

Assessment of the quality of digital terrain models with global and local approaches: a case study using fine spatial resolution data from alternative sources

Michal GALLAY

Abstract: This paper assesses absolute accuracy of four different digital terrain models (DTMs) which were generated from data which acquired by four modern technologies. The accuracy was measured as differences between a ground truth DTM and the four evaluated DTMs. A range of methods for characterizing the data distribution of elevation residuals surfaces was applied. The objective is to promote the use of local and spatial methods of geographical data analysis.

Keywords: elevation residuals, global, local, spatial, aspatial

Background and motivation

Nowadays, the elevation data can be collected by a variety of methods including ground surveying and remote sensing techniques. The latter are capable to collect the data quicker, on a finer scale and with a higher degree of accuracy than ever before. Digital terrain models can be easily created from secondary data sources such as topographic maps. Furthermore, there is also a wide range of ready-to-use elevation data available for anyone to purchase and use, usually in the form of digital elevation model (DEM). Depending on the application, the users are interested in digital terrain (bare earth) surface models (DTMs) or models which include the surface of the vegetation canopy and man-made structures (digital surface models, DSMs, see *Fig. 1*). The Central and Eastern European geography prefers a digital model of (geo)relief (DMR) instead of a DTM. After Krcho, J. (2001), georelief surface is more or less a synonym of terrain and it refers to the surface of lithosphere and pedosphere.

Tab. 1. Specifications of the analysed data sets

Acquisition method, year	Data Type	Data format	Spacing	Vertical accuracy	Supplier and copyright
LiDAR 2001	LR points	points	≈ 2 m	0.25 m ¹	Environment Agency UK
LiDAR, 2001	DTM	grid	2 m	0.25 m ¹	Environment Agency UK
Interferometric SAR, 2002	DTM	grid	5 m	1 m, 2.5m ₂	Intermap Tech. UK Ltd.
Photogrammetry, 2003	DTM	grid	10 m	1.5 m ³	Infoterra Ltd. © GeoPerspectives
Map digitizing, 1987 ⁵	DTM	grid	10 m	2.5m, 5m ⁴	Ordnance Survey GB © Crown Copyright

LR points – points of the last return of a laser pulse, DTM – digital terrain model (bare earth surface), CI – contour interval, stated vertical accuracy is a RMSE for: 1 flat land Environment Agency (2008), 2 unobstructed and obstructed flat land, Intermap (2002), 3 GeoPerspectives (2006), 4 with respect to the contour interval 5 or 10m, respectively, Ordnance Survey (2001), 5 general info on contour line currency Ordnance Survey (2001).

The quality of a DEM is often specified by the provider only as root mean squared error (RMSE) in vertical and horizontal direction. The RMSE is a statistics calculated using more accurate data set of point measurements of elevation also referred to as a 'ground truth'. RMSE is often the only information on the quality which potential users can get before they get the DEM. Additional specifications as can be seen in Table 1, usually comprise the time of data collection and spatial resolution. Ideally, this

information applies to a restricted area (a particular tile) but sometimes it is to the whole national coverage. One can call such approach of data characterization as global, numerical and aspatial. The only measure of the quality is the value of the RMSE globally describing the data distribution. It is not revealing how the error performs in the terrain model. It is an aspatial statistics.

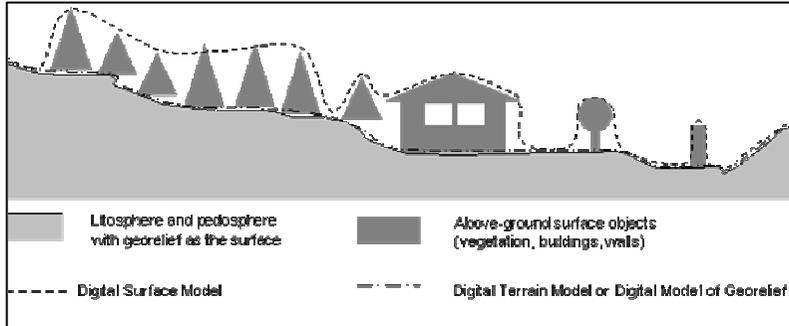


Figure 1: Schematic display of different surface levels modelled by a digital surface model (DSM), a digital terrain model (DTM). After Krcho, J. (2001), DTM represents the interface between the atmosphere and hydrosphere on one hand, and lithosphere and pedosphere on the other hand. The interface is termed as the georelief (relief, terrain). Thus, DTM is often referred to as a digital model of relief (DMR).

