The Effect of Changing Grid Size in the Creation of Laser Scanner Digital Surface Models

Smith, S.L[1], Holland, D.A[1], and Longley, P.A[2]

[1] Research & Innovation, Ordnance Survey, Romsey Road,
Southampton, SO16 4GU
[2] Centre for Advance Spatial Analysis, University College London,
1-19 Torrington Place, Gower Street London, WC1E 6BT

Biography of Principal Author

Sarah Smith is currently in the second year of her Ph.D. researching 3D modelling from laser scanning data, with particular focus on the spatial variation of error within each stage of the modelling process. She works as a Research Scientist in Research and Innovation at Ordnance Survey, and is studying at University College London.

---

1 Introduction

Laser scanner data are obtained as an irregularly spaced set of points. For many software packages this irregular format is unsuitable and the points need to be converted to a regular grid for analysis and visualisation. Despite the fact that laser scanner points are typically sampled at very small separation distances, the subsequent transformation from points onto a grid can introduce a degree of error into the model. The level of this error can fluctuate greatly with different interpolation methods and with a varying grid size. Consequently, the choice of grid size in the creation of digital surface models is a fundamental one. Identifying the most accurate sampling distance is essential where the model is to be used for detailed analysis. The choice of grid size is also important in applications where file size and computation times are of primary concern. Identification of how different interpolation algorithms and spatial resolutions affect the representation and accuracy of derived surfaces will ultimately help users process their data efficiently.

This paper aims to contrast the effects of varying grid spacing on the representation of interpolated urban surface models derived from first return laser scanner data. It concentrates on understanding the changing 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 here. Conclusions regarding optimal grid spacings for different applications are suggested.

2 Background

Whilst the effect of different interpolation methods on the form of the surface has been investigated in the past (Zinger et al, 2002; Morgan and Habib, 2002; Lloyd and Atkinson, 2002; Smith et al, 2003) , 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. It was suggested that this was due to the fact that at larger grid spacings the exact shape of features was difficult to reconstruct as occlusions, caused by the slant of the laser path, resulted in loss of data and higher errors at the boundaries of features. Behan's (2000) study looked at global or average error differences between two interpolation methods. This paper aims to further Behan's (2000) work, by comparing four interpolation methods, and importantly also investigating the changing patterns in the spatial distribution and magnitudes of error.

3 Methodology

The data used were captured from an airborne sensor, at a point density of ~2m. The dataset was supplied by the Environment Agency for England and Wales, and is shown diagrammatically below.

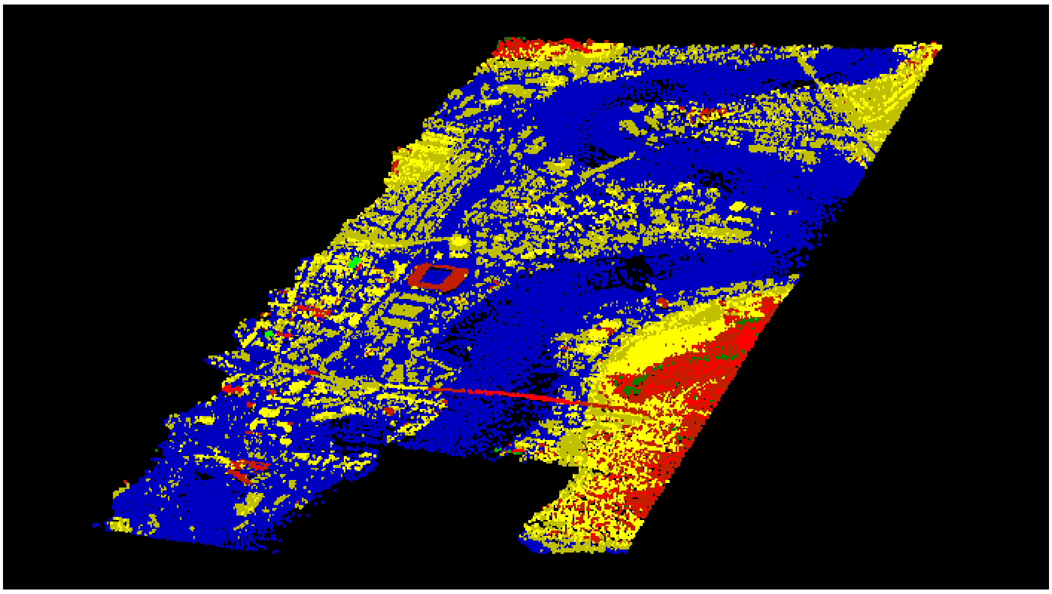


Figure 1: The area of LiDAR used for analysis. Data supplied by the Environment Agency for England and Wales. Figure produced using TerraScan software.

