ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) 103–110

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Modelling vertical error in LiDAR-derived digital elevation models

Fernando J. Aguilar[a,*], Jon P. Mills[b], Jorge Delgado[b], Manuel A. Aguilar[a], J.G. Negreiros[d], José L. Pérez[c]

[a] Department of Agricultural Engineering, Almería University, 04120 Almería, Spain
[b] School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, UK
[c] Department of Cartographic Engineering, Geodesy and Photogrammetry, Jaén University, 23071 Jaén, Spain
[d] Institute of Statistics and Information Management, New University of Lisbon, 1070-312 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 25 February 2009
Received in revised form
8 September 2009
Accepted 18 September 2009
Available online 2 October 2009

Keywords:
LiDAR
Vertical accuracy assessment
Spatial modelling
Topographic mapping

ABSTRACT

A hybrid theoretical–empirical model has been developed for modelling the error in LiDAR-derived digital elevation models (DEMs) of non-open terrain. The theoretical component seeks to model the propagation of the sample data error (SDE), i.e. the error from light detection and ranging (LiDAR) data capture of ground sampled points in open terrain, towards interpolated points. The interpolation methods used for infilling gaps may produce a non-negligible error that is referred to as gridding error. In this case, interpolation is performed using an inverse distance weighting (IDW) method with the local support of the five closest neighbours, although it would be possible to utilize other interpolation methods. The empirical component refers to what is known as "information loss". This is the error purely due to modelling the continuous terrain surface from only a discrete number of points plus the error arising from the interpolation process. The SDE must be previously calculated from a suitable number of check points located in open terrain and assumes that the LiDAR point density was sufficiently high to neglect the gridding error. For model calibration, data for 29 study sites, 200 × 200 m in size, belonging to different areas around Almeria province, south-east Spain, were acquired by means of stereo photogrammetric methods. The developed methodology was validated against two different LiDAR datasets. The first dataset used was an Ordnance Survey (OS) LiDAR survey carried out over a region of Bristol in the UK. The second dataset was an area located at Gador mountain range, south of Almería province, Spain. Both terrain slope and sampling density were incorporated in the empirical component through the calibration phase, resulting in a very good agreement between predicted and observed data ($R^2 = 0.9856$; $p < 0.001$). In validation, Bristol observed vertical errors, corresponding to different LiDAR point densities, offered a reasonably good fit to the predicted errors. Even better results were achieved in the more rugged morphology of the Gador mountain range dataset. The findings presented in this article could be used as a guide for the selection of appropriate operational parameters (essentially point density in order to optimize survey cost), in projects related to LiDAR survey in non-open terrain, for instance those projects dealing with forestry applications.

1. Introduction

Accurate and high spatial resolution digital elevation models (DEMs) from airborne light detection and ranging (LiDAR) data are in increasing demand for a growing number of mapping and GIS tasks related to applications such as forest management, urban planning, bird population modelling, ice sheet mapping, flood control, road design, etc. (Lim et al., 2003). Indeed, the world of mapping and data visualization is quickly evolving into a three and four dimensional context, and the utility of LiDAR technology has been proven in a variety of scientific communities (Stoker et al., 2008). For example, the creation and display of 3D models representing the bare earth, vegetation, and surface structures have become a major focus of LiDAR application, being used to reconstruct forest height changes over long-term time series (Véga and St-Onge, 2008). LiDAR-derived DEMs are also being increasingly used for new applications relating to change detection and geopositioning (James et al., 2006; Miller et al., 2008). However, many end users of DEMs are unaware of the issues surrounding the quality of the underlying height data and their influence in derived calculations such as slope and aspect (Smith et al., 2005) or canopy height models (Clark et al., 2004).

In fact, many variables are known to contribute to the accuracy of LiDAR-derived DEMs. Among them, LiDAR sampling density

* Corresponding author. Tel.: +34 950015339; fax: +34 950015491.
E-mail addresses: faguilar@ual.es (F.J. Aguilar), j.p.mills@ncl.ac.uk (J.P. Mills), jdelgado@ujaen.es (J. Delgado), maguilar@ual.es (M.A. Aguilar), c8057@isegi.unl.pt (J.G. Negreiros), jlperez@ujaen.es (J.L. Pérez).

0924-2716/$ – see front matter © 2009 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2009.09.003

is considered to be a significant contributor to the vertical error (Hodgson et al., 2004), although terrain morphology and land cover, due to their sensitive effect on filtering performance to transform a digital surface model (DSM) to a bare-earth digital terrain model (DTM), have to be quoted as other important factors influencing the final DEM accuracy. Unlike the last two variables, however, sampling density or post-spacing represents a significant portion of overall survey costs (Raber et al., 2007) that can be selected as an operational parameter of the project.