Objective

The objective of the research is quantifying the absolute accuracy of four different DTMs with respect to a ground truth DTM interpolated from LiDAR last return points. Subtracting the ground truth DTM from the four evaluated DTMs produced residual surfaces which can be regarded as error fields with respect to the ground truth LIDAR based DTM. The main objective of the paper is to present the methods which were used in the quantitative description of the residual surfaces and show their weaknesses and the strengths. The focus is not to discuss the quality of the DTMs or the methods of their data acquisition. If it is done so, it is only with respect to the etalon as was applied.

Datasets and methodology

All four datasets used were ready-to-use commercial DTMs covering an area of a mountainous region of the Lake District, Cumbria (Fig. 2). The site is approximately 1500 by 1500 metres in extent. The DTMs were generated from primary data collected by airborne remote sensing using light detection and ranging (LiDAR), interferometric synthetic aperture radar (InSAR), photogrammetry and by digitizing contour lines which are a secondary data. All the data sets are proprietary DTMs provided as a fine resolution grids (a point every 2 – 10 m) and they were projected in the OSGB36 coordinate system using Ordnance Datum Newlyn. The specifications can be found in Table 1 and the visualisations as shaded maps are in Figure 3. The paper is not-attempting to evaluate the DTMs, the reader is advised that any conclusion drawn should be regarded with respect to the DTM as a product of the provider and not to the original method of data acquisition.

Despite terrain being sampled at different spatial scales, the DTMs were regarded as representing a similar spatial scale. To keep the comparison as consistent as possible, all the DTMs were re-sampled to the same spatial resolution of 5 metres. The impacts of data currency and land-cover at the time of the data collection are neglected. However, the areas with the presence of trees, buildings stone walls or roads were masked in each DTM and they were not included in the analysis to avoid statistical bias.

Raw LiDAR points were also available for this research and were used to produce a ground truth DEM. This data come from the same survey mission as the data from which the LiDAR DTM was created. The vertical accuracy of the points on flat land is expected to be less than 25 centimetres (Environment Agency, 2008). The recorded values represent the last returns of a laser pulse. In cases where the laser light penetrated down to the ground, these elevations should be the most accurate samples

of the ground surface amongst the DTMs analysed. While there are limitations to this assumption, the last return points were used as a ground truth in this study.