Thirty six DSMs were created using different interpolation algorithms and varying grid spacings. Four interpolation methods were used: bilinear, bicubic, nearest neighbour, and biharmonic splining. Each surface was created using 1m, 2m, and 4m grids, using different percentages of raw data values.

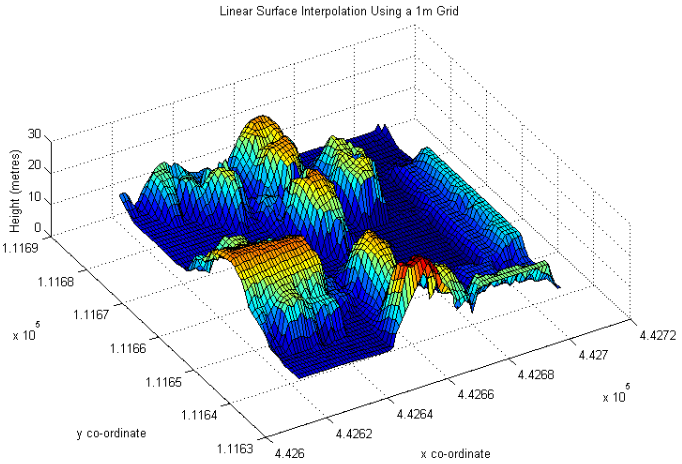
The success of the interpolated surfaces was measured using a technique called jack-knifing. Jack-knifing (Deutsch and Journel, 1998) is the process of randomly selecting subsets of the original data points, and using these points to predict the values at the remaining locations. The extracted data can then be used to assess the accuracy of the predictions. It was anticipated that this methodology would show local differences in error – which could be used in a comparison of grid sizes.

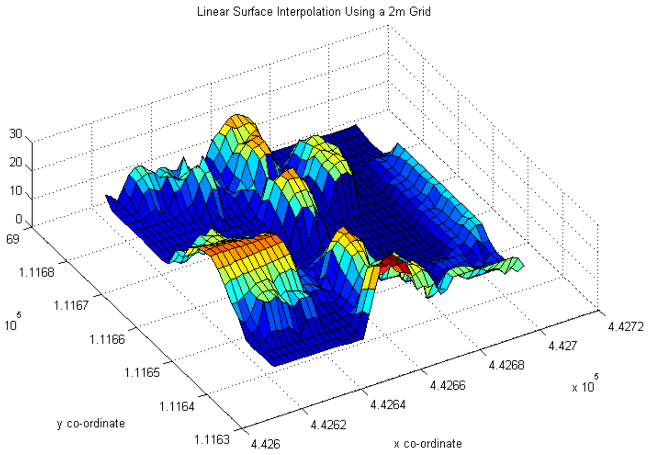
Data were analysed in the Matlab environment. As it is an interpreted language, Matlab is very quick to programme – however, because of this, it does not perform as quickly as compiled languages. This means that analysis can be problematic for large volumes of data, especially for the processing of laser scanner files which can contain many millions of points. For this reason a small subset of the original dataset was selected for this investigation. The subset chosen contained a number of buildings of varying heights, a variety of vegetation, and an area of bare earth and was considered to be a representative sample of the study region.

The surface models were created by interpolating height values at regular grid locations. A prediction of the height at the locations of the original data points was then found by interpolating from the regular surface model. By comparing the raw data values at these points with the interpolated values, a measure of error was obtained. Accuracy statistics were calculated to contrast the effects of the different grid sizes – these included: root mean squared error (RMSE), standard deviation, mean absolute difference, and the range of predicted values. In addition the pattern of the spatial variation in the magnitude of error was analysed.

