OPTIMAL INTERPOLATION OF AIRBORNE LASER SCANNING DATA FOR FINE-SCALE DEM VALIDATION PURPOSES

Michal GALLAY[1], Chris LLOYD[2], Jennifer MCKINLEY[2],

[1]Ústav geografie, Prírodovedecká fakulta, Univerzita Pavla Jozefa Šafárika, Jesenná 5, 040 01 Košice,

michal.gallay@upjs.sk

[2]School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN Belfast,

c.lloyd@qub.ac.uk, j.mckinley@qub.ac.uk

Abstract

Airborne laser scanning (ALS) is becoming a widely used method for detailed remote sensing across European and other countries. It provides dense sets of highly accurate data points which can be further used for creation of digital elevation models (DEMs). However, the high density also determines higher processing power demands. In case grid-based DEMs are concerned and interpolation is required, it is therefore useful to test the performance of particular interpolation methods using smaller samples in order to find the suitable approach and provide guidelines for other users. In this paper, inverse distance weighting (IDW) and regularized spline with tension and smoothing (RST) were tested to find the optimal DEM based on last return airborne laser scanning (ALS) data. Data from four sites differing in terrain complexity were used. Summary statistics of cross-validation errors were used to select the parameters for generating the optimal ALS DEM. The results show that IDW and RST generate very similar cross-validation statistics. Cross-validation errors increase with increasing short-range variation of elevation for IDW as well as for the RST interpolation method. Rescaled tension between 200 – 400 with smoothing provided acceptable cross-validation statistics for RST. Supplemental validation of the interpolated ALS DEMs against ground surveyed measurements provided means for absolute accuracy assessment. It suggests that both IDW and RST are equivalent for the outlined purpose. For IDW being a simpler method demanding less processing time, this was chosen for deriving the optimal ALS based DEM. The choice of interpolation method is less influential when a surface is interpolated to coarser or similar resolutions than the resolution of the input data which supports finding by other authors.

Keywords: inverse distance weighting, spline, laser scanning, lidar, interpolation

INTRODUCTION

Airborne laser scanning (ALS) can be considered as the most accurate method for mapping land surface by remote sensing. It provides rapid and dense collection of data points with submeter to subdecimeter measurement precision. Such data properties are difficult to achieve in an efficient way by other remote sensing methods such as photogrammetry, synthetic aperture radar (SAR). ALS is capable of collecting altitude of several surface levels depending on the penetration of laser beam down through the ground. The height data are further used to generate digital elevation models (DEM) and the recorded intensity of the backscattered laser beam can be used for classification of surface objects. Generally in vegetated areas, the first returns correspond to the upper landscape canopy level (e.g. vegetation tops) and the last returns to the ground (terrain surface). While the first returns are used to generate digital surface models (DSM) the last returns are used for generation of digital terrain models (DTM). In cases, where impenetrable objects such as buildings are present or the ground is exposed the first returns refer to the last returns. Detailed background on ALS can be found in Baltsavias (1999) or Pfeifer and Briese (2007). Nowadays, ALS as a technology is becoming more accessible for a wide range of users who deal with the problem of effective processing of millions of data points. This has also stimulated research on suitable approaches for generation of digital elevation model (DEM) from the ALS data (Rees 2000, Lloyd and Atkinson 2002, Mitášová et al. 2005). The studies compared performance of several interpolation methods used to derive gridded DEMs for a certain region. Comparison of various kinds of interpolation methods for DEM generation for various terrain types is presented in Carrara et al. (1997) or Svobodová and Tuček (2009) who, however, used digitized contour line data. A systematic methodology for assessment of ALS DEMs is outlined for example in Šíma (2011).

The paper presented here extends the findings of the previously mentioned authors by (i) considering four different types of terrain complexity (flat, undulating, hilly, mountainous), (ii) comparing inverse distance weighting and regularized spline with tension and smoothing as described in Mitášová and Mitáš (1993) and Mitášová and Hofierka (1993). The aim is to find an optimal DEM derived from the ALS last return data which can be further used as a reference DEM ('ground truth') for assessing vertical accuracy of DEMs derived from data acquired by different methods (contour digitizing, photogrammetry, SAR interferometry) which was done in Gallay (2010).

