

INVESTIGATING THE SPATIAL STRUCTURE OF ERROR IN DIGITAL SURFACE MODELS DERIVED FROM LASER SCANNING DATA

Smith, S.L.*¹, Holland, D.A¹, and Longley, P.A²

¹ Research & Innovation, Ordnance Survey, Romsey Road, Southampton, SO16 4GU e-mail
sjsmith@ordsvy.gov.uk, dholland@ordsvy.gov.uk

² Centre for Advance Spatial Analysis, University College London, 1-19 Torrington Place, Gower Street
London, WC1E 6BT e-mail plongley@geog.ucl.ac.uk

KEY WORDS: digital surface models, surface roughness, segmentation

ABSTRACT:

Laser scanning produces a set of irregularly spaced data points, the values of which indicate the height of the surface above datum. These data are frequently used for creating digital surface models (DSM) which represent the upper surface of the earth, or features on that surface. As the first stage in this modelling process, many commercial software packages require that the irregularly spaced raw data points are interpolated onto a regular grid before a DSM can be created. This process of interpolation introduces error into the surface model. Errors within height models have, in the past, been communicated in terms of global measures of accuracy for the model. Such quantification ignores the spatial structure of errors across the surface, hindering subsequent analysis. This paper investigates the structure, spatial patterns, and magnitude of these errors, with the aim of aiding future decision-making based on the derived DSM. This paper focuses on a number of processes which affect the level and the location of errors within a model, including comparing a variety of interpolation methods and altering the size of grid spacings for the resampling. Understanding both the quantities and the spatial structure of the error is important where models are to be used for subsequent detailed analysis, such as for telecommunications planners and flood modellers. The ability to anticipate where the largest errors are likely to occur is of paramount importance in order to assess the confidence value which may be attached to any subsequent analysis based upon the models.

RÉSUMÉ

Laser scanning produit une série de points irrégulièrement espacés, dont les valeurs indiquent la hauteur à la surface de la terre. Ces données sont fréquemment utilisées pour créer les Digital Surface Models (DSM) qui représentent la surface supérieure de la terre. Beaucoup de logiciels commerciaux exigent que les points irrégulièrement espacés soient interpolés sur une grille régulière avant qu'un DSM puisse être construit. Ce procédé d'interpolation introduit l'erreur dans le modèle de surface. Les erreurs dans les modèles de hauteur ont, dans le passé, été communiqués sous la forme de mesures globales de précision pour le modèle. Une telle quantification néglige la structure spatiale des erreurs à travers la surface, freinant l'analyse et l'identification de clusters de points. Cet article examine la structure, la répartition spatiale, et l'amplitude de ces erreurs, dans le but d'améliorer les futures prises de décision basées sur le DSM dérivé. Ces papiers se concentrent sur plusieurs traitements qui affectent le niveau et l'emplacement d'erreurs dans le modèle, dont une comparaison entre différentes méthodes d'interpolation et la modification de la taille des mailles de grille pour le reechantillonnage. La compréhension de la quantité et de la structure spatiale des erreurs sont importants pour les modèles destinés à une analyse détaillée, telle que la planification pour les télécommunications et la modélisation d'inondations. La capacité de pouvoir prévoir où les plus importantes erreurs vont probablement se produire est extrêmement importante pour évaluer le niveau de confiance qui peut être attachée à toute analyse basée sur un modèle donné.

* corresponding author

1 Introduction

Laser scanner data are obtained as an irregularly spaced set of points. Many software packages require that these points are interpolated onto a regular grid for analysis and visualisation of a height model. Despite the fact that laser scanner points are sampled at very small separation distances, the interpolation from points onto a grid can introduce a degree of uncertainty into the model. The level of this uncertainty can vary greatly with different interpolation methods and grid sizes. The aim of this paper is to identify these differences, and to suggest how they may cause the propagation of error in later stages of the modelling process. The interpolation techniques used within this investigation are: bilinear, bicubic, nearest neighbour, and biharmonic splining.

