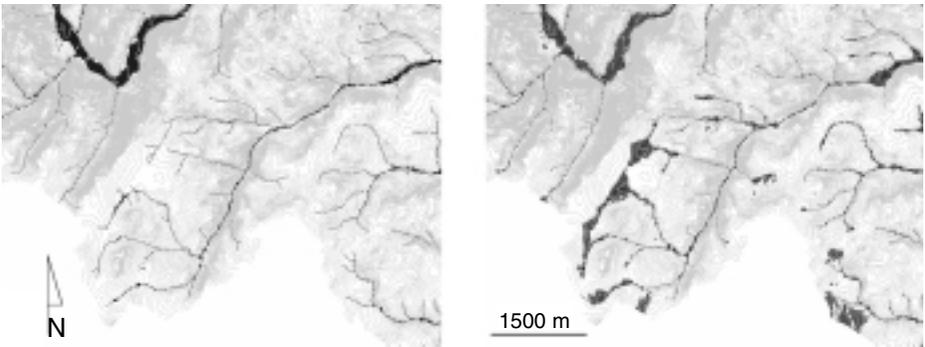
Fig. 15
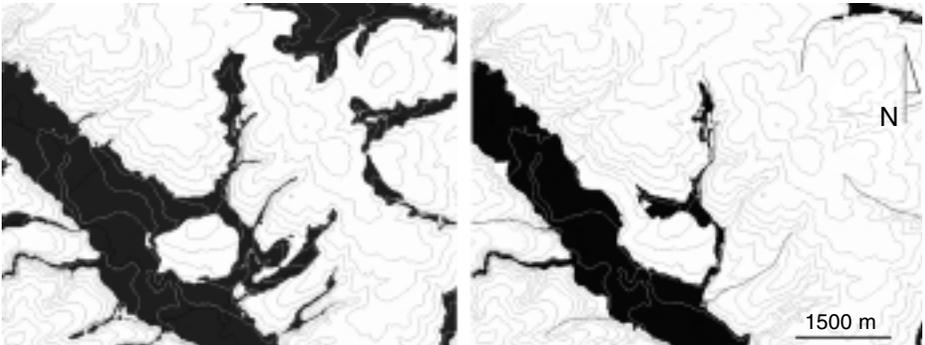
Fig. 16

Other type of problematic relief morphology is the overestimating of the floodplain of smaller streams in the flat parts of the lower course. The PCP model reduces this problem as the parameters are calculated from the source of these smaller streams, thus the floodplain remains restricted to a size natural for these streams (Fig. 16).

The model performance depends on many factors, the sources of errors come mainly from three main sources: from the input data character and quality (both spatial inaccuracy and temporal inactuality), from the intrinsic model simplifications and from the subjective decisions of the operator, necessary in the process.

Primarily, it is the eternal source data quality problem. As the model performance is naturally directly dependant on the accuracy of the input data, it can never exceed the accuracy of the DEM. Secondly, the character of the relief and also anthropogenic interference in the floodplain decreases the accuracy of the modelled floodplain. In general, it can be stated that the more natural state of the floodplain, the better are the results of the model. This is due to two factors: anthropogenically induced changes are both new (therefore not included in the maps) and sharp (thus the gradual interpolation-based models fail to mark their exact borders). Particularly industrial zones and communication ramparts, not contained in the DEM and topological data, are the sources of the major errors.

Another problematic point is the setting of the probability values and distribution for each given width of the buffer / relative height interval values and the function of the growth of these parameters along the course. The higher point value, the more likely is the given pixel to be a part of the floodplain (Hartvich 2007). So far, it is necessary to asses the values manually, based on the expert knowledge of the operator, which may, however, introduce subjectivity and thus possible error into the process.

The possibilities of workaround for this problem are limited. However, we are currently working on two solutions which would decrease the subjective factor:

i) to calibrate the process based on field-collected data. In particular, the information from several geodetic profiles on different sites along the course might be used to calibrate both the interval point values and the formula of the progressive parameter change function. Thus it will be possible to reduce the uncertainty caused by the subjective setting of the parameters.

ii) to design a user-friendly interface of the model, which would allow the user to test quickly various settings of the model parameters and to decide – based on comparison of the performance – about the most suitable settings for a given catchment.

There are also certain problems in the model scripting left. Even though the script for river routing has shown improvement in stability and reliability during the route river layer creation, there are, however, still some limitations. One condition for the input line layer is that each single river polyline must be represented by just one line feature (one river can not be composed of several line features). To fulfil this condition, the Dissolve tool must be used prior to the calculation. Another condition which must be met is that no feature can be a multipart feature. Handling the multipart features is not supported in this stage of development, we consider, however, including the support for multipart features in the future model development.

However, the above mentioned drawbacks are either inevitable (input data shall influence any such model) or are balanced by the advantages of this attitude to the floodplain delineation. These include fast coverage by large catchments, repeatability

of the simulation, reasonably accurate results and possibility of cross-control of the field mapping.

5. Conclusion

It can be concluded that the current model represents another step towards the real situation and thus to the reliability of the resulting floodplain. As was illustrated on the analysis of the results on the example of the Opava R., the PCP model performs better particularly in the previously problematic areas. The analysis of the field control shows that the overall performance of the PCP model is satisfactory; the average error in the floodplain extent is lower than 7%, which represents distance of 40 m, which represents only 4 raster pixels at the resolution of 5 m.