STUDY SITES AND DATA

It is fortunate that different types of fine-scale digital elevation datasets exist for the territory of Great Britain which is why this European region is the focus. Not only the terrain but also topographic surface data exist on a national basis. Up to date ALS, SAR interferometry, photogrammetry and digitizing of topographic maps have been used to generate nation-wide relatively high-detail DEMs. For that reason it was possible to compare data of a similar scale overlapping over different types of landscapes which is presented in Gallay (2010). As figure 1 shows four sites were selected in England where the data generated using the different acquisition methods overlap in space. In this paper ALS last return data for the sites were used for the object of the analyses. Their extent was adjusted so that all data types overlap at the site and variability of terrain is represented sufficiently. The maximal extent of the sites was limited to 2 km by 2 km with respect to the intended working scale of 1:10 000 (5 metre cell size) for the analyses. This is a manageable size if one calculates that full coverage at 5 m cell size means million data points to be analysed bearing in mind that permutations of this number will be involved.

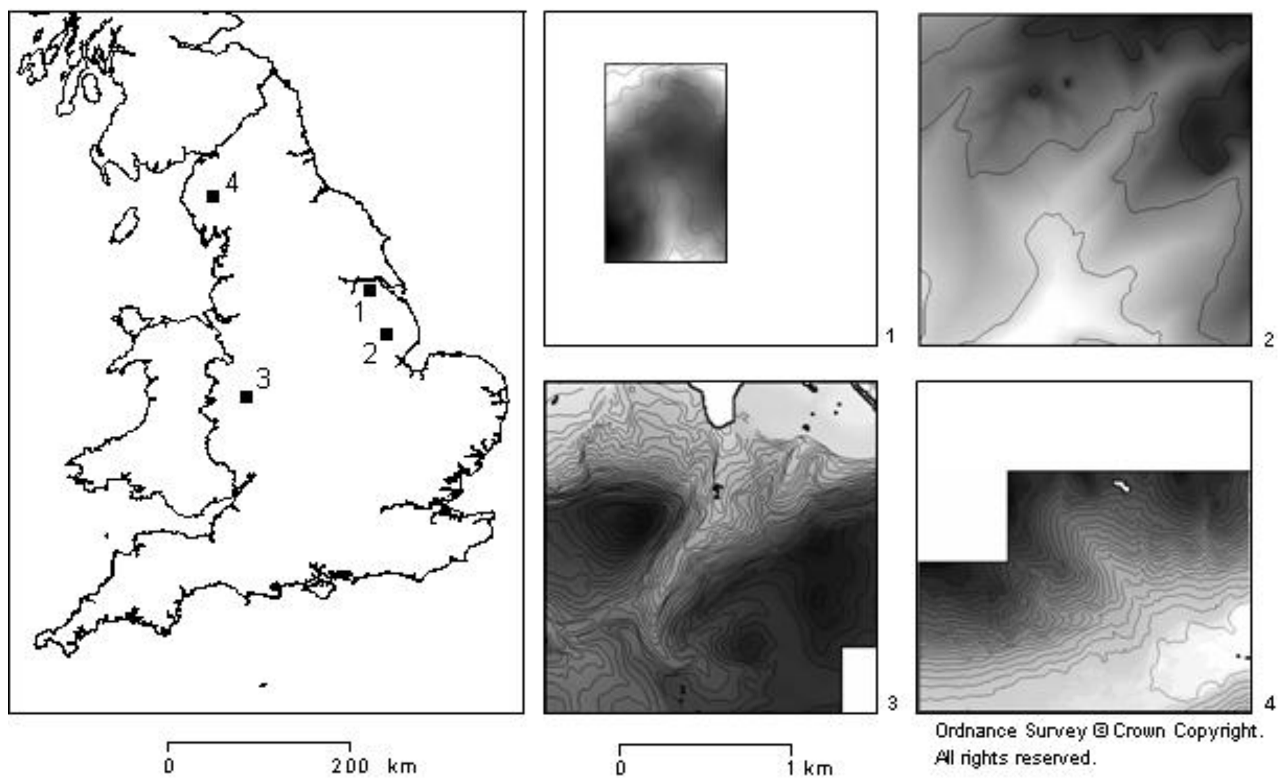


Fig. 1: The location and extent of the study sites in Great Britain. 1 - Barton upon Humber (flat terrain), 2 – Lincoln (undulating terrain), 3 – Telford (hilly terrain), 4 - Great Langdale (mountainous terrain). Contour interval 1, 2, 3 is 5 metres, for 4 it is 20 metres. Elevation increases from white to black shade.

The data were provided by the Environment Agency UK (http://www.environment-agency.gov.uk) and they originated during missions flown by the in July 2003, March 2001, August 1999, December 2000, respectively for the site 1 to site 4.