Each of the methods produces slightly different height values across the surface. Several global characteristics of the various surfaces have been noted previously in the literature. Zinger et al (2002) commented that linear interpolation will tend to overly smooth and deform building edges. However, such general characteristics reveal little about the exact spatial pattern of error within a surface model. Lloyd and Atkinson (2002) further investigated the quantification of error within interpolated surfaces. The authors focused on a comparison of Inverse Distance Weighting (IDW) and kriging interpolation, and quantified the inaccuracy in each surface. Such measures are useful general indicators of error within surface models, again however they do not reveal anything about the spatial pattern of errors across the surface. It is argued here that understanding this pattern

2 Methodology and Results

The aim of the investigation was to better understand the pattern of errors in DSMs derived from laser scanning data. It was considered that there are two principal processes which affect the spatial structure of the errors: choice of interpolation algorithm, and choice of grid size for resampling. The effect of both processes was explored in this paper. The methodology for each investigation is detailed further below.

2.1 A Comparison of Different Interpolation Algorithms

Four DSMs were created from a subset of a first return laser scanning dataset, supplied by the Environment Agency. The data were captured from an airborne sensor, at a point density of ~2m. The area used for modelling for this paper is shown in Figure 1, which shows that this sample region comprises a complex roof structure (church), some bare earth, a flat

is important for many data users who wish to use the interpolated surface in subsequent analysis. This is the rationale for the investigation presented in this paper.

Given this requirement, the errors in four DSMs were identified and quantified, and the spatial pattern explored. The effect of changing grid size on the patterns of error is also investigated, and for this thirty six DSMs were created. Particular emphasis is placed on the significance of the spatial pattern of the error for 3D feature extraction and reconstruction.

Modelling from interpolated, or gridded, data is generally considered to be less accurate than modelling from the raw data, and most recent academic research has not seriously considered this approach (such as Rottensteiner and Briese, 2002; Vosselman, 2000; Maas and Vosselman, 1999; Gruen and Wang, 1998, who have all modelled from raw data). Yet, the gridded approach remains the one favoured in many commercial and business environments, mainly as a result of the speed of processing and accessibility of software which requires regularly spaced data. Quantifying the difference in errors between the two modelling approaches remains largely unexplored. This paper forms part of ongoing research at Ordnance Survey, Great Britain, which aims to investigate and model the propagation of error in the modelling from both gridded and raw data. The work presented in this paper goes some way to understanding the spatial structure of error in the gridded approach. Further work will quantify the differences in errors between this and the raw data modelling approach, and will subsequently compare the two methods with the aim of understanding which approach is the more suitable for particular applications.

roof and a variety of vegetation. Despite being a small area (1315 points over a 80m by 50m region), the surface was considered to be representative of the typical types of structure found in the wider region.

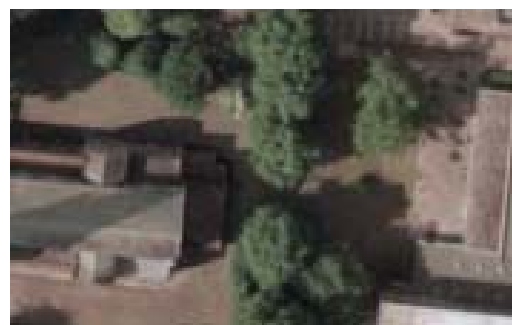


Figure 1: Orthorectified photograph of the corresponding area. Aerial photography reproduced with permission of Ordnance Survey © Cc Ordnance Survey. All rights reserved.

The raw points were first resampled onto a regular 1m grid, using four interpolation methods: bilinear, bicubic, biharmonic splining, and nearest neighbour. All interpolation was conducted in the Matlab environment, and the resultant surface forms produced are shown below in Figure 2.

