Using of local indicators of spatial association for evaluation of spatial accuracy of DEM

Jana Svobodova[1], *JakubMirijovsky*[1], *Ales Vavra*[1], *Jan Brus*[1] *and Helena Kilianova*[1]

[1] Palacky University in Olomouc, tr. Svobody 26, Olomouc, 771 46
j.svobodova@upol.cz, jakub.mirijovsky@upol.cz, ales.vavra@upol.cz, jan.brus@upol.cz, helena.kilianova@upol.cz

Abstract

In the study, whose results are presented in the paper, the statistical method LISA (local indicators of spatial association) has been used for the delimitation of the areas with the statistical significant error values. The error values obtained by subtraction of estimated values of altitude (DEM) from the reference surface were used as the input data for the local spatial cluster analysis. The aim of an interpretation of the results of local cluster analysis was to determine the behaviour of errors (location and size) from the perspective of different quality DEMs, i.e. DEMs created by different interpolation methods and their settings.

The main significance of LISA lies in the spatial expression that allows you to monitor the extent, distribution and overall structure of clusters. When comparing the results derived from selected high-quality and low-quality DEM or DEM created by the different interpolation methods it is necessary to realize the different importance of these three aspects. This was reflected in the creation of rules for interpretation and evaluation the results of local cluster analysis, which is one of the main results of the study.

Keywords: DEM, LISA, clusters, statistical significance, spatial accuracy.

1. Introduction

A relief is a key factor in many environmental processes. It has a significant impact on climate variables like temperature and precipitation. It also influences a degree and spatial extent of weathering, erosion or accumulation on the Earth's surface etc. (Moore *et al.*, 1991; Wilson and Gallant, 2000; Arrell, 2005). Within a GIS environment, digital elevation models represent the real image of a relief. The most quality DEM must be used to get best quality of the results from environmental analysis.

The DEM quality (in terms of metric accuracy) mainly depends on the underlying data source and interpolation method (Weng, 2002). In connection with the underlying data the method of collecting these data affects the quality of the resulting DEM. Even if it is chosen the best source of elevation data it is not guaranteed accurate DEM on the output, because for its generation it is also necessary to choose a suitable interpolation method and setting of its parameters. Systematic and unsystematic errors can occur during the process of the DEM creation. These errors can negatively influence the next computation of the morphometric attributes.

There exists couple of methods for the determination of the inaccuracy in the DEM. The fast way how to uncover the majority of the errors is the visual control of the DEM. This leads to the recognition of the errors caused by the deficient in the vertical resolution and the problems caused by the local abnormality. There are also other more sophisticated ways how to compare the resulting DEM. These methods are based on the statistical approach. Among the mostly used methods belong root mean square error (RMSE), total absolute error (AE) or hammock index (H). The root mean square error (RMSE) expresses an extent to which an interpolated value differs from a real value. A higher value corresponds to a greater difference between two datasets (Wood, 1996). The total absolute error (AE) shows a real sum of all deviations from reference data in both positive and negative directions (Svobodová *et al.*, 2009). The hammock index (H) examines the regularity of location of interpolated values between the known values, as well as excessive incidence of pixels whose altitude tally with original data values (Wood, 1996).

These common methods used for assessment of a metric accuracy of DEMs (like RMSE, AE, H) are only global non-spatial measures. Their unquestionable advantages are their easy calculation and interpretation. However, as to environmental applications, the knowledge of spatial error variability is very important. Many authors (Desmet, 1997; Erdogan, 2009; Fischer, 1998; Gao, 1997; Hofierka *et al.*, 2007; Hunter and Goodchild, 1997; Pechanec *et al.*, 2011; Svobodová, 2010; Wise, 2000) claim that the extent of errors in a DMR depends on a character of a terrain and is spatially variable.