The data represent last return echoes which are considered in the paper as terrain surface samples. The Agency is responsible for official management of flood defence, visualising flood risk and flood forecast. For that reason, the Agency opted for ALS acquisition of high accuracy elevation data. The ALS missions have been flown since 1998 by plane.The measurement accuracy stated by the provider is 0.25 meters (1σ) for a flat open ground and the points have 2 meters approximate spacing. The main constraint of the study sites selection was the spatial coverage of the datasets at the time of the research. The coverage of the airborne ALS data was the main limitation in choosing the particular sites. The ALS data acquisition by the Environment Agency was limited to the regions of the Great Britain under potential risk of flooding such as coastal zones, valley floors, villages, towns and cities. For that reason, the sites were chosen on the basis of expert judgement considering the condition of differing landscape types. Providing all DEM data types have full national coverage stratified random sampling could be used for the selection.

Another dataset used for the purposes of this paper comprised elevation data collected with a total station in the Great Langdale valley (site 4). The measurements were used to assess the accuracy of the ALS points of the last return (see Gallay et al. 2011) and to assess the DTMs produced within the tuning of IDW and RST parameters. The ground survey data were acquired for two sections representing different types of terrain situated at the Middle Fell farm (inclined slope) and at the Rossett Bridge (flat alluvial plain). The height was measured every 2-5 metres on approximately a regular pattern with a denser sampling on terrain break lines. Specifications of the datasets can be found in Gallay et al. (2011) or Gallay (2010). However, it would be more appropriate to use data spread across the whole area, since the original purpose of the ground survey was to look at the performance of the ground survey methods on a finer scale rather than collect the ground truth for validating the airborne data.

METHODS

Finding the optimal interpolation parameters for generation of ALS DTM involved tuning the parameters of two interpolation methods. Inverse distance weighting (IDW) and regularized spline with tension and smoothing (RST) were chosen and tested to find the optimal DTM. IDW is a standard and fast, although not a very flexible method of local spatial prediction. The main principle is in assigning weight in the interpolation equation based on the distance to the prediction location. More details on the method can be found for example in Lloyd (2006). On the other hand, calculations with RST are more complex than with IDW and for that reason the method has higher processing time demands. RST is explained in detail by Mitášová and Mitáš (1993) and Mitášová and Hofierka (1993). The method is fully implemented in GRASS GIS software while IDW is an integral part of a wide range of analytical software ArcGIS, QGIS, Surfer. For both IDW and RST interpolation methods, there is the option to conduct the 'leave-one-out' version of K-folded cross-validation procedure, thus calculate prediction errors at the locations of the original points (see e.g. Lloyd 2006). In the special form where K equals the number of data points, the method is embedded in the Geostatistical Analyst tool of ArcGIS as well as in *v.surf.rst* in GRASS GIS which were both employed in this research. Summary statistics of these errors were used to select the parameters for generating the optimal ALS DTM. The most important criterion was to find the interpolation settings which cross-validation error statistics approximate to zero. Mean error, interquartile range, root mean squared error and mean absolute error were considered the most indicative statistics. If the requirement of cross-validation errors approaching zero is met the interpolation is as robust and reliable as possible but it does not imply that it is producing the most accurate DTM surface. For this reason the validation of the interpolated ALS DEMs was supplemented by an assessment of absolute accuracy against ground surveyed measurements on site 4.

IDW interpolation and cross-validation

IDW based interpolation and its cross-validation was conducted in Geostatistical Analyst of ArcGIS 9.0 (© ESRI). Squared distances ($p=2$) were used in the weighting function and the only varying parameter was the number of neighbouring points entering the prediction at the grid nodes. The initial setting of the number of the closest points was four and it was further incremented by 4 up to 32 points, so that eight DEMs were produced for each site. With the increment of 4 it is more likely that the number of neighbours increases evenly around the predicted location. This condition is applicable as the distribution of the raw ALS points follows approximately a regular pattern.

RST interpolation and cross-validation

Finding the optimal interpolation parameters for the RST involved tuning of tension and smoothing, while the other parameters remained constant and were used as follows:

```
v.surf.rst -t segmax=40 npmin=300 dmin=1.0 dmax=5.026359 zmult=1.0.
```

Due to the settings the total number of input points was reduced to 33 – 50% of the original, but the cross-validation and the DEM interpolation still required several hundred times longer processing time than IDW. Tension and smoothing parameters were incremented and cross-validation errors were calculated for each setting. In our case, rescaled tension ranging from 100 to1000 and smoothing values of 0.0, 0.1, 0.2, 0.5 were applied, thus 40 different DEMs were generated. Similar RST tuning procedures were carried out by Mitášová et al. (2005) and Hofierka et al. (2007).