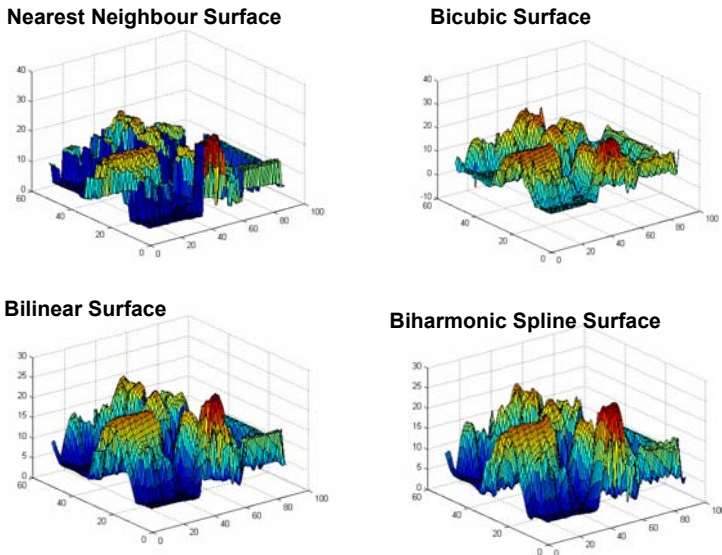


Figure 2: Showing the different surfaces created using the interpolation methods. Surfaces created from raw first return laser scanner data supplied by the Environment Agency for England and Wales.

The four interpolation methods vary in their complexity and in the time taken to compute them. Whilst this paper focuses on understanding which interpolation method produces the least error, it should also be noted that different user requirements may dictate that the accuracy of a model is not always the primary consideration. In such cases, file sizes and computation times may take priority – and a relatively simple, but less accurate, interpolator may provide the *optimal* solution. For information regarding the specifics of the four algorithms see Sandwell (1987), Watson (1992), and Smith et al (2003b).

Given the variety of methods of interpolating at unsampled sites, some significant differences in the forms of the resulting surfaces were anticipated. The forms of the surfaces in the four DSMs were compared both qualitatively and quantitatively in order to gain some understanding of differences between the interpolation errors.

2.1.1 Comparison of different interpolation algorithms

From visual inspection it was noted that the nearest neighbour surface was noticeably blocky in appearance (Figure 2), whereas the bilinear and the biharmonic splined surfaces were much smoother. The form of the bicubic surface appeared to be somewhere between the two. In order that the characteristics of the errors introduced by interpolation are fully understood, some quantification was required.

For the purposes of this investigation, the error (ϵ) at each investigated point within the surfaces was considered to be the difference between the raw data point ($Z(x)$) and the interpolated value ($Z_i(x)$) for that location (see eq.1 below).

$$\epsilon(x) = Z(x) - Z_i(x) \tag{1}$$

- where ϵ = predicted error
- x = location of point
- Z = height value
- Z_i = interpolated height value

In accordance with this definition, the error for each point on the surface was calculated. The error measurement procedure involved using different proportions of the data points to create the DSM. For this a process called jack-knifing was employed. Lloyd and Atkinson (2002) suggest that the jack-knifing technique can be used to identify which interpolation algorithm is the most accurate. The jack-knifing technique (Priyakant et al 2003; and Deutsch and Journel, 1998) involves extracting a randomly selected set of data points, and using the remaining data to create a surface. The interpolated surface is then used to predict the values of the extracted data, and the differences between the extracted and predicted values considered to be a measure of the error. The central idea underpinning this technique is that more robust algorithms will perform well despite a reduction in the number of raw data points used to form the surface. This technique was repeated using 95%, 50%, and 25% of the raw data points.

The values on the interpolated surface were then compared to each of the extracted raw data values – and the difference between the two calculated. This technique can expose general characteristics of the surface, and can provide some initial indication of differences between the surfaces. Preliminary statistical analysis was conducted on the error predictions. Statistics used included standard deviation, root mean squared error (RMSE), maximum positive error, maximum negative error, and mean. In order to check the reliability of these observations, the test was run three times (each time with a different random

selection of points). The repeat tests confirmed these observations.