The main result of the of the present stage of the model development is represented by a functional set of tools, able to delineate the extent of the floodplain with considerable accuracy. The particular improvements in the current version are in progressive changing of the input parameters and in partial automatisation of the process.The next development of the model will aim particularly on complete automatisation and debugging of the procedure and on the lowering of the subjective interferences into the model.

References

ANDERSON, D.J. (2000): GIS-based hydrologic and hydraulic modelling for floodplain delineation at highway river crossings. MS, University of Texas at Austin, http://www.crwr.utexas.edu/online.shtml

ANDRYSIAK, P.B. AND MAIDMENT, D. (2000): Visual Floodplain Modeling with Geographic Information Systems (GIS). MS, University of Texas at Austin, http://www.crwr.utexas.edu/online.html

ESRI (2004): Writing Geoprocessing Scripts With ArcGIS. ESRI 380 New York Street Redlands, California 92373-8100, USA, 88 pp.

ESRI (2007): Virtual Campus – Getting Started with Scripting in ArcGIS 9. http://training.esri.com

ESRI (2007): Virtual Campus – Creating and Editing Linearly Referenced Features with ArcGIS 9. http://training.esri.com

HARTVICH, F. (2007a): Hodnocení morfometrických parametrů nivy, vymezení nivy v GISu a využití příčných profilů. In: Langhammer, J. (ed.) a kol. (2007): Povodně a změny v krajině. Monografie, KFGG, PřF UK, Praha, 2007.

HARTVICH, F. (2007b): Character and dynamics of the floodplain of the Losenice River, Šumava Mts. Silva Gabreta, NP Šumava Authority, Vimperk (v tisku).

HARTVICH, F., JEDLIČKA, J. (2007): Metodika vymezování údolní nivy v prostředí GIS. In: Langhammer, J. (ed.): Sborník z konference „Povodně a krajina", 5. 6. 2007, PřF UK Praha.

HUDSON, P. F. AND COLDITZ, R.R. (2003): Flood delineation in a large and complex alluvial valley, lower Pa'nuco basin, Mexico. Journal of Hydrology, Volume 280, Issues 1–4, September 2003, 229–245.

KŘÍŽEK, M., HARTVICH, F., CHUMAN, T., ŠEFRNA, L., ŠOBR, M., ZÁDOROVÁ, T. (2006): Floodplain and its delimitation. Sborník ČGS – Geografie 111 (3), 260–273.

LEWIN, J. AND MANTON, M.M.M. (1975): Welsh floodplain studies: The nature of floodplain geometry. Journal of Hydrology, Volume 25, Issues 1–2, April 1975, 37–50.

TATE, E. AND MAIDMENT, D. (1999): Floodplain Mapping Using HEC-RAS and ArcView GIS. MS, University of Texas at Austin, http://www.ce.utexas.edu/centers/crwr/reports/online.html

TOWNSEND, P.A. AND WALSH, S.J. (1998): Modeling floodplain inundation using an integrated GIS with radar and optical remote sensing. Geomorphology 21 (1998) 295–312.

R é s u m é

Progresivní nárůst vstupních parametrů při vymezování údolní nivy na základě DMR – aplikace a hodnocení modelu na příkladu povodí Opavy

Pro vymezení říční nivy byla vyvinuta procedura, která na základě morfometrických dat (vzdálenosti od říčního koryta, relativní výšky nad hladinou řeky a sklonitosti svahů) umožňuje připravit výstupy pro určení rozsahu a polohy údolní nivy. Výhodou tohoto přístupu je malé množství požadovaných vstupních dat. V této publikaci jsou porovnávány výsledky dvou vývojových fází modelu, se stálými a s proměnnými vstupními parametry šířky bufferu vzdálenosti od koryta a relativní výšky nivy. Model je implementován v prostředí ArcGIS jako sada skriptů v jazyce python. Tyto skripty umožňují automatizaci jednotlivých dílčích úkonů nutných pro vymezení říční nivy. Jako vstupní data jsou požadovány pouze liniová vrstva říční sítě a digitální model terénu ve formě rastru. Jediným vstupem, který je nutné nakalibrovat, je parametr, určující rychlost progresivního růstu relativní výšky/vzdálenosti v závislosti na vzdálenosti podél toku. Výsledným rastrům progresivního bufferu, relativní výšky nivy a sklonitosti svahů je přiřazena relativní hodnota, určující pravděpodobnost výskytu nivy. Součtem těchto rastrů vznikne kvantitativní vyjádření pravděpodobnosti výskytu nivy. Díky automatizaci celého procesu do několika skriptů však může řešitel opakovaně určit oblasti pravděpodobného výskytu nivy s různými vstupním parametry progresivního nárůstu šířky / relativní výšky, a následně zvolit variantu nejlépe odpovídající realitě, ověřené například na vybraných profilech v terénu.

Filip Hartvich
Ústav struktury a mechaniky hornin AV ČR, v.v.i.,
V Holešovičkách 41
182 09 Prague 8,
Czech Republic
e-mail: hartvich@irsm.cas.cz
phone: +420 266009233