Masking the ALS data

The raw ALS data used in the presented research represent point height measurements of the surface hit by a single laser pulse as the last, hence the term “last returns”. 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 DEMs analysed. However, this assumption is not applicable for dense vegetation cover or man-made structures such as buildings, roads, cars, etc. For that reason, the assessment is carried out in two parallel workflows for the unmasked and masked above-ground surface features. The objective was to observe the effect of the presence of the masked objects to the differences in the analyses with regard to the outputs of the statistical and geostatistical analyses. To identify the problematic objects, a DTM based on the last return ALS points was subtracted from an ALS based DSM which was also provided by the Environment Agency. The above-ground surface objects were expected to stand out in the residual surface which was visually analysed and compared against a detailed aerial orthophotograph of 25 cm spatial resolution. The objects identified as the above-ground surface features were manually traced and on-screen vectorized over the orthophotograph of each site. Further the polygons were converted to raster of 5 m cell size and the mask regions was expanded by 2 pixels (10 m) to avoid possible edge effects of the points capturing non-ground surface heights in the subsequent analyses.

Supplemental validation against ground surveyed data

This validation should be regarded as supplementary to the cross-validation and not equivalent. The heights measured by surveying with total station were subtracted from DEMs derived from the ALS data. Ideally, one would prefer TPS measurements at the locations of ALS points or within a radius corresponding to the horizontal accuracy of the ALS data (ca. 15-30 cm) depending on the height of flight). However, there were only few TPS locations within such a close distance. For that reason the following approach was applied. The DEMs produced during the tuning of IDW and RST.were generated with the spatial resolution of 5 metres at which the accuracy assessment of the airborne datasets was carried out in Gallay (2010). Elevations at TPS points were subtracted from the IDW based DEMs as well as from the RST based DEMs. Thus, elevation residuals at TPS locations were produced at 5 m spatial resolution and their statistics are provided in Table 5.3-4. It can be argued that if the elevation is sampled exactly at the locations of the grid nodes the values would differ. Interpolation of the TPS to the grid nodes would provide a solution but it would also increase number of variables in the assessment.

RESULTS

IDW interpolation

The outcomes of the IDW cross-validation displayed in figure 2 suggest that the varying number of neighbours has a relatively small effect on the statistics of each site. The differences of the central values are of the order of millimetres both for the unmasked and for the masked datasets.

In general for all sites, the mean and median error values around zero suggest unbiased predictions. The statistics fall below a centimetre level with the exception of Telford site where the predictions are systematically underestimated by less than 2 cm if mean errors are taken into account. The statistics deviate more from zero when no mask is applied. They are rather similar for Barton, Lincoln and Great Langdale datasets but higher for Telford with no masking. Although mountainous Great Langdale has the highest overall variation of elevation, it was identified in variograms that the short range variation at Telford is higher due to presence of trees (forest) and the high buildings of the power station with tall cooling towers.