Table 1: Accuracy statistics from the jack-knifing investigation

jk95					
	max	min	std dev	rmse	mean
linear	7.5799	-6.0513	2.2351	2.2242	-0.17507
cubic	7.5057	-13.086	2.6394	2.6341	-0.29322
spline	7.5525	-5.0021	1.9776	1.9666	0.30066
nearest	10.09	-8.66	2.5813	2.5683	-0.18946
jk50					
	max	min	std dev	rmse	mean
linear	12.418	-14.563	2.7011	2.7107	0.25266
cubic	14.434	-14.841	2.789	2.7949	0.21302
spline	13.043	-16.057	2.7986	2.8126	0.30066
nearest	15.99	-8.66	3.4575	3.4711	0.33574
jk25					
	max	min	std dev	rmse	mean
linear	17.112	-14.947	3.4867	3.4976	0.29758
cubic	13.269	-15.917	3.3914	3.3968	0.22166
spline	14.079	-15.662	3.1523	3.1634	0.28308
nearest	16.47	-8.66	4.2932	4.3047	0.34312

N.B In Table 1 jk95 refers to the results from the test run using 95% of the points to create the surface, and the remaining 5% to assess the accuracy of the surface. The same applied for the 50% (jk50) and the 25%(jk25) investigations.

2.1.2 Findings

The biharmonic splining algorithm was found to create the least error of all the interpolators. However, it should be noted that the relative accuracy offered by the splining method declined as the number of raw data points for the interpolation was reduced. This was due to the strong artificial oscillations in the surface reconstruction in unconstrained regions (ie. those with few data points). The results also showed that, in general, the nearest neighbour interpolator produced the highest RMSE results, meaning that it repeatedly introduced the highest amount of error across the surfaces. Despite preserving the true discontinuities across the surface, the nearest neighbour algorithm was found to introduce a significant amount of error, due to its inability to model oblique surfaces - as there are no slopes the changes between groups of values are very steep and create discontinuities which are not necessarily present in the raw data. Oblique surfaces, such as roofs in the urban environment, cannot be represented as a continuous slope in this approach.

Interestingly, there was very little difference in the magnitude of errors introduced by the bilinear and the bicubic methods – and surprisingly in many cases the bicubic method produced greater errors than the bilinear prediction. This was considered to be a result of the discontinuities in the urban environment – in such areas the bicubic interpolator will attempt to overly smooth the surface, overshoot (greater range of errors), and attempt to extrapolate at the edge of the

data – all contributing to slightly higher errors over the surface. The errors within the bilinear surface were not the result of overshoots, as this method cannot predict values greater or less than the data points within the plane at any particular location. This means that where data values are missing, the bilinear interpolator cannot ever predict these – thus creating errors in the resultant surface. Despite the fact that the errors were created by different processes, the surfaces for the bilinear and the bicubic methods introduced very similar amounts of error across both discontinuous and flat terrain.

It was noted that the higher errors which occurred in the biharmonic splined surface created using 25% of raw data points may have been caused by edge effects, where the method tried to extrapolate. To test this hypothesis the spatial pattern of the errors was analysed.

2.1.3 The Spatial Pattern of Errors

The pattern of individual errors was examined by plotting the locations of the interpolated points and assigning them a size in proportion to the error calculated for that point (Figure 3). It was observed that there was a strong spatial dependence of the highest magnitude errors in the biharmonic splined surface, and that many of the highest magnitude errors occurred at the edges of the dataset. It was considered that these edge errors were skewing the statistical analysis – perhaps incorrectly suggesting that the bilinear interpolation method was more accurate overall than either the bicubic or the biharmonic splining.

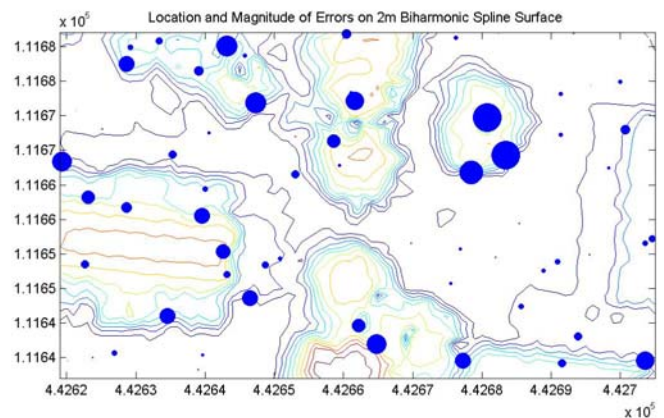


Figure 3: The Location and Magnitude of the Errors in the Biharmonic Splined Surface created on a 2m Grid.