Although there have been a number of studies dealing with LiDAR-derived DEM error, most can be classified as empirical work in which the influence of different variables on DEM error were analysed (Hopkinson et al., 2004; Hodgson and Bresnahan, 2004; Su and Bork, 2006; Goodwin et al., 2006). An exception is the work published by Kraus et al. (2006) that deals with theoretical and practical aspects of the point density, terrain curvature and resulting accuracy expectations for DTMs. In this research, a step-by-step approach is proposed to generate local and relative accuracy measures for every interpolated grid point in a DTM.

The traditional nominal accuracies of 0.15 m RMSE are very difficult to achieve except in particular cases related to flat open terrain and low altitude data collections. A few empirical studies have been conducted to date, suggesting accuracies of 0.14–1.50 m RMSE for large-scale mapping applications (Hodgson and Bresnahan, 2004; Hodgson et al., 2003; Adams and Chandler, 2002; Raber et al., 2007), depending on the platform parameters and environmental conditions. Therefore, there is a need to develop models for estimating global and absolute accuracy measures of LiDAR data under operational conditions. In this way, there is an increasing trend to propose robust and non-parametric statistical methods to estimate the accuracy of DEMs under non-open terrain, where error distribution is usually far away from the normal distribution (Aguilar and Mills, 2008; Höhle and Höhle, 2009).

Since laser scanning requires good reflectivity at the terrain surface, and blunders may occur due to multi-path in the neighbourhood of buildings, some grid posts are recorded without heights and areas of data voids can arise. Such absences in data demand an interpolation procedure be carried out to infill the voids. The amount of the missing area in relation to the whole working area is usually known as DEM completeness, and can be used as a very valuable parameter for checking DEM quality (e.g. Höhle, 2007). For example, US Federal Emergency Management Agency guidelines (FEMA, 2007) propose a maximum 5 m posting criteria for using LiDAR data to construct DEMs in the floodplain mapping process. Furthermore, FEMA's guidelines tolerate a minimum percentage of data voids or areas where the point-to-point distance is larger than a previously established threshold.

Low completeness is especially common when dealing with LiDAR forestry applications (non-open terrain), where the laser beam penetration through canopy can be limited and so the ground sampling density is consequently reduced (Fig. 1). Hence, there is a frequent need to densify the initial LiDAR point cloud (last echo) when the surveyed area presents dense vegetation (Lim et al., 2003) and new ground points have to be interpolated to infill gaps and construct accurate DTMs and canopy height models (CHMs). However, the interpolation methods used for infilling gaps may produce a non-negligible error that is henceforth referred to as gridding error (Smith et al., 2005). That is, the propagation of the sample data error (SDE) towards interpolated points. Obviously, gridding error depends on, among other variables, the interpolation method employed (Aguilar et al., 2005; Fisher and Tate, 2006).

A hybrid theoretical–empirical model has been developed for modelling the error of LiDAR-derived DEMs. It may be especially useful, in a practice sense, when applied under non-open terrain where void areas can arise due to the lack of laser beam penetration through vegetation, etc. The goal of this research is to establish a relationship between LiDAR sampling density, terrain morphology, interpolation method and LiDAR-derived DEM accuracy. The remainder of the article is organized as follows: after presenting the fundamentals of the proposed model in Section 2, the study sites, datasets and methodology employed to calibrate and validate the model are described in Section 3. In Section 4, the results corresponding to the calibration and validation phases are presented, introducing a discussion about the reliability and robustness of the two different datasets used. Section 5 is devoted to the final conclusions.

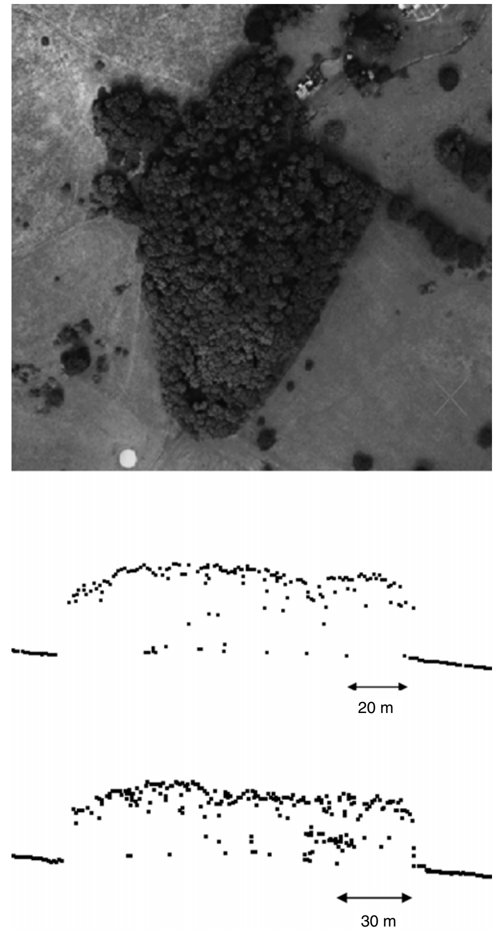

Fig. 1. Aerial image of a very dense vegetated area and two corresponding profiles of LiDAR point clouds (last pulse).