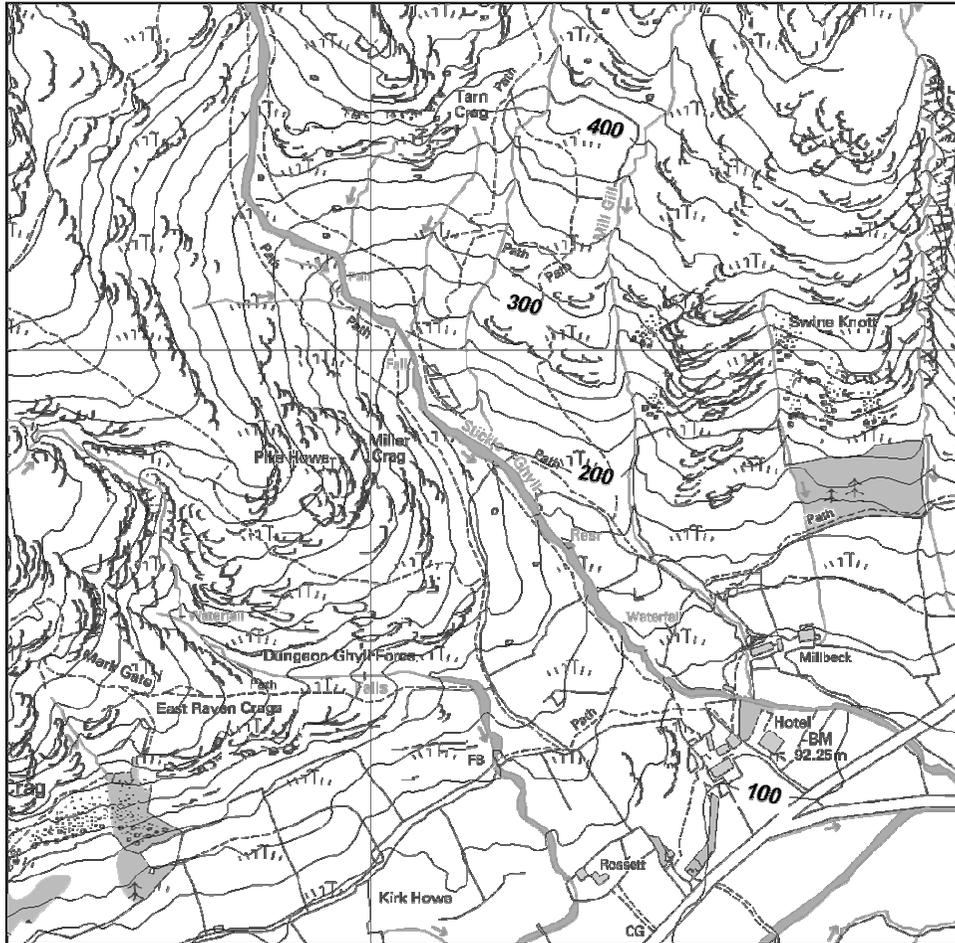


Fig. 2. Topographic map of the Great Langdale Valley at the Dungeon Ghyll, Cumbria, England.

A DTM was interpolated from the points (ca. 2 m spacing) with inverse distance weighting and the cell-size of 5 metres to match the resolution of the resampled DTMs (see Rees, W. G., 2000). The reason of the interpolation of the relatively dense field of LiDAR points was to estimate the elevation exactly on locations of the observed values which were the centres of grid cells of the resampled DTMs. The quality of the interpolated DTM was assessed with respect to total station heights on a small section of the study area and also by the cross-validation errors. Due to the different focus of the paper this issue is not discussed further. The DTM was used as the ground truth (DMT_{GT}) and was subtracted from the four DTMs (5×5 m) ($Residuals_{DMT} = DMT - DMT_{GT}$). Thus, the difference surfaces of the same spatial resolution were generated (Fig. 1). All the data were analysed and visualized in R (R Development Core Team, 2008) and QGIS with the GRASS GIS plug-in. Details on the calculation of Moran's I can be found in Lloyd (2007), Bivand et al. (2008) and manuals of *spdep* R package.

Results

At first, it is characterized by simple summary statistics (Tab. 2). They provide a fast and easy to comprehend idea of the properties of the DTMs errors. However, it is a very general and perhaps misleading picture. Let us take a look at the RMSEs. If one would decide only upon them, which is still the

only measure of error given to the user, one would not choose the LiDAR DTM. Instead, we would go for the InSAR or the photogrammetric DTM which shows lower RMSEs.

Additional statistics which describe the shape of the distribution (quantiles, skewness, kurtosis) reveal that the errors of the LiDAR DTM are more concentrated around zero than for the DTM based on contour lines. Thus, more knowledge about the shape of the distribution of errors is more useful than a single statistic. A boxplot in *Figure 4* is a graphical display of the most important distribution quantiles (quartiles, median, etc.). A histogram in the same figure visualises the shape of the data distribution. The two displays combined provide an easy-to-understand image of the datasets which is difficult to grasp from the global numerical expressions. On the other hand, measures of skewness, kurtosis, or normality quantify the shape of the distribution.

Tab. 2. Summary statistics of the masked elevation residual surfaces in metres.

DTM origin	Cells	Min	Max	1. Qrt	3. Qrt	IQR	ME	SDE	MAE	RMSE
LiDAR	81 873	-23.6	63.0	-0.2	0.4	0.6	1.1	6.6	2.8	6.7
InSAR	81 875	-27.7	38.8	-1.2	0.5	1.6	-0.1	3.7	1.8	3.7
Photogrammetry	81 875	-32.9	29.8	-1.3	0.9	2.2	-0.4	2.7	1.7	2.7
Contour based	81 875	-33.1	19.4	-2.0	1.3	3.3	-0.4	3.0	2.2	3.0

Cells – number of grid cells, Min – minimum error, Max – maximum error, 1.Qrt – lower quartile, 3.Qrt – upper quartile, IQR – inter-quartile range, ME – mean error, MAE – mean absolute error, SDE – Standard deviation of error, RMSE – root mean squared error.