In order to determine if these apparent edge effects were affecting the error statistics a buffer zone was created, being at least one grid cell inside the interpolated surface, and the analysis rerun within the buffer zone. The resultant error statistics for the

bilinear and the biharmonic splined surfaces are shown below:

Table 2 The Statistics Using the Buffer Zone

	Max error	Min error	Std Dev	RMSE	Mean Error
Bilinear	5.9724	-5.3674	2.1020	2.0877	0.1230
Spline	5.9883	-5.5568	2.0466	2.0345	0.1186

As can be seen from Table 2, the errors from the biharmonic splining methods using a buffer zone are lower than those produced using the bilinear method confirming that some errors within a splined surface are caused by edge effects, and that a buffer zone should therefore be used for interpolation.

Further visual inspection of the corresponding orthorectified photograph of this area showed that the patterns of highest error within the buffer zone occurred where there were discontinuities (such as the occurrence of vegetation in the scene). This correlation was confirmed by statistical analysis. It may be possible to use this pattern in subsequent feature extraction.

2.2 The Effect of Changing Grid Size

Whilst the effect of different interpolation methods on the form of the surface has been investigated in the past (eg. Zinger et al, 2002; Morgan and Habib, 2002; Lloyd and Atkinson, 2002; Smith et al, 2003a) there has been little research into the effect of changing grid size in the interpolation stage save for that of Behan (2000). Behan (2000) quantified error within models produced from different interpolation algorithms. It was found that the most accurate surfaces were created using grids which had a similar spacing to the original points. Behan's (2000) study looked at global or average error differences between two interpolation methods. This paper aims to extend Behan's (2000) work, by comparing *four* interpolation methods, and investigating the magnitudes of error created by each method at *three* different grid spacings. Identifying the most appropriate sampling distance is essential where the model is to be used for detailed analysis. The choice of grid size is especially important in applications where file size and computation times are of primary concern. Identification of how different spatial resolutions affect the representation and accuracy of derived surfaces will ultimately help users process their data efficiently.

2.2.1 Comparison of Different Grid Sizes

This investigation analysed the effects of varying grid spacing on the surface and described the characteristics of error with differing grid sizes. Thirty six Digital

Surface Models (DSMs) were created using four interpolation algorithms and 3 different grid spacings. The error patterns and statistics in each were analysed and the results presented below.

Figure 4 shows three of the surfaces created using a bilinear interpolation algorithm at different grid spacings. The surfaces are created using three grid resolutions: 1m, 2m, and 4m.

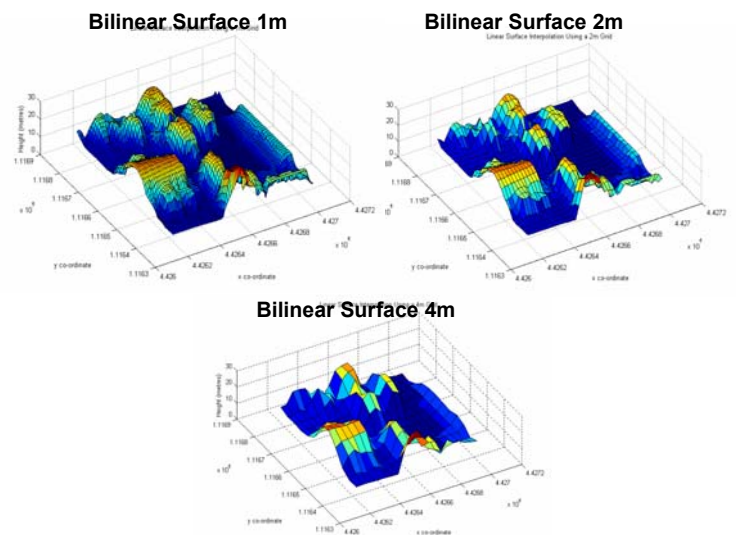


Figure 4: Showing the surfaces created from three different grid sizes

As would be expected, Figure 4 shows that there is a loss in surface detail at lower resolutions. The effect of this generalisation on the level of introduced error will clearly be more significant in areas where surface form changes are greater – such as in clusters of vegetation.