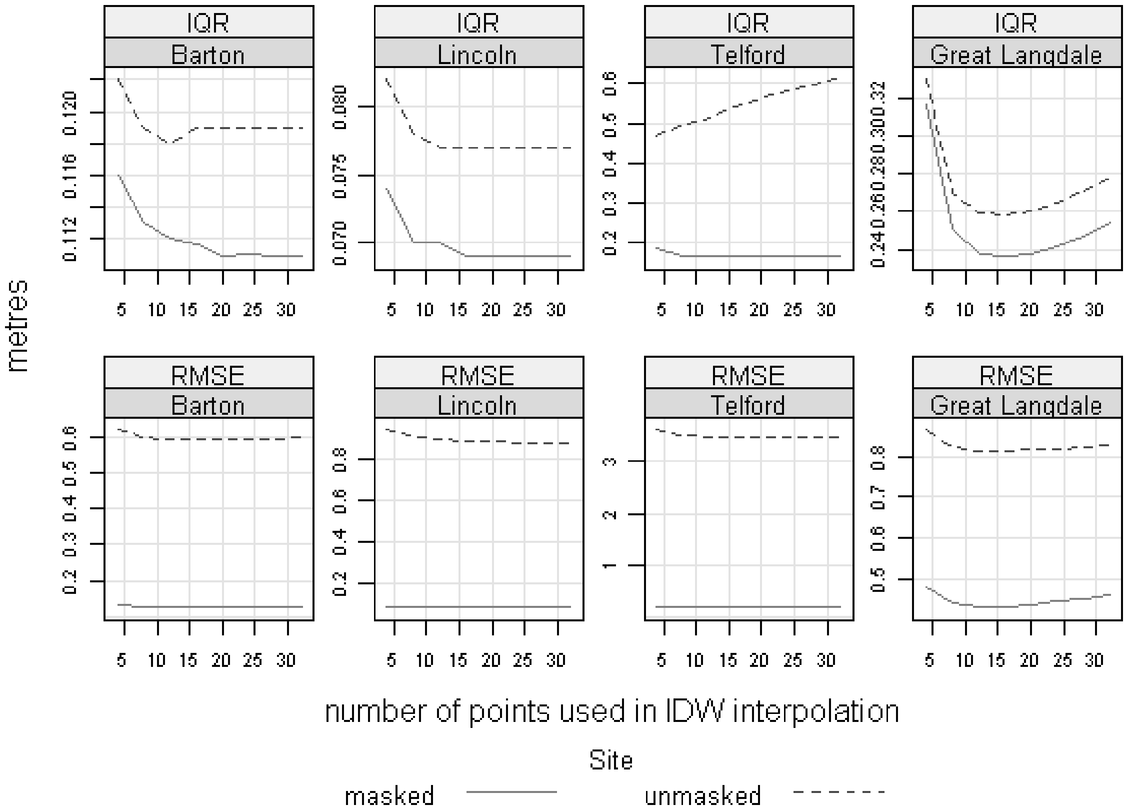


Fig. 2: IDW cross-validation statistics of interpolation of ALS data with different number of neighbours when the above-ground surface objects were masked and when they were not. Note the scale of y axis is different for each site to highlight the trends. Interquartile range (IQR) and root mean square (RMSE) of cross-validation errors.

Predicted values react more sensitively when the variation of the input values is larger at the given resolution of prediction. The above-ground surface objects are captured by the ALS on each site which biases the performance of the interpolation in the areas of bare ground surface.

One can conclude that fitting the surface by the IDW interpolation becomes less reliable in the areas of sudden increase of short range variation (higher slope angle) which are the edges of the masked areas (building walls, forest edge) or in their interiors (surface height variation within a forest, car park, settlement). Also the number of neighbours entering the interpolation with which the cross-validation statistics are favourable differs at each site. Thus, masking these objects would better describe how well the interpolation method fits the ground surface. Fig.2 also reveals that the cross-validation errors are markedly less for all sites after the masking. Overall, the RMSEs are lower for smooth terrains of Barton (lowest: 0.124 m) and Lincoln (lowest: 0.085 m) while the RMSE is higher for rougher terrain at Telford (lowest: 0.184 m) and Great Langdale (lowest: 0.431 m). There appears to be no obvious preference for optimal number of points used in the IDW prediction as the differences among the settings are very subtle in the order of millimetres. However, a local minimum of RMSE can be observed between 10 to 24 neighbours for all sites. Based on the statistics after masking the datasets, the interpolation with 12 neighbours appeared to be the optimal solution as the RMSE, MAE and ME are the closest or at least the second closest to zero for all four areas. Since it was found that this parameter does not vary for any of the sites the number of variables was not increased. The purpose of the DEMs was to be a representation of the real surface and for that reason fewer number of input points was preferred if the statistics are very similar as the variation of the points further apart would not influence the predicted values at the grid nodes.

RST interpolation and cross-validation

The RST cross-validation statistics are displayed in figure 3. The residuals seem to be unbiased for all sites and follow similar tendencies and order as the IDW cross-validation residuals. However, there are extremely large errors (above 1 metre) produced until tensions reach 200. Increasing tension values keep the residuals at a similar level with a slight sinusoidal fluctuation in the order of centimetres depending on the dataset. Local minimum is reached at tension value of 300 and local maximum tension of 600. This tendency can be seen when the above-ground surface features are unmasked and when they are excluded. Introduction of smoothing has a positive effect on reducing the residuals. There is a marked difference especially at lower tensions when no smoothing and smoothing of 0.1 is introduced. A larger smoothing value makes RMSEs at higher tensions to be more similar. Masking the above-ground surface objects also reduces the residuals. The RMSE is several times lower than for the unmasked datasets and the order is again comparable to the case of IDW cross-validation. The RMSE and the other statistics of the RST cross-validation suggest that tension of 300 and smoothing of 0.5 are optimal to be used for all sites. Even though tensions close to 1000 generate comparable RMSEs, visually such DEMs look too noisy.