4 Results

Figure 2 below shows the surfaces created using linear interpolation methods at three different grid spacings. From this diagram it can be noted that there is a significant loss of surface form with decreasing grid size. Whilst this is to be expected, the quantification of the potential increase in error is of paramount importance to the user who is assessing which grid size is optimal for a particular application. A quantification of this error is therefore critical to any informed decision making process.


Linear Surface Interpolation Using a 1m Grid


Linear Surface Interpolation Using a 2m Grid

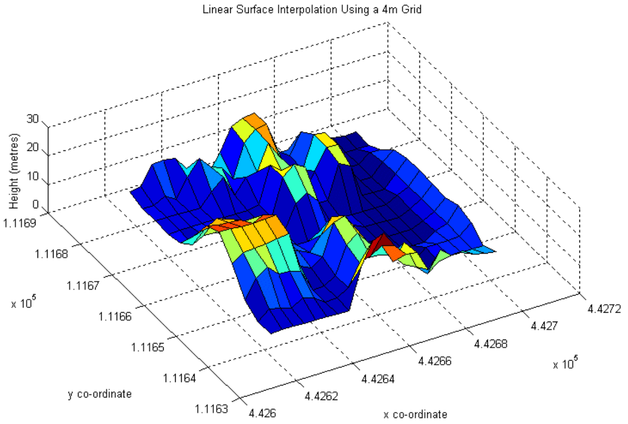

Linear Surface Interpolation Using a 4m Grid

Figure 2: Showing the surfaces created from three different grid sizes
As would be expected, figure 2 shows that there is a loss in detail in surface form 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.

The quantification of error analysis revealed some major differences between the interpolation methods in terms of the amounts of error they introduced to the DSM. It is shown in this paper that the linear, cubic, and 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 produced the lowest error was the 2m grid. This was the closest spacing to the original point density of the raw data. Interestingly, the more detailed grid, 1m spacing, appeared to introduce slightly higher errors particularly in the maximum and minimum predictions. This was not anticipated, and is being investigated further. It is shown in this paper that the nearest neighbour interpolator produced significantly higher errors at larger grid spacings than any of the other methods. Reasons for this are presented and discussed. The greater magnitude of the errors across the nearest neighbour surface created at 4m, are compared in the figure below to the errors in the linear surface.

Figure 3 shows the difference in the pattern of error between the linear and the nearest neighbour surfaces. In this diagram, the circles represent the location and magnitude of error across the interpolated surface. The colour of the circles relates to the grid size used to create the surface, while the size of the circles relates to the magnitude of the error at that location. It can be seen that the magnitude of error on the nearest neighbour surface is greater that on the linear surface. It can also be noted that the spatial distribution of errors is similar – although the highest errors on the 4m grids are slightly displaced from feature edges owing to the lower resolution.