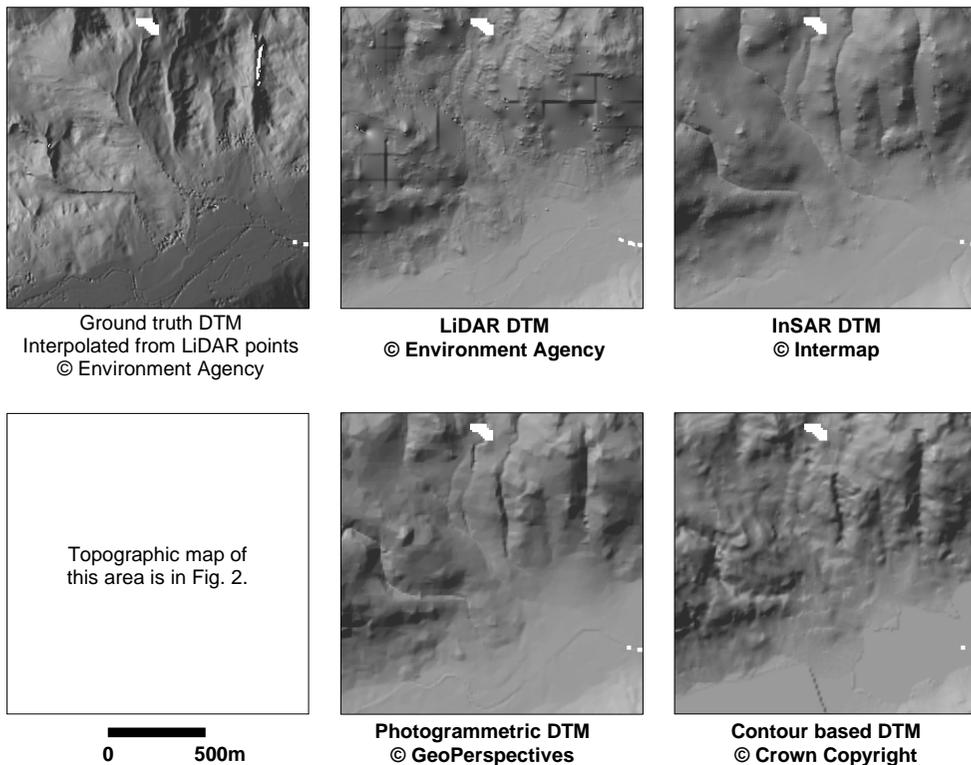


Fig. 3. Shaded relief maps of the datasets used. Great Langdale Valley at the Dungeon Ghyll, Cumbria, England.

The most complex idea about the distribution of the errors in space is provided by displaying the error field in the form of a map (Fig. 5). It brings an important, geographical aspect to the data assessment if one can assign a location to a data observation. Visualising the surface of the elevation residuals shows that the errors are spatially autocorrelated and not random. Its magnitude is higher for the LiDAR DTM residuals as there are larger patches of similar values across the region as oppose to e.g. the photogrammetric DTM residuals.