Absolutely the easiest way how to get the value of local deviations is the subtraction of estimated values of altitude (DEM) from the reference surface, if available, or from reference points. These differences can then be used for example to calculate the total absolute error (Svobodova *et al.*, 2009). Gousie (2005) or Erdogan (2009) states other examples (methods) to express the spatial distribution of error. The basic problem of the above mentioned methods is mere spatial visualization of absolute or relative values without any determination of significance of the size of errors. These surfaces, however, can be used by local cluster analysis (LISA) to determine the values of errors and locate statistically significant clusters of high error values, clusters of low error values or so-called outlier.

2. Application of LISA

To assess DEMs spatially, a local cluster analysis LISA which enables to observe a spatial distribution of errors can be applied. In contrast to a mere representation of size of errors in form of an accuracy surface (gained by a mere difference between estimated data and reference data), which does not allow for exact determination of boundaries between high and low error values, a local cluster analysis determines statistically significant high and low error values (or outlying values) accurately enough (Figure 1). The accurate determination can be further used for a research into relationships between occurrence of high error values (or outlying values) in DEMs and values of morphometric parameters derived from these DEMs (Svobodová, 2011).

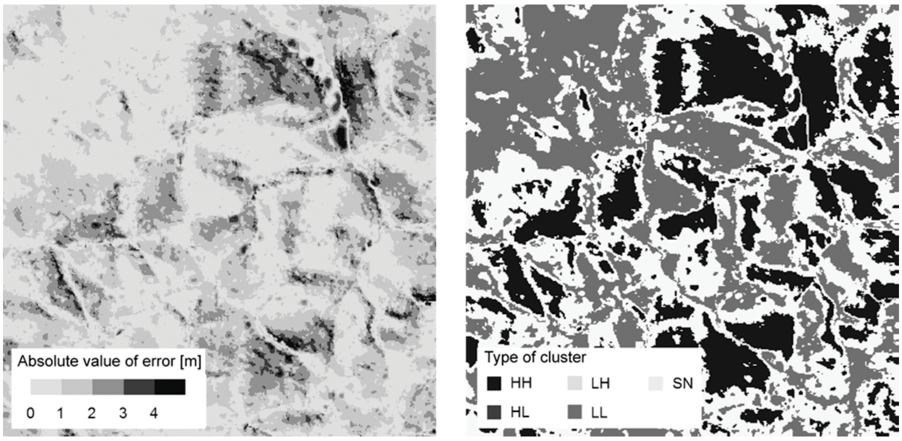

Figure 1: Comparison of common accuracy surface with absolute error values (left) and localization of statistically significant clusters of high (HH) or low (LL) values gained by LISA (right) - the example of the Rusavská highland

The aim of spatial assessment of metric accuracy of DEMs is to verify the results of non-spatial assessment. However, even if the local cluster analysis LISA is used to assess the spatial metric accuracy of DEMs it is necessary to use some non-spatial values. It came out during the evaluation of LISA results that when interpreting results it is desirable to focus not only on a spatial visualization of clusters of statistically significant high or low error values, but also on quantitative (or sometimes global) values of indicators like a range of values in a set of errors, a mean error value or maximal value in the clusters of high error values. On the basis of assessment of spatial extend of statistically significant clusters of errors together with using of above-mentioned quantitative indicators the rules for assessment of quality of DEM (at one area) in the terms of spatial accuracy were made and they are state in Table 1 (Svobodová, 2011).

Table 1: Evaluation of quality of DEMs according to a spatial assessment of LISA results supplemented with quantitative non-spatial indicators

Area of clusters at comparison of two DEMs by visual evaluation	Range of error values, mean value of errors, max. value in HH* clusters	Quality of DEM
bigger	higher	less-quality
	lower	more-quality
smaller	higher	less-quality
	lower	more-quality
comparable	higher	less-quality
	lower	more-quality
	comparable	comparable

*cluster of high values (high error values)

It is possible to demonstrate a need to supplement spatial assessment of quality of DEMs with quantitative non-spatial indicators on the results of local cluster analysis of highland relief (Rusavska highland RH). The results of LISA are in case of the high-quality DEM compact and aerially more extensive clusters of high or low errors. As for the low-quality DEM, the clusters are aerially less extensive and their types vary in areas more often (see Figure 2). Interpretation by visual evaluation only could lead us to the conclusion that there are more statistically significant errors in the DEM which is considered to be high-quality one than in the lower-quality DEM (Figure 3). Nonetheless, this way of differentiating between high-quality and low-quality DEM is not correct.