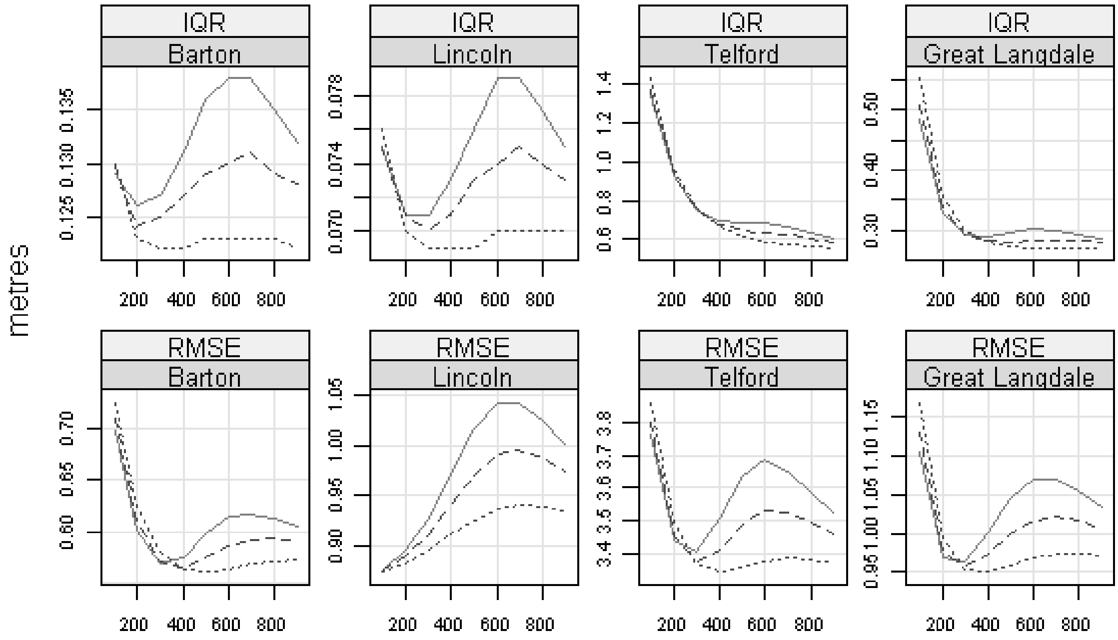
Supplemental validation against ground survey data

The lowest RMSEs for given IDW and RST interpolation settings at 5 m spatial resolution are similar for the flat meadow (about 0.30 m) and very similar on the inclined slope (IDW: 0.62 m, RST: 0.60 m). However, a greater difference was expected not only due to the use of different prediction methods but also because of the increased uncertainty of the measured values by interpolating into a coarser grid of 5 metres cell size. In fact overall, they are similar and the lowest RMSEs for a given IDW and RST settings are about 0.30 m for the flat meadow at the Rossett Bridge and 0.60 m for the slope at the Middle Fell Farm, respectively. RST RMSE as well as other statistics are the lowest when smoothing is applied and are most similar. RST with no smoothing generates larger errors which tend to decrease as tension increases. When slight smoothing is applied the residuals reach the lowest values and level out for tensions of 300 and higher.