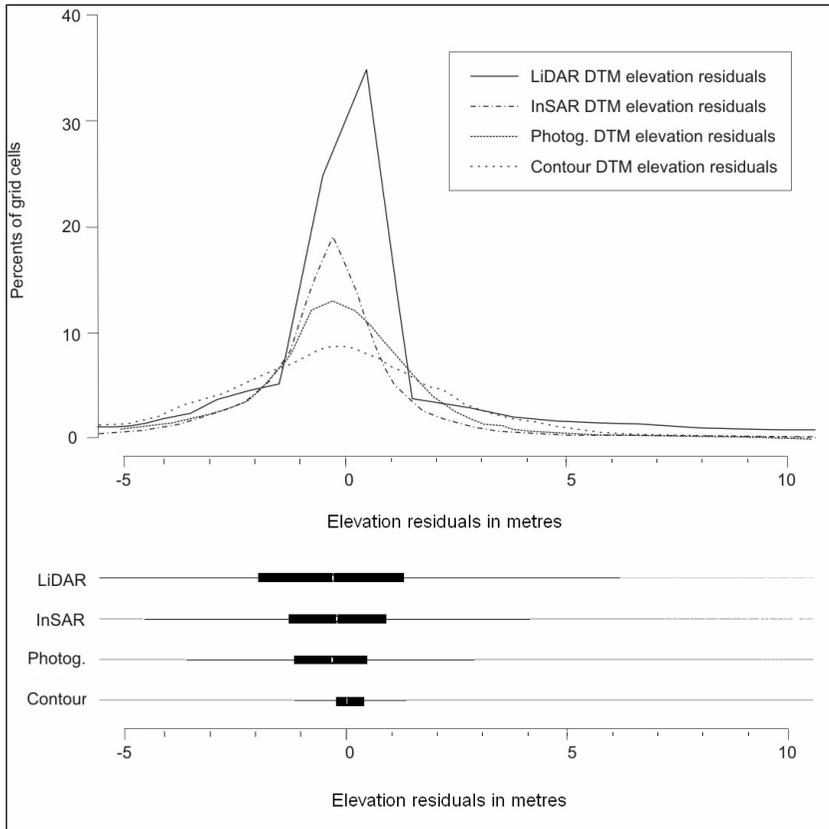


Fig. 4. Histogram of the elevation residuals of each analysed DTM. The histogram curves do not extent up to the extremes of the distributions.

Spatial autocorrelation can be described by Moran's I (Lloyd, C. D., 2007). Although it is a global statistics it parametrizes the spatial behaviour of the errors. After Moran, P. A. P. (1950), values lower than -2 or greater than $+2$ infers significant spatial autocorrelation at $p=0.95$. The values for all the analysed residual surfaces ($0.84 - 0.96$) suggest weak positive spatial autocorrelation (Fig. 5). The greater the value of Moran's I is stronger is the autocorrelation of data. Such general information enables to model the uncertainty of elevation measurements (errors) in more realistic fashion as if just the RMSE values are available (Wechsler, S. P., Kroll, C., N., 2006, Hunter, G. J., Goodchild, M. F., 1997). Knowing that a particular autocorrelation of measurement error can be expected given the data acquisition method a potential user of a DTM can more accurately quantify the uncertainty in any application of the DTM in use.

It is possible to calculate Moran's I values locally in a moving window and get an idea of its local variation. The maps of local Moran's I (I_i) in Figure 6 show the I_i values are close to zero in some areas whereas higher than 10 to 120 in others. Again, the geographical approach enables to describe the spatial pattern of errors quantitatively.

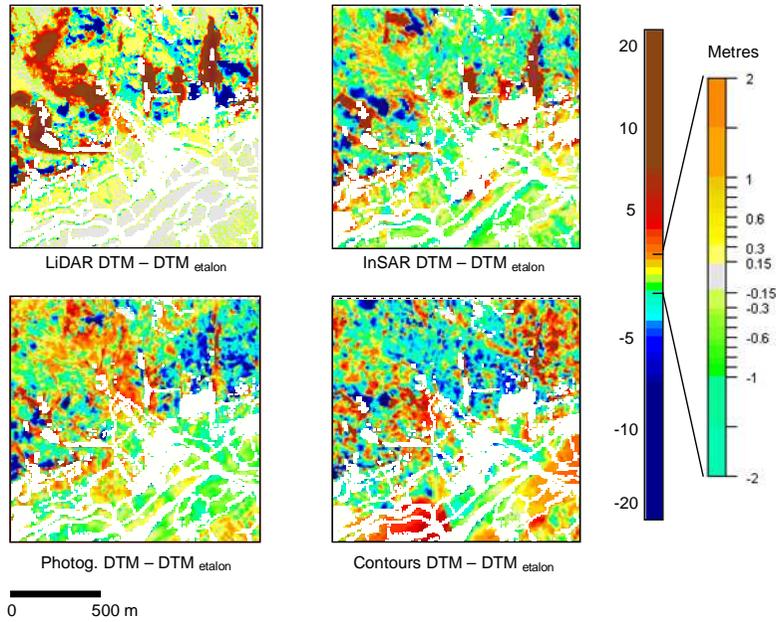


Fig. 5. Maps of elevation residuals of each analysed DTM with respect to the etalon, ‘ground truth’ DTM based on last return LiDAR point measurements. The above ground surface features are excluded from the calculation and masked with the white colour.