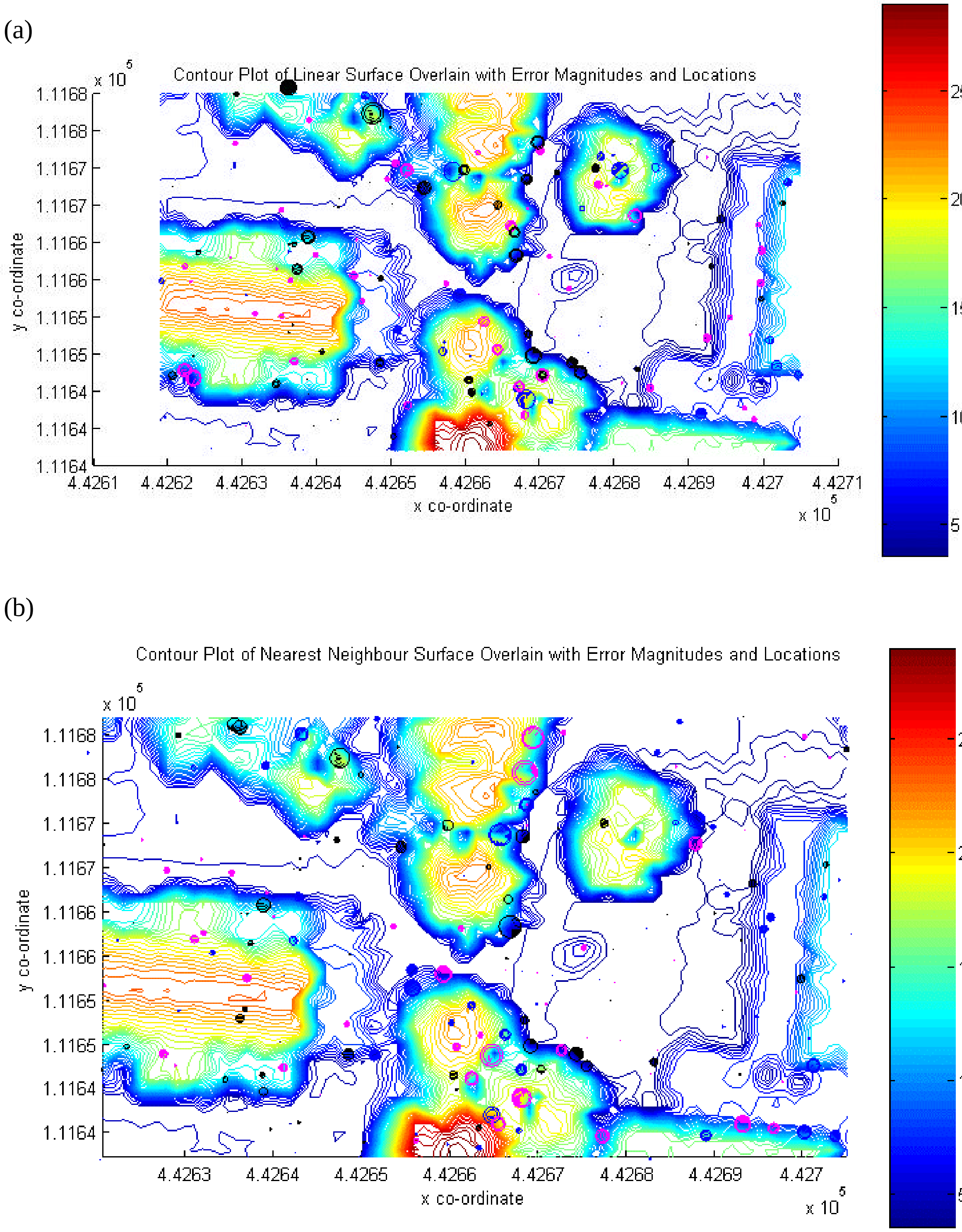


Figure 3: A comparison of error magnitude between the nearest neighbour, and the linear surfaces produced using a 4m grid. The blue circles represent the location and magnitude of errors on the 1m surface, the black circles represent those on the 2m surface, and similarly the magenta circles show the results from the 4m surface. It can be seen that the magenta circles are much larger on the nearest neighbour surface (b), than on the linear surface (a).

5 Conclusions

It is shown in this paper that changes in grid sizes have very different effects on interpolation surfaces. It is suggested that the optimal sampling spacing for minimising error is that which is as close as possible to the original point spacing. It is demonstrated that the nearest neighbour interpolator produces much greater errors than alternative methods at lower resolution grid spacings.

The choice of optimal grid spacing for DSM creation is discussed in light of the above conclusions. The importance of error introduced within the interpolation stage is discussed. The significance of relative file sizes and computation time for analysis for the varying grid sizes is also reviewed in the paper.

It is shown that the choice of grid sizes has an important baring on the magnitude and the location of higher order errors. The relevance of the changing pattern of errors is demonstrated using a number of specific examples.

References

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 edn, New York: Oxford University Press

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

Morgan, M and Habib, A (2002) Interpolation of LiDAR Data and Automative Building Extraction. *ACSM-ASPRS 2002 Annual Conference Proceedings*

Smith, S.L, Holland, D.A, and Longley, P.A (2003) Interpreting Interpolation: The Pattern of Interpolation Errors in Digital Surface Models Derived from Laser Scanning Data. *Proceedings of GISRUK 2003.* In press.

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