2. Model development

The full model developed in this work can be expressed as the sum of three components:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{SDE}} + \sigma^2_{\text{gridding}} + \sigma^2_{\text{filtering}}. \tag{1}$$

Where σ_{total} is the total vertical error estimated by the model, σ_{SDE} is the SDE, i.e. error from LiDAR data capture of ground sampled points in open terrain, σ_{gridding} is the aforementioned gridding error, and $\sigma_{\text{filtering}}$ is the error due to filtering non-terrain objects (buildings, vegetation, etc.) to obtain the final bare-earth DTM. All the terms are expressed as error variances.

Filtering error can be neglected in open terrain because of the absence of non-terrain objects. Similarly, gridding error can also be excluded in open terrain because of the high point density usually offered by modern LiDAR technology. Hence, SDE can be calculated from a suitable number of check points located in open terrain.

However, the question then arises as to what happens in non-open terrain? It should be noted that Eq. (1) is based on general error-propagation theory, assuming that the sources of error are linearly independent or uncorrelated and that the errors are randomly distributed. In fact, this last hypothesis is very difficult to assume when dealing with non-open terrain, due to the presence of local outliers as a result of possible filter-induced bias. Thus, $\sigma_{\text{filtering}}$ cannot be considered as randomly distributed. Instead, it is necessary to compute it by means of non-parametric methods such as that recently proposed by Aguilar and Mills (2008) for coping with this likely leptokurtic and biased error distribution. In this sense, filtering error is certainly not negligible over non-open terrain and very cumbersome to model because it depends on the algorithm used to filter the LiDAR data, the type and density of vegetation, terrain complexity and so on. Hence, it is recommended to take filtering error into account by adding the expected error to the SDE in Eq. (1) as a specific SDE_{LC} for every type of land cover distinguished over the entire working area. The expected error over non-open terrain arising from the filtering process can be effectively measured from an adequate sample of check points corresponding to every land cover type such as high grass and crops, bush lands and low trees, urban areas, etc., following the methodology proposed by Aguilar and Mills (2008).

Bearing in mind the difficulty of obtaining high precision check points over densely vegetated areas, an approximation based on experience can be recommended to estimate a value for SDE_{LC}. A number of exhaustive empirical studies are, therefore, needed to accomplish reasonable estimates for filtering error depending on the land cover characteristics and the filtering method used.

Taking into account all these considerations, Eq. (1) can then be rearranged as:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{SDE}_{\text{LC}}} + \sigma^2_{\text{gridding}}. \tag{2}$$

The remaining problem then relates to how to model the gridding error. A hybrid theoretical–empirical approach is proposed in this work. The theoretical component seeks to model the propagation of the SDE_{LC} obtained for every land cover type towards interpolated points. In this case, interpolation is performed using the inverse distance weighting (IDW) method using an exponent of two and with the local support of the M closest neighbours, although it would be possible to utilize other interpolation methods. A brief explanation of the deduced equation is presented through Eqs. (3), (3a), (4)–(8) for the case of $M = 3$ closest neighbours.

$$Z_0\left(\frac{1}{d_1^2}+\frac{1}{d_2^2}+\frac{1}{d_3^2}\right) = \frac{Z_1}{d_1^2}+\frac{Z_2}{d_2^2}+\frac{Z_3}{d_3^2}. \tag{3}$$

The variance of Z_0 (the interpolated point) and Z_i (the three sample points in this case) can be estimated using general error-propagation theory, supposing the heights Z as random second-order stationary variables and omitting covariance between sample point elevations. If the interpolation weights for every point take the values a, b and c respectively, being $a + b + c = 1$ and $a, b, c > 0$, it can be rewritten as:

$$Z_0 = aZ_1 + bZ_2 + cZ_3. \tag{3a}$$

From which it can be deduced that:

$$\sigma^2_{Z_0} = \sigma^2_{\text{surface}}\left(a^2+b^2+c^2\right). \tag{4}$$

Where σ_{surface} is an absolute and global estimation of the surface error for every land cover type, based on LiDAR-derived ground points (sampling density and accuracy) and terrain complexity, which will be developed later. Rearranging Eq. (4), the following expression can be obtained:

$$\sigma_{Z_0} = \sigma_{\text{surface}}\sqrt{\left(a^2+b^2+c^2\right)} = \sigma_{\text{surface}}K. \tag{5}$$