2.2.2 Quantifying the Effect of Changing Grid Size

As with the investigation into the effect of using different interpolation algorithms, the grid size investigation used the jack-knifing methodology to assess the performance of the algorithms. The statistics describing the error in the resultant surfaces are presented in Table 3 below:

Table 3 Error Statistics for Surface Created Using Different Grid Sizes

		Min Error	Max Error	Std Dev	RMSE	Mean Error
Bilinear	1m	5.4815	-6.9604	2.0519	2.0348	-0.0042
	2m	5.9724	-5.3674	2.102	2.0877	0.123
	4m	5.0303	-9.7882	2.8802	2.854	-0.1121
Bicubic	1m	7.4669	-5.7762	2.0627	2.0453	0.0265
	2m	7.3289	-5.5331	2.10601	2.1504	0.2225
	4m	4.4489	-7.1808	2.6438	2.6597	0.5327
Biharmonic Splining	1m	5.9965	-5.9152	2.137	2.1232	0.1019
	2m	5.9883	-5.5568	2.0345	2.0466	0.1186
	4m	5.309	-8.2747	2.5488	2.5299	0.0514
Nearest Neighbour	1m	9.45	-8.74	2.5592	2.5402	0.0484
	2m	9.4	-13.34	3.4035	3.3778	0.0351
	4m	12.35	-13.34	3.4344	3.4346	-0.4249

The quantification of errors revealed differences between the interpolation methods in terms of the amounts of error they introduced to the DSM. The bilinear, bicubic, and biharmonic splining interpolators produced relatively stable range and mean errors, and there appeared to be only a minimal difference between the surfaces produced at different resolutions. Despite this, it was noted that the grid spacing which generally produced the lowest error was the 1m grid. This was not the closest spacing to the original point density of the raw data, and thus does not concur with the findings of Behan (2000). However, the differences between the errors on the 1m and the 2m grids were not thought to be significant. Further tests will be needed to verify this. The nearest neighbour interpolator produced significantly higher errors at larger grid spacings than any of the other methods. In all cases interpolation onto the 4m grid resulted in higher errors, due to the loss of information in this approach. The increase in error was in the order of 50-80cm. In such cases it remains the decision of the end-user as to whether the decreases in accuracy are outweighed by faster computation and smaller file sizes.

3 Recommendations and Conclusions

This paper has shown that there is significant variation between the forms of DSMs created using different interpolation algorithms and different grid sizes. It was found that the most error was introduced by the nearest neighbour algorithm, and the least error was introduced by the biharmonic splining method. Of the methods investigated, the splining method is thus considered to be the optimal method for minimising error caused during the resampling of points in urban areas. Despite being shown to produce the least overall error, biharmonic splining and bicubic interpolation were observed to over smooth building edges and may therefore be unsuitable for some applications. Ultimately the choice of optimal algorithm for a particular application must be decided by the user – understanding spatial variations in accuracy can therefore promote better informed decision making.

This investigation has shown that changes in grid sizes have very different effects on the magnitude of error introduced by different interpolation algorithms. On the basis of the findings of this investigation, the following recommendations are made for the modelling of DSMs in the highly discontinuous mix of artificial and natural structures that characterise urban environment:

- Where accuracy is the most important factor, optimal grid spacing for any interpolation method should be as close as possible to (or slightly less than) the original point spacing. This is similar, but not identical, to the findings of Behan (2000).
- Where file size and/or computation times are the most important factor it is suggested that the bilinear interpolator be employed. In terms of computation time, the fastest method is the nearest neighbour algorithm, which works up to twice as fast as the bilinear method. However, it was found that this method produced errors of between 50 and 150cm greater than the bilinear method in urban areas. As such it is recommended that the bilinear algorithm be employed for the combination of optimal computation time and accuracy offered.

Finally, the pattern of highest magnitude error was shown to correlate with areas of greater surface roughness, whilst this pattern would be expected, a quantification of this has not previously been offered. There is some potential for this pattern to be used in subsequent image segmentation as an indicator of surface roughness. This is currently being investigated by the authors.