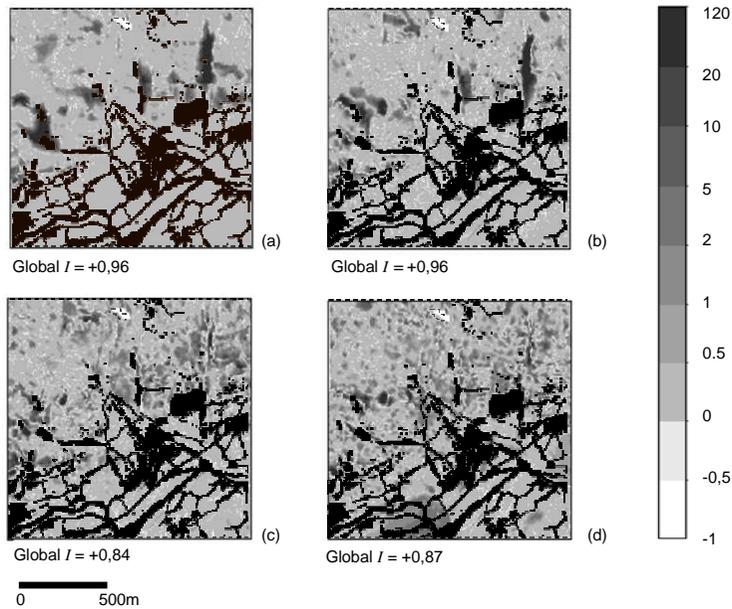


Fig. 6. Maps of local Moran's I of each masked elevation residuals surface and its global Moran's I value. Local Moran's I for (a) LiDAR DTM, (b) InSAR DTM, (c) Photogrammetric DTM, (d) Contour DTM. Values have no units, masked regions are in black.

A visual comparison of the spatial pattern of the elevation residuals and the spatial pattern of the I_i implies there is a relationship between the residuals and I_i . A scatter plot of local I_i against the absolute value of elevation residuals reveals a logarithmic relationship between the residuals and I_i (Fig. 7). Solely based on a visual inspection of the scatter, in general, the larger measurement errors are spatially more autocorrelated than the small residual values. More advanced techniques such as ordinary least squares regression or geographically weighted regression could be used to further analyse the relationship of the two variables.