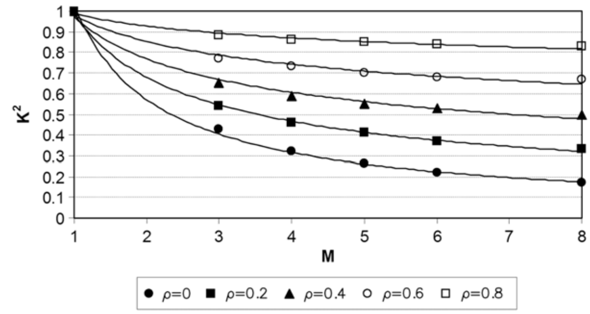

Fig. 2. Relationship between the propagation coefficient (K^2 in Eq. (8)) and the number of closest neighbours M.

A Monte Carlo numerical simulation was conducted to obtain the coefficient K in Eq. (5), randomly varying the values of the interpolation weights and the number of closest neighbours (M) over 5000 runs. The range of closest neighbours tested in Monte Carlo simulation varied between 3 and 8. Obviously, if $M = 1$ then K must be very close to 1, being taken as another point during regression analysis. The relationship between the coefficient K and M presented a very good agreement ($R^2 = 0.9972$) to the potential expression $K^2 = 1.032\ M^{-0.864}$, from which the following equation can be derived:

$$\sigma^2_{Z_0} = 1.032M^{-0.864}\sigma^2_{\text{surface}}. \tag{6}$$

This approach holds when there is no correlation between the sample point elevations. But, sample point elevations usually show a significant positive correlation depending on the terrain roughness. Hence, an extra term must be added to express the increase in error propagation due to correlation. This new approach can be expressed by the known general formulae (for the particular case of three closest neighbours):

$$\sigma^2_{Z_0} = \sum_{i=1}^{3}\sum_{j=1}^{3}\rho_{ij}\sigma_{z_i}\sigma_{z_j}\frac{\partial z_0}{\partial z_i}\frac{\partial z_0}{\partial z_j}. \tag{7}$$

Where ρ is the correlation coefficient between sample point elevations i and j. Applying this formulae for our case, the following equation can be obtained:

$$\begin{aligned}\sigma^2_{Z_0} &= (a^2+b^2+c^2+2\rho(ab+ac+bc))\sigma^2_{\text{surface}};\\ \sigma^2_{Z_0} &= K^2\sigma^2_{\text{surface}}.\end{aligned} \tag{8}$$

Notice that we are supposing that correlation coefficient is similar over all the working area (at least for every land cover type). Obviously, Eq. (5) can be easily derived when correlation is null. Following the same previously described Monte Carlo numerical simulation method, the curves showing the relationship between the number of closest neighbours and the term K^2 (Eq. (8)) can be computed for different correlation coefficients (Fig. 2). Similar potential expressions to that one presented in Eq. (6) could be derived depending on the correlation coefficient estimated for the sample point elevations (Fig. 2 shows their graphic adjustment).

Now, we have a general equation to compute the elevation error of each node in a grid DEM. It is known that if this elevation error is $\sigma^2_{Z_0}$, the propagation error of the DEM from the linear, biquadratic or bicubic interpolation is given by (Shi et al., 2005):

$$\sigma^2_H = \sigma^2_{\text{gridding}} = \frac{4}{9}\sigma^2_{Z_0}. \tag{9}$$

Where σ_H becomes the interpolated surface error of the gridded DEM, i.e. σ_{gridding}.

Finally, an expression for the error term σ_{surface} is necessary. In this work, an empirical formula is proposed based on the sum of two components:

$$\sigma^2_{\text{surface}} = \sigma^2_{\text{SDE}_{\text{LC}}} + \sigma^2_{\text{IL}}. \tag{10}$$

Where σ_{IL} would be the empirical component of the (hybrid) model which refers to what is known as "information loss" (Li, 1993; Huang, 2000). Through this work, this term can be understood as the error purely due to modelling the continuous terrain surface from only a discrete number of points, thus assuming the impossibility in finding an analytical method to compute the standard deviation for all the height differences between the real terrain surface and the DEM surface. The hypothesis here is that information loss is a global DEM error estimation related to ground point density and terrain complexity, which is propagated by means of σ_{gridding} (see Eq. (8)) into Eq. (1). Indeed, the final gridding error formulation is given by (for the case of no correlation between sample point elevations):

$$\sigma^2_{\text{gridding}} = \frac{4}{9} 1.032 M^{-0.864} (\sigma^2_{\text{SDE}_{\text{LC}}} + \sigma^2_{\text{IL}}). \tag{11}$$