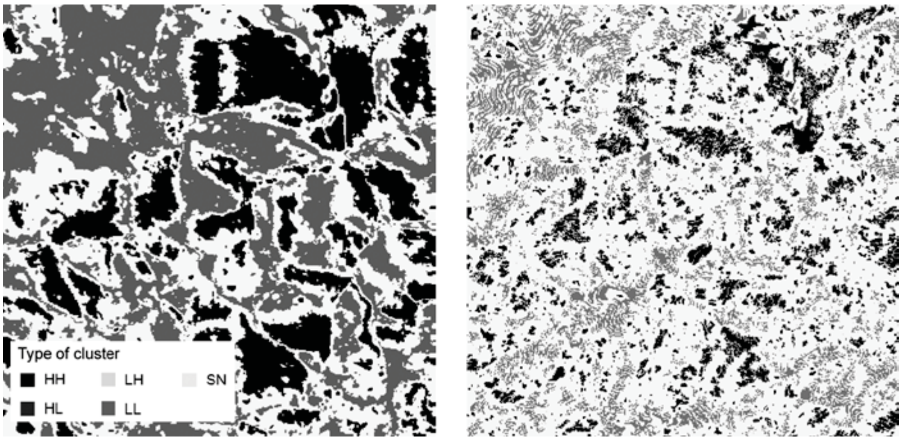

Figure 2: Typical structure of clusters in results of LISA derived from the high-quality (left) and low-quality (right) DEMs of highlands (the example of the Rusavska highland)

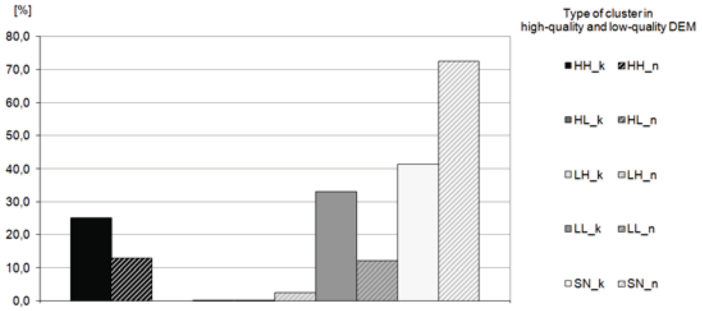

Figure 3: Percentage rate of features that belong to the individual types of clusters with the use of high-quality and low-quality DEMs (the example of the Rusavska highland)

When considering the various extent of error values, different mean error values in files entering the local cluster analysis and different size of maximal value in the clusters of high error values (Table 2), however, we find out that clusters of high errors in high-quality DEM which at first sight appear to be more extensive, in reality cover much lower values of errors than aerially smaller clusters of high errors derived from low-quality DEM. Moreover, taking into account the structure of clusters, it can be supposed that more compact shapes allow for a smoother change of error values, which can preserve shape fidelity of an actual relief. On the contrary, a frequent change of little clusters of low and high errors can bring about occurrence of local minims and maxims in DEM resulting in prevention of fluent outflow.

Table 2: Range and mean value of absolute error values gained by subtraction of high-quality (HQ) or low-quality (LQ) DEM from referential DMR and maximal value of error in clusters of statistically significant high error values (max. HH) - the example of the Rusavska highland

range [m]		mean [m]		max. HH [m]	
HQ	LQ	HQ	LQ	HQ	LQ
6,53	17,42	1,22	1,53	6,53	17,42

3. Conclusion

The using the LISA method allows you to accurately determine as in value as in space the high error values (for a given level of statistical significance). This gives it a huge advantage over other methods used for spatial assessment of the accuracy of DEM. The inclusion of the selected global characteristics such as a range of values in a set of errors, a mean error value or maximal value in the clusters of high error values is necessary for the correct interpretation of the results.