It is suggested here that an understanding of the spatial pattern of error in digital surfaces is as important as the provision of the surface model itself. Global accuracy statements alone are of little use for many applications, and as such there is a real requirement for an understanding of the spatial variation in error across the surface. For reliable modelling and feature extraction based on DSMs, it is of paramount importance to determine where errors are, particularly in regions of interest for city modelling such as in the vicinity of buildings and vegetation.

Ultimately the choice of optimal grid size, and interpolation method depend entirely on the application for which the surface is to be used. Studies such as the one presented in this paper are merely designed to aid with this decision making and to allow the user access to more information about both the magnitude and the spatial patterns of errors created by different methods.

4 Further Work

The potential of geostatistical interpolation is currently being investigated. Whilst the utility and accuracy of this approach have been widely publicised in the past (Armstrong, M, 1998.) the suitability of this technique for interpolating in urban environments remains to be proved. In addition, the conclusions based on the grid size investigation are being corroborated through repeat tests with alternative datasets, and tests of significance of the results.

Other future work by the authors will assess the potential for altering the grid spacing across the scene to minimise error created during interpolation. Additional work looking at a two stage interpolation process is also underway. In this approach, the breaklines within the study area are first identified, and these used to constrain the Delaunay triangle creation for the bilinear and bicubic methods. The results of this new method will then be compared to Ordinary Kriging.

References

- Armstrong, M (1998) Basic Linear Geostatistics. Springer, Germany
- Behan, A (2000) On the Matching Accuracy of Rasterised Scanning Laser Altimeter Data. *IAPRS, Vol. XXXIII, Amsterdam*
- Deutsch, C.V and Journel, A.G (1998) *GSLIB: Geostatistical Software and User's Guide*, 2nd ed. New York, Oxford University Press
- Gruen, A. and Wang, X. (1998) CC-Modeler: a topology generator for 3-D city models. *ISPRS Journal of Photogrammetry and Remote Sensing* 53 286-295
- Lloyd, C.D and Atkinson, P.M (2002) Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing* (23), 12. Pg. 2519 – 2524
- Maas, H-G and Vosselman, G (1999) Two Algorithms for Extracting Building Models from Raw Laser Altimetry Data. *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (1999) 153-163
- Morgan, M and Habib, A (2002) Interpolation of LiDAR Data and Automative Building Extraction. *ACSM-ASPRS 2002 Annual Conference Proceedings*
- Priyakant, N.K.V., Rao, L.I.M., and Singh, A.N. (2003) Surface approximation of Point Data using different Interpolation Techniques – A GIS approach. <http://gisdevelopment.net/technology/survey/techgp0009pf.htm> (accessed March, 2003)
- Rottensteiner, F and Briese (2002) A New Method for Building Extraction in Urban Areas from High-Resolution LiDAR Data. *ISPRS Commission III, Vol.34, Part 3A, Photogrammetric Computer Vision, Graz, 2002*
- Sandwell, David T., "Biharmonic Spline Interpolation of GEOS-3 and SEASAT Altimeter Data", *Geophysical Research Letters*, 2, 139-142, 1987
- Smith, S.L, Holland, D.A, and Longley, P.A (2003a) Interpreting Interpolation: The Pattern of Interpolation Errors in Digital Surface Models Derived from Laser Scanning Data. *Proceedings of GISRUK 2003*.
- Smith, S.L, Holland, D.A, and Longley, P.A (2003b) The Effect of Changing Grid Size in the Creation of Laser Scanner Digital Surface Models, *Proceedings of Geocomputation 2003 (in press)*
- Vosselman, G (2000) Slope Based Filtering of Laser Altimetry Data. *IAPRS, Vol. XXXIII, Amsterdam*
- Watson, David E., *Contouring: A Guide to the Analysis and Display of Spatial Data*, Tarrytown, NY: Pergamon (Elsevier Science, Inc.): 1992
- Zinger, S, Nikolova, M., Roux, M, and Maître, H (2002). 3D Resampling for Airborne Laser Data of Urban Areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV, 3B, Commission III

Acknowledgment

The authors wish to thank Ordnance Survey, Great Britain, for funding this research and for providing aerial photography, and the Environment Agency for kindly supplying the raw laser scanning data.