Where σ_{gridding} is the gridding error across the whole surface and σ_{IL} is the information loss. Information loss was empirically modelled by means of a calibration process described subsequently in Section 3.1, which yielded the following expression:

$$\sigma_{IL} = 0.276 \text{Slope}^{0.973} N^{-0.499}. \tag{12}$$

Where Slope is the average terrain slope, a measure of terrain complexity, computed for the whole working area (dimensionless expression) and N is the LiDAR ground resolution (points/m^2). A well-known characteristic of observed elevation error for terrain mapping is the relationship with terrain slope, especially in the case of DTM generation by means of laser scanning, where planimetric error may be relatively high (around 0.5 up to 1 m) and also may be directly translated to vertical error on sloping surfaces. In this regard, better horizontal accuracies can be reached by applying a suitable fine geo-referencing by strip/block adjustment with self-calibration. In practice, due to the fact that determination of horizontal accuracy for LiDAR data is cumbersome, LiDAR accuracy is generally only stated in vertical direction. It is worth noting that the term relating to point density in Eq. (10) practically equals $1/N^{1/2}$, being morphologically quite similar to the empirical formula proposed by Karel and Kraus (2006) for computing vertical accuracy of LiDAR-derived DTMs.

Therefore, the empirical component regarding information loss embraces two variables directly implied in the final gridding error: terrain complexity and original sampling density. In this sense, a recent work published by Hu et al. (2009) has offered a theoretical explanation, based on approximation theory, for the empirical observation that relates DEM accuracy with terrain complexity and sampling density.

Referring back to Eq. (2), and substituting σ_{gridding} with the expression shown in Eq. (11), the final model formulation can now be written as:

$$\sigma^2_{\text{total}} = \sigma^2_{\text{SDE}_{LC}} + \frac{4}{9} 1.032 M^{-0.864} (\sigma^2_{\text{SDE}_{LC}} + \sigma^2_{IL}). \tag{13}$$

This last expression is only valid for the case of no correlation between sample point elevations. A simple change of the potential term should be accomplished to take into account the effect of different coefficients of correlation, as it was depicted in Fig. 2.

3. Study site and datasets

3.1. Datasets used for model calibration

Twenty-nine study sites, 200 × 200 m in size, belonging to different areas around Almeria province,south-east Spain, were acquired by means of stereo photogrammetric methods for model calibration purposes. The range of slopes and geomorphologic conditions of these 29 study sites can be considered as very diverse (from 3% up to 82% average slope) and so represented an excellent data source for calibrating the aforementioned empirical component of the model termed "information loss".

It is necessary to clarify that we only use these photogrammetric DEMs for the calibration of the empirical component of the model (information loss in Eq. (12)) and could be discussible to apply that empirical formulae, deduced from photogrammetric data, for the purpose of modelling error in LiDAR-derived DEMs. In fact, it is known that LiDAR-derived DEMs are less sensitive to terrain slope than those derived from digital photogrammetry (Hodgson and Bresnahan, 2004). In this regard, it is important to note that this empirical model only estimates the error generated as a result of modelling real terrain using a numerical abstraction (i.e. it establishes only the relationship between terrain complexity and sample point resolution), in this case a regular grid DEM. The intrinsic error arising from the method used to capture the original sample point elevations (e.g. LiDAR, stereo photogrammetry, etc.) is incorporated into the model through the term SDE$_{LC}$, which, as already mentioned, can include other specific errors (e.g. due to filtering). As a result, the empirical model can be derived independently of the method employed to acquire the raw data.

The DEM of each topographic surface was obtained automatically by stereo image matching and subsequent manual editing, introducing break lines, ridges and so on to ultimately obtain a bare-earth DEM. In this sense, it is important to highlight that the working area is an arid zone with only very scarce vegetation. The photogrammetric flight was carried out with a Zeiss RMK TOP15 metric camera with a focal length of 153.33 mm and resulted in imagery at an approximate scale of 1:5000. The average flying height was around 760 m and the base/height ratio 0.6. The negatives were digitized with a Vexcel 5000 photogrammetric scanner, with a geometric resolution of 20 μm and a radiometric resolution of 24-bits (8-bits per RGB channel). The final ground pixel size was 0.1 m. In this way, DEMs with a grid spacing of 2 m were obtained. The grid points were on UTM map projection (zone 30 North; European Datum 1950) and elevation data were stored as orthometric heights. A DEM vertical accuracy assessment was conducted using high precision check points measured by differential GPS methods, resulting in a vertical RMSE of 0.3 m (see Aguilar et al., 2007a, for more details).