Acknowledgments

This paper was created within the project "The small format aerial photography in the study of the effect of surface heterogeneity on the habitats" (project number PrF_2012_007) with the support of Internal Grant Agency of Palacky University in Olomouc.

References

Arrell, K. (2005), *Predicting glacier accumulation area distributions*. Durham University.

Desmet, P.J.J. (1997), “Effects of interpolation errors on the analysis of DEMs”. *Earth surface processes and landforms*, Vol. 22:569-580.

Erdogan S. (2009), “A comparsion of interpolation methods for producing digital elevation models at the field scale”. *Earth surface processes and landforms*, Vol. 34:366-376.

Fisher, P. (1998), “Improved Modeling of Elevation Error with Geostatistics”. *GeoInformatica*, Vol. 2(3):215-233.

Gao, J. (1997), “Resolution and accuracy of terrain representation by grid DEMs at a micro-scale”. *International Journal of Geographical Information Science*, Vol. 11:199-212.

Gousie, M. B. (2005), “Digital Elevation Model Error Detection and Visualization”. *In ISPRS - The 4th Workshop on Dynamic & Multi-dimensional GIS*, Wales, UK, pp. 42-46.

Hofierka, J., Cebecauer, T., Šúri, M. (2007), “Optimisation of Interpolation Parameters Using a Cross-validation”. *In Digital Terrain Modelling, Development and Applications in a Policy Support Environment*, Series: Lecture Notes in Geoinformation and Cartography, Springer, ISBN: 978-3-540-36730-7, pp. 67-82.

Hunter, G. J., Goodchild, M. F. (1997), “Modeling the uncertainty of slope and aspect estimates derived from spatial databases”. *Geographical Analysis*, Vol. 29:35-49.

Moore, I. D., Grayson, R. B., Landson, A. R. (1991), “Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications”. *Hydrological Processes*, Vol. 5:3-30.

Pechanec, V., Brus, J., Kilianova, H. (2011): “Modelling spatial distribution of ecotones in GIS”. *In: Conference Proceedings SGEM 2011, 11th International Multidisciplinary Scientific GeoConference STEF92 Technology Ltd.*, Sofia, Bulgaria, pp 637-644, ISSN: 1314-2704.

Svobodová J., Tuček P. (2009), “Creation of DEM by kriging method and evaluation of the results”, *Geomorphologia Slovaca et Bohemica*, Vol. 9(1):53-60.

Svobodová J., Voženílek V. (2010), “Relief for Models of Natural Phenomena”. *In Landscape Modelling: Urban and Landscape Perspectives*, Part V., Vol. 8, Springer Science + Business Media B.V., pp. 183-196.

Svobodová J. (2011), *Quality assessment of digital elevation models for environmental applications*. Ph.D. thesis, Department of Physical Geography and Geoecology, Faculty of Science, University of Ostrava. Czech Republic.

Weng, Q. (2002), “Quantifying Uncertainty of Digital Elevation Models Derived from Topographic Maps”. *In Symposium on Geospatial Tudory, Processing and Applications*. Ottawa, pp. 403-418.

Wilson, J.P., Gallant, J.C., eds. (2000), *Terrain Analysis: Principles and Applications*. John Wiley & Sons, ISBN 0-471-32188-5.

Wise, S. (2000), “Assessing the quality for hydrological applications of digital elevation models derived from contours”. *Hydrological processes*, Vol. 14:1909-1929.

Wood, J. D. (1996), *The geomorphological characterisation of digital elevation models*. PhD thesis, Geography Department, University of Leicester, UK.