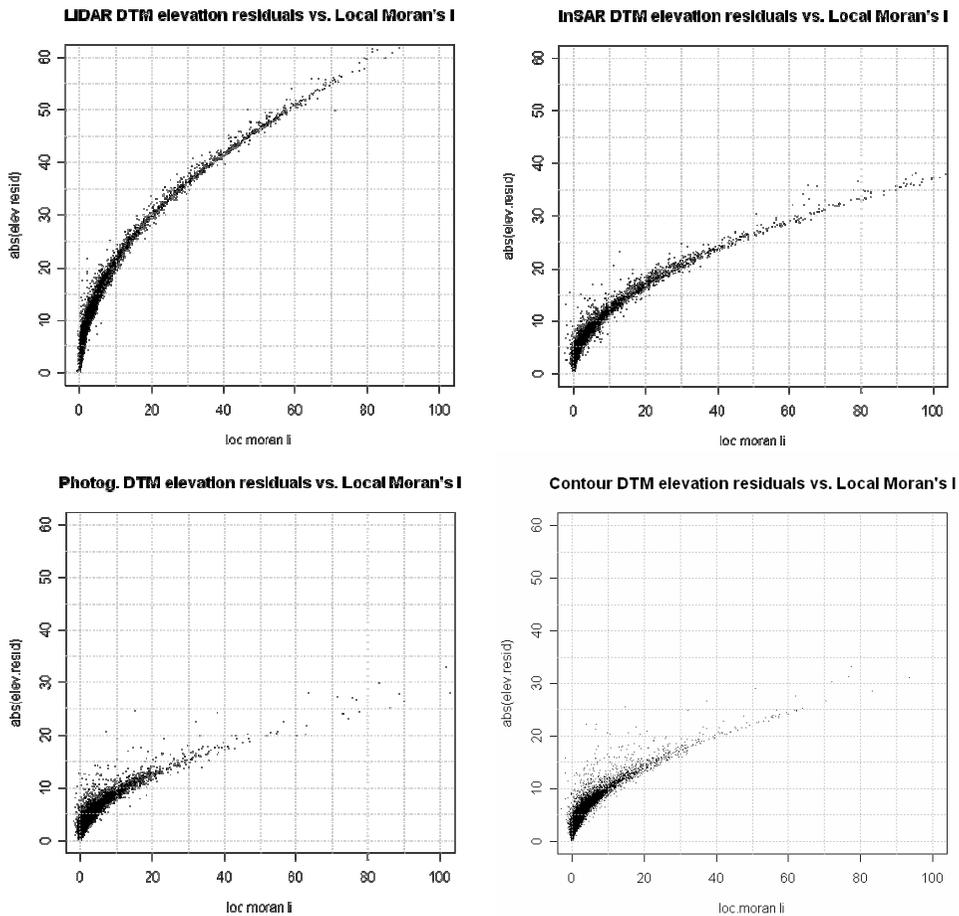


Fig. 7. Scatter plots of absolute values of elevation residuals against local Moran's I values.

Discussion and conclusions

The range of methods employed in the characterization of the properties of the errors has shown that measures which parameterized the data set globally offer a very general idea about the data distribution. Statistics as numbers give insufficient information and hide important behaviour of the data in space. Analysis of a geographical type of data requires local approaches such as locally calculated errors, statistics (Moran's I) or to conduct a geographically weighted regression. Once the local parameters are calculated, it is more appropriate to display them graphically in the form of histogram or map. The information gain of a particular approach applied in this study with respect to the spatial, graphical or numerical aspect is shown in Figure 8. The schematic graph can be argued but surely summary statistics reveal less about the data as oppose to a histogram or geographical display of

the data in the form of map. On the other hand, it is difficult to get a general picture about the data values from the map itself. Thus, one can recommend use global summary statistics and numerical methods together with their graphical visualisation. The greatest exploratory power is achieved if the methods as are facing each other along the horizontal axis in Figure 8 are used together. Also, one should be careful when using local statistics such as Moran's I, semivariogram or geographically weighted regression as the definition of 'what is local' i.e. the extent of neighbourhood is critical with this kind of data characterization. There is a plethora of textbooks dealing with global and local methods of spatial data analysis and the reader is advised to Bivand, R., S. et al. (2008) or Lloyd, C. D. (2007). Methods of ground surveying error analysis are discussed in Bitterer, L. (2006).

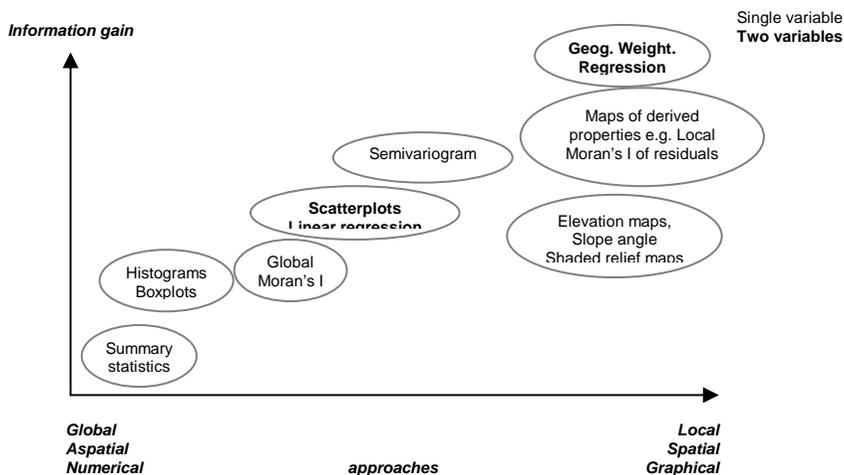


Fig 8: A general information gain of the applied approaches of data characterization.

References

- BITTERER, L., 2006: Vyrovnávací počet. Žilinská univerzita v Žiline. Stavebná fakulta. 279 p.
- BIVAND, R. S., PEBESMA, E. J., GÓMEZ-RUBIO, V., 2008: Applied Spatial Data Analysis with R. Springer New York. 378 p.
- FISHER, P. F., TATE, N. J., 2006: Causes and consequences of error in digital elevation models. Progress in Physical Geography, 30, 4, 467–489.
- HUNTER, G. J., GOODCHILD, M. F., 1997: Modeling the Uncertainty of Slope and Aspect Estimates. Geographical Analysis, 29, 1, 35–49.
- KRCHO, J., 2001: Modelling of georelief and its geometrical structure using DTM: positional and numerical accuracy. Q111, Bratislava. 336 p.
- LLOYD, C. D., 2007: Local Methods for Spatial Analysis. CRC Press New York. 244 p.
- MORAN, P. A. P., 1950: Notes on Continuous Stochastic Phenomena. Biometrika, 37, 17–33.
- RAAFLAUB, L. D., COLLINS, M. J., 2005: The effect of error in gridded digital elevation models on the estimation of topographic parameters. Environmental Modelling & Software, 21(5). 710–732.
- REES, W. G., 2000: The accuracy of Digital Elevation Models interpolated to higher resolutions. International Journal of Remote Sensing, 21, 7–20.
- WECHSLER, S. P., KROLL, C., N., 2006: Quantifying DEM Uncertainty and its Effect on Topographic Parameters. Photogrammetric Engineering and Remote Sensing, 72, 9, 1081–1090.
- R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. (softvér)