The different sampling densities used to calibrate the empirical component of the model were extracted from each original grid DEM by stratified random sampling (4 × 4 sampling quadrants, i.e. 16 quadrants) that guaranteed a homogenous distribution of the sampled data over the whole working area (Burrough and McDonnell, 1998). Each sampling density tested was composed of four replicates randomly extracted and ranging from 0.25 point/m^2 (2 m average grid spacing) to 0.0008 points/m^2 (35 m average grid spacing). Residuals, or differences between the original and interpolated DEMs, were computed by the true validation method (Voltz and Webster, 1990) over a sample of 169 check points previously extracted by random sampling from the original datasets. Bilinear interpolation was used to compute the heights corresponding to the random positions of the check points from the 2 m grid spacing DEMs. Given that the reference grid spacing is relatively small, the error introduced by this interpolation approach was neglected.

The information loss, that is the response or dependent variable in regression analysis for calibrating the empirical component, was hence computed as the standard deviation of those residuals. During the calibration process, original sampling points were assumed to be free of error and hence σ_{IL} was fully coincident to the value of σ_{surface} (see Eq. (10)).

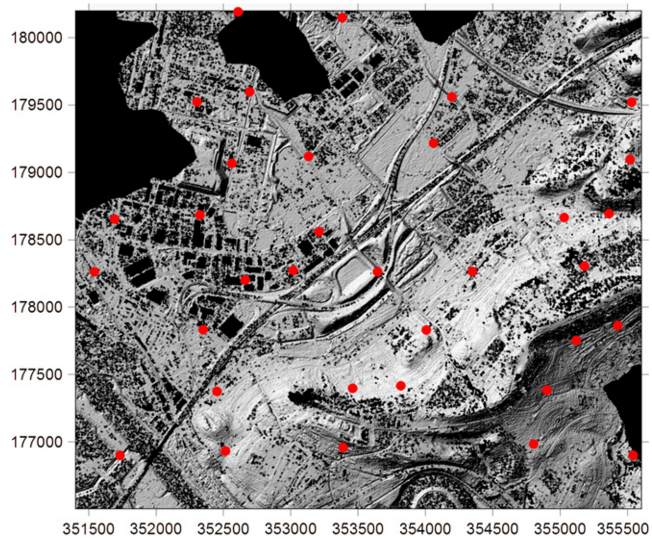

Fig. 3. Aerial view corresponding to the validation area in Bristol, UK, depicting the GPS check point distribution. Reproduced with the permission of the Ordnance Survey.

3.2. *Datasets used for model validation*

The developed methodology was validated against two different LiDAR datasets, as described below.

3.2.1. *Ordnance Survey data (Bristol, UK)*

The first dataset for model validation corresponds to an Ordnance Survey (OS) LiDAR survey carried out over a region of Bristol in the UK. The data were captured in August 2006 with a Riegl Q560 sensor resulting in an original ground spacing of between 0.5 and 1 points/m^2 over flat to hilly terrain. A set of 49 ground check points, natural features distributed throughout the study area (Fig. 3), were surveyed by OS using differential GPS methods. A number of points had to be removed because they were located in areas outside the LiDAR data coverage or situated on top of posts, fences or bollards. The remaining 34 check points, all clearly placed on the ground, enabled the assessment of the LiDAR survey vertical accuracy.

Different LiDAR point densities were acquired from the original LiDAR data by a thinning or decimation process carried out with Terrascan^{TM}software using the central point algorithm. This yielded several average grid spaced DEMs of 4.4 m, 5.3 m, 7 m, 8.4 m, 11.1 m, 13.2 m, 16.9 m and 23.5 m. The SDE computed for the Bristol area was 0.124 m and the average slope throughout the whole area was approximately 11%. Starting from the eight previously extracted LiDAR samplings, the interpolation by means of the IDW method (five closest neighbours) allowed eight dense DEMs to be obtained (0.5 points/m^2) and their accuracy assessed by means of the DGPS check points (observed error).

3.2.2. *Gador data (Almería, Spain)*

The second dataset corresponds to an area located at the Gador mountain range, south of Almeria province, Spain. LiDAR data were collected in a survey carried out in September 2007 using a Leica ALS50-II scanning system with a flying height of 1500 m, providing an approximate resolution of around 0.5 points/m^2. In this case, a high resolution image dataset (0.2 m ground pixel size) was also obtained with a digital camera, an Intergraph DMC (RGB plus IR). The processing of the second dataset consisted of working on one of the seven strips composing the LiDAR survey. Last return raw data were filtered to segment the ground surface from vegetation, buildings and any gross errors embedded in the general point cloud. The filtering algorithm developed by Axelsson (2000) and implemented in the Terrascan^{TM}software was used to segment ground points by means of a progressive TIN densification method, where the surface was allowed to fluctuate within certain values. The digital imagery was utilized to help filter non-terrain objects and bare-earth terrain.