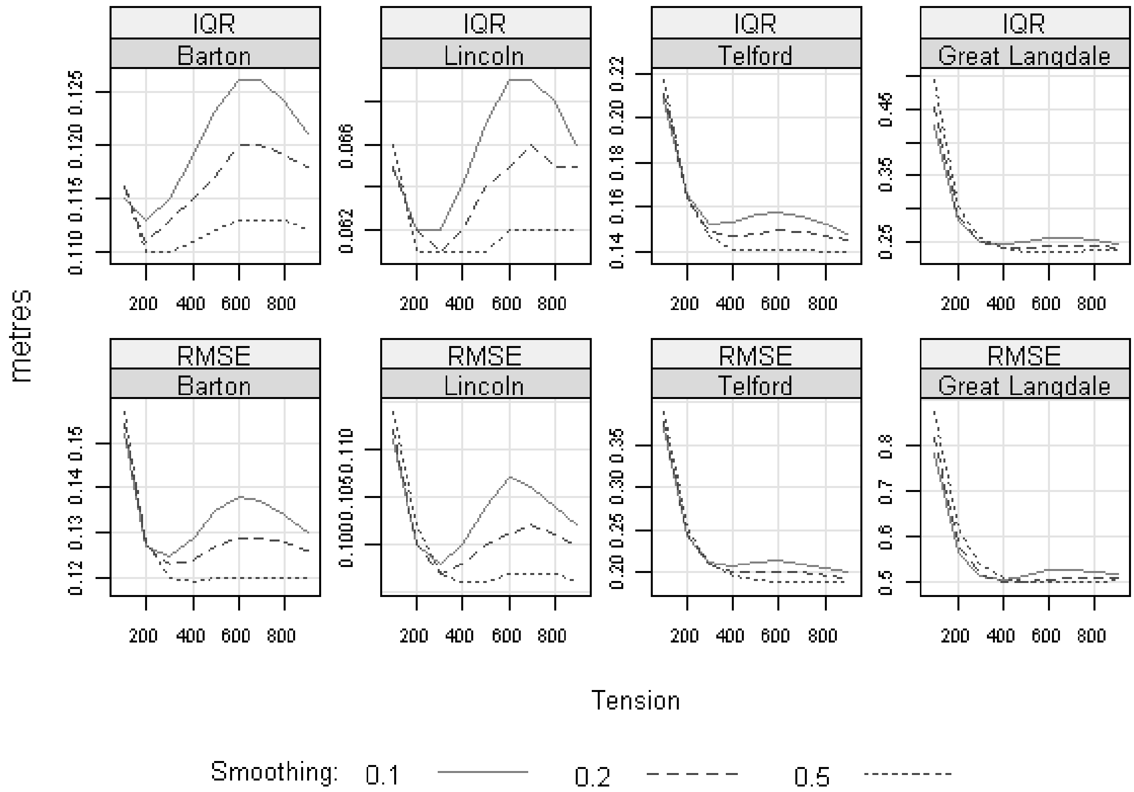
DISCUSSION AND CONCLUSIONS

In summary, the data used with IDW and RST and smoothing generate very similar cross-validation statistics when only tensions are higher than 100 are considered only. However with more detail, comparison of the lowest RMSE achieved in RST cross-validation and the lowest RMSE of the IDW cross-validation suggests better performance of RST when the datasets were unmasked. The RMSE is smaller by a few millimetres (with the exception of Barton, site 1). Conversely, the IDW performs slightly better with the masked datasets generating RMSEs lower by a few millimetres (with the exception of Great Langdale). This implies that RST generates a slightly more reliable surface compared to IDW when the short-range variation is high due to presence of the above-ground surface objects. The main drawback of RST is the processing time which is several hundred times longer for the RST than with IDW. The RST calculation used 300 neighbours to fit surface per one prediction whereas IDW worked with a ten times smaller number. The number of points entering the RST per prediction can be reduced by setting the *npmin*, *dmin* and *segmax* parameters to lower values, but tests would be required to see how RMSE reacts to such settings. Unarguably, RST can produce smooth surfaces if the settings are adjusted to do so. This kind of DEMs is desirable for many types of terrain related analysis and IDW is not capable to fine control the surface properties. For the purposes of further analysis, it was decided that IDW interpolated using 12 closest neighbours would be used to generate the

DEM to be used as the ground truth. The main reasons were that to avoid smoothing with RST and to produce exact interpolation function in case both methods produce very similar RMSE. IDW is an approachable method in proprietary as well as open source GIS systems or statistical software and it is simple to code.



No masking of the above-ground surface objects



Masking of the above-ground surface objects

Figure 3: RST cross-validation statistics of interpolation of ALS data with different tension and smoothing values *without and with applied mask* to the above-ground surface objects. Note the scale of the y axis is different for each site to highlight the trends. Values of zero smoothing are not displayed due to larger range of y axis required. Interquartile range (IQR) and root mean square (RMSE) of cross-validation errors.

Thus, the results produced can be tested more easily by other researchers. RST is recommended if both smooth and at the same time still reliable surfaces are required. The properties of a DEM can be finely tuned and controlled with RST. Testing the RST based DTM would be advised for expanding on the current outcomes of this paper. Additionally, there are further conclusions which can be drawn from this analysis with respect to the method and data used:

(i) There is a very small difference in applying IDW or RST on the ALS data in terms of the accuracy level. This just supports the findings of Rees (2000), Lloyd and Atkinson (2002), Liu et al. (2007) and Smith et al. (2005) in that the choice of interpolation method is less influential when a surface is interpolated to coarser resolutions than the resolution of the input data.

(ii) Cross-validation errors increase with increasing short-range variation of elevation for IDW as well as for the RST interpolation method. Most important is the variation within the range encompassing the neighbours which enter the prediction at a location. This relationship is suggested by the decrease of the cross-validation residuals and their statistics after the above-ground surface objects captured by the ALS were excluded. Variograms ranges changed accordingly. The user should be aware that even if the terrain is flat the presence of any buildings, walls or trees will bias the cross-validation statistics and the summary statistics should be calculated also after excluding the above-ground surface objects. It depends on the purpose of the DEM, whether the interpolation should model every feature with the highest accuracy possible or just the ground surface.