Acknowledgements

This paper was produced within a doctoral project funded by the European Social Fund at Queen's University Belfast. I would also like to thank Dr. Chris Lloyd and Dr. Jennifer McKinley for supervising the project. Last but not least, I thank the companies and institutions which provided the data for this research, namely: Environment Agency UK, Intermap Technologies UK Ltd., Infoterra Ltd., Ordnance Survey Great Britain.

Hodnotenie metód kvality digitálnych modelov terénu globálnymi a lokálnymi postupmi: príkladová štúdia s použitím údajov s vysokým rozlíšením z alternatívnych zdrojov

Michal GALLAY

***Zhrnutie:** Hlavným cieľom príspevku je poukázať na nedostatočnú výpovednú hodnotu globálnych numerických nepriestorových metód, akými sú globálne štatistické parametre (minimum, maximum, priemer, kvartily, stredná kvadratická chyba), ktoré charakterizujú údajový súbor ako celok. V prípade geografických údajov majúcich priestorový aspekt sú jednoduché opisné štatistiky nepostačujúce, respektíve treba ich použiť spolu s vizualizáciou dát aspoň formou histogramu. Optimálnym prostriedkom vyjadrujúcim kvalitu (vlastnosti) geografických údajov je však mapa a lokálne štatistické ukazovatele, ktoré analyzujú aj ich priestorové vzťahy (Moranovo I, semivariogram, geograficky vážená regresia). Príspevok sa zaoberá prístupmi hodnotenia kvality štyroch rôznych typov digitálnych modelov terénu (DTM). DTM pokrývajú identické územie z údolia Great Langdale v Lake District v Anglicku a predstavujú voľne dostupné dáta. Problém je predstavený na príklade analýzy poľa reziduí nadmorskej výšky (zvýškových povrchov), ktoré boli získané odčítaním etalónového DTM od hodnotených DTM na princípe mapovej algebry. Etalónový DTM bol vytvorený v rozlíšení 5 metrov interpoláciou hodnôt nadmorských výšok z bodov posledného odrazu laserového impulzu. Špecifikácie použitých dát možno nájsť v Table 1. Pre účely analýzy boli aj hodnotené DTM prevzorkované do rozlíšenia 5 metrov. Osobitosťou tohto výskumu je použitie hustého a relatívne presného výškového bodového poľa etalónových meraní predstavujúceho výškový model veľmi blízky priebehu reálneho terénu. Tak bolo možné analyzovať 'skutočnú' priestorovú autokoreláciu chýb merania aj jej vzťah ku metóde zberu primárných dát. Hoci elektronická tachymetria alebo diferenciálne globálne navigačné systémy umožňujú získať presnejšie údaje, dĺžka zberu pre ekvivalente husté výškové bodové pole a pre väčšie územia je niekoľkonásobne vyššia a okrem toho je potrebné uvažovať aj prístupnosť terénu. Zámerom príspevku však nie je hodnotenie kvality jednotlivých DTM, ale metódy analýzy zvýškových povrchov a ich lokálny resp. globálny aspekt. Z tohto hľadiska sa im venujú napríklad Bivand, R. S. et al. (2008), Lloyd, C. D. (2007) alebo Bitterer, L. (2006).*

Address:

Mgr. Michal Gallay
Ústav geografie, Prírodovedecká fakulta,
Univerzita P.J. Šafárika
Jesenná 5, 040 01 Košice
michal.gallay@upjs.sk

School of Geography, Archaeology and Palaeoecology,
Queen's University Belfast,
BT7 1NN, Belfast, United Kingdom