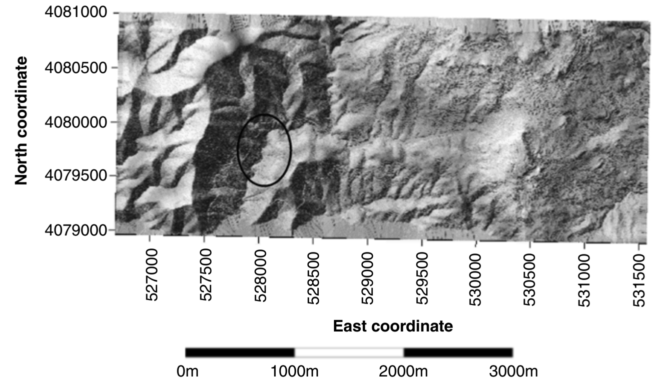

Fig. 4. Aerial view corresponding to the Gador mountain range LiDAR dataset (strip 3). The marked ellipse highlights the validation area where DGPS check points were surveyed.

After the filtering process, a field survey was carried out in November 2008 to determine the planimetric and elevation coordinates of 204 check points located across the strip by means of differential GPS method (Topcon Hiper Pro receivers) based on three base stations belonging to a previously established geodetic network measured in static mode. The coordinates of check points were referred to European Terrestrial Reference System, ETRS89 using the UTM projection. The vertical datum took the geoid as reference surface, adopting the medium level in the calm seas of Alicante (Spain) as the null orthometric height point. The average slope of the Gador working area was approximately 42.8%. Fig. 4 shows an aerial view concerning the studied LiDAR strip after filtering.

Total LiDAR error for validation purposes was computed as the differences between the heights from the 204 aforementioned check points and the ground DEM heights corresponding to the filtered LiDAR data. Further, the 3-sigma rule was applied to remove outliers which can corrupt the true statistical distribution of the errors (Daniel and Tennant, 2001), leaving 199 check points belonging to three different land cover types: bare earth, pine cover with a slightly dense vegetation understory and highly dense shrub cover. Table 1 shows statistical parameters related to error distribution at check points for every considered land cover type after outlier removal. It can be observed how systematic errors have not been removed and so the mean is not zero for vegetation land cover (pines and shrubs in this case). This phenomenon is frequently encountered when working with LiDAR data in afforested areas, where it is usual to find a bias due to dense low lying vegetation (Su and Bork, 2006; Kraus and Pfeifer, 1998; Goodwin et al., 2006). That is, LiDAR-derived DEMs tend to overestimate the reference ground elevation. Furthermore, the application of Smirnov–Kolmogorov normality test to the whole dataset revealed that error distribution was slightly non-normal for the case of non-open terrain, as has been recently highlighted by Aguilar and Mills (2008).

Finally, the filtered Gador LiDAR dataset was interpolated using IDW (five closest neighbours) to obtain a regular grid with a constant spacing of 1.4 m (equivalent to the initial nominal density for raw LiDAR data). Starting up from that original grid, a decimation process was carried out to gradually remove rows and columns obtaining nine different LiDAR point densities from 0.0015 points/m^2 up to 0.1218 points/m^2. Residuals, or LiDAR-derived DEM errors, for model validation were computed over

Table 1
Statistical characteristics of error distribution at check points for the Gador dataset regarding the different considered land cover types. Error determined by the difference between check point and LiDAR heights.

	Bare-earth cover	Pine cover with a slightly dense vegetation understory	Highly dense shrub cover	Results regarding all covers
Number of check points	50	61	88	199
Minimum error (m)	−0.36	−0.84	−1.59	−1.59
Maximum error (m)	0.28	0.48	0.20	0.48
Mean error (m)	−0.05	−0.27	−0.48	−0.31
Standard deviation (m)	0.11	0.25	0.31	0.31

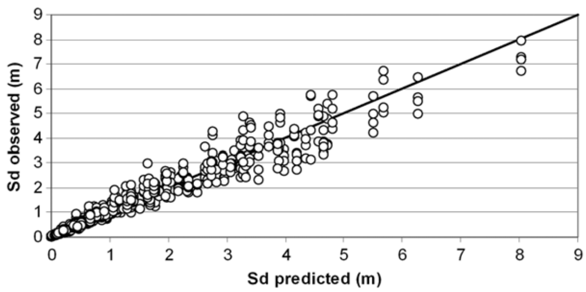

Fig. 5. Results corresponding to the calibration phase of the empirical component "information loss" (see Eq. (10)). $R^2 = 0.9856$ ($p < 0.001$).