(iii) Rescaled tension (flag –t in v.surf.rst command) between 200 – 400 with smoothing provides acceptable cross-validation statistics when data of comparable spacing and accuracy are used as inputs to RST. Higher tensions cause the presence of pits and peaks around the input points even though the cross-validation errors are low. Lower tension than 200 produces overshoots, surface is very flexible and 'free to bend' unless smoothing is applied too. Similar findings were published in Mitášová et al. (2005)

(iv) Smoothing greater than zero decreases cross-validation errors of RST. It should be supported by validation against precise data for establishing the parameter value.

The supplemental comparison of ALS point datasets against the ground surveyed measurements with a total station provided insight into the absolute accuracy of the point data and DEM further used as the ground truth. It suggests that both IDW and RST are equivalent in terms of representation of the real surface for the outlined purpose. For IDW being a simpler method it was chosen for further analysis.

REFERENCES

Baltsavias, E.P. (1999) Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3), 199-214.

Carrara, A., Bitelli, G. and Carla, R. (1997) Comparison of techniques for generating digital terrain models from contour lines. International Journal of Geographical Information Science, 1997. 11(5), 451-473.

Gallay, M. (2010) Assessing alternative methods of acquiring and processing digital elevation data, in School of Geography, Archaeology and Palaeoecology (PhD thesis). Queen's University Belfast, Belfast, 389 p.

Gallay, M., Lloyd, C. D., McKinley, J. and Barry, L. (2011) Comparing the vertical accuracy of digital elevation models derived using modern ground survey and airborne laser scanning (in Slovak). Kartografické listy, 19, 61-71.

Hofierka, J., Cebecauer, T. and Šúri, M. (2007) Optimisation of Interpolation Parameters Using a Cross-validation. In Peckham, R.J., Jordan, G. (eds.) Digital terrain modelling: development and applications in a policy support environment, Springer, New York. 67-82.

Liu, X., Zhang, Z., Peterson, J. and Chandra, S. (2007) The effect of LiDAR data density on DEM Accuracy. In Oxley, L. and Kulasiri, D. (eds.) MODSIM07 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand Inc.: Christchurch, New Zealand, 1363-1369.

Lloyd, C.D. and Atkinson, P. M. (2002) Deriving DSMs from ALS data with kriging. International Journal of Remote Sensing, 23(12), 2519-2524.

Mitášová, H. and L. Mitáš, Interpolation by regularized spline with tension (1993) I. Theory and implementation. Mathematical Geology. 25, 641-655.

Mitášová, H. and J. Hofierka, Interpolation by regularized spline with tension (1993) II. Application to terrain modeling and surface geometry analysis. Mathematical Geology, 25, 657-659.

Mitášová, H., L. Mitáš, and R.S. Harmon (2005) Simultaneous spline approximation and topographic analysis for ALS elevation data in open source GIS. IEEE Geoscience and Remote Sensing Letters, 2(4), 375-379.

Pfeifer, N. and Briese, C. (2007) Geometrical aspects of airborne laser scanning and terrestrial laser scanning. In Procceedings of the ISPRS Workshop on Laser Scanning and SilviLaser 2007, Espoo, Finland, ISPRS, 311-319. URL: < http://foto.hut.fi/ls2007/final_papers/Pfeifer_2007_keynote.pdf >

Rees, W.G. (2000) The accuracy of Digital Elevation Models interpolated to higher resolutions. International Journal of Remote Sensing, 21(1), 7-20.

Smith, S. L., Holland, D. A. and Longley, P. A. (2005) Quantifying Interpolation Errors in Urban Airborne Laser Scanning Models. Geographical Analysis, 37, 200-224.

Svobodová, J., Tuček, P., Ondráčková, J. (2009): Evaluace digitálních modelů reliéfu metodami statistické analýzy. Geodetický a kartografický obzor, 55/97 (9), 227-231.

Šíma, J. (2011) Příspěvek k rozboru přesnosti digitálných modelů reliéfu odvozených z dat leteckého laserového skenování celého území ČR. Geodetický a kartografický obzor 57/99 (5), 101-106.

Vosselman, G., Kessels, P. and Gorte, B. (2005) The utilisation of airborne laser scanning for mapping. International Journal of Applied Earth Observation and Geoinformation, 6, 177-186.