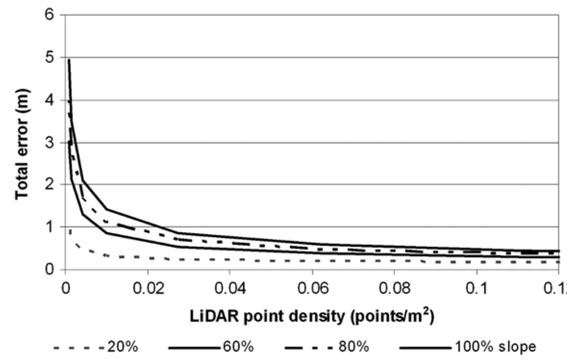

Fig. 6. Graphical representation of the proposed model performance for different average slopes. SDE takes a value of 0.15 m in this case.

the corresponding sample of 199 DGPS check points previously surveyed.

4. Results and discussion

4.1. Model calibration

The empirical component of the model, or information loss in Eq. (12), showed a good fit to the experimental data, with an R^2 regression coefficient of 0.9856 ($p < 0.001$) (Fig. 5). It must be highlighted that the standard deviation (Sd) observed and predicted in Fig. 5 refers to the term σ_{IL} in Eq. (12). DEM information loss grows almost linearly with increasing slope (rugged terrain) and presents a non-linear inverse relationship with LiDAR ground sampling density. It should be pointed out that a decrease in LiDAR ground resolution only provokes a significant increase of DEM error when considering low LiDAR post-spacing values. Furthermore, the break-point location (the LiDAR ground sampling density above which practically no gridding error is propagated to the interpolated DEM) can vary depending on the average terrain slope, which is the parameter used by the model to account for the terrain complexity. Hodgson and Bresnahan (2004) found that, when dealing with LiDAR-derived DEMs, observed elevation error in steeper slopes (e.g. 25°) was estimated to be twice as large as those on low slopes (e.g. 1.5°).

Fig. 6 shows graphically the aforementioned performance for different types of terrain slopes (from 20% up to 100%) according to Eq. (11). It is worth noting that the total error in LiDAR-derived DEMs turns out not to be very sensitive to change in LiDAR nominal post-spacing when the surveyed area presents low average slope. Nevertheless, the influence of LiDAR resolution becomes relevant when terrain complexity increases, and it is, therefore, recommended to increase LiDAR point density in order to reduce the final total error. These results coincide with those obtained by Hu et al. (2009) who, from a theoretical point of view, conclude that increasing sampling density implies a decrease in interpolation error. Thus, if very high density source data are available and the source data are highly accurate, as in the case of LiDAR data over open terrain, high accuracy is guaranteed regardless of the interpolation method used.

This finding could be deemed as a plausible explanation to the experimental results published by Raber et al. (2007), where they describe an intriguing absence of a significant pattern relating error in DEM accuracy and different LiDAR post-spacing. The likely explanation is that the terrain where they developed the research presented an average slope ranging between 2% and 5%, a much too low slope to detect significant effects on DEM accuracy due to LiDAR point density variation.

4.2. Validation of the developed model

The full model was validated using the datasets described in Section 3.2. Validation using the Bristol dataset provided the results depicted in Fig. 7. As can be seen in Fig. 7, the results offered by the developed model fit reasonably well to the observed errors, reproducing quite accurately the experimental relationship plotted against total error and LiDAR point density. It is necessary to point out that a mean correlation coefficient for sample point elevations was previously computed for the Bristol area by means of the grid correlogram, which indicates how well height values correlate across the grid DEM. This value can notably vary depending on the separation distance and terrain complexity. In the case of Bristol area, the maximum separation distance between sample points was about 35 m and the terrain can be considered as quite smooth, resulting in a mean correlation coefficient ρ of 0.80 for a separation distance lower than 35 m. This value was incorporated to Eq. (13) by means of its corresponding curve in Fig. 2.

Regarding the validation carried out over the Gador dataset, a mean correlation coefficient of around 0.40 was estimated within the separation distance range employed mainly due to the high terrain complexity of this area. It should be underlined that the estimated values performed by the proposed model fitted the observed data closely (see Fig. 8), depicting an even better agreement than that obtained using the Bristol dataset. In this regard, it is worth noting that the accuracy assessment of LiDAR-derived DEMs for the Bristol area presented an estimated average error of around 14.3%, according to the theoretical method proposed by Aguilar et al. (2007b), basically due to the low number of check points used. In the case of the Gador area, where there was a greater number of check points, the corresponding estimated error was about 5.1%, much lower than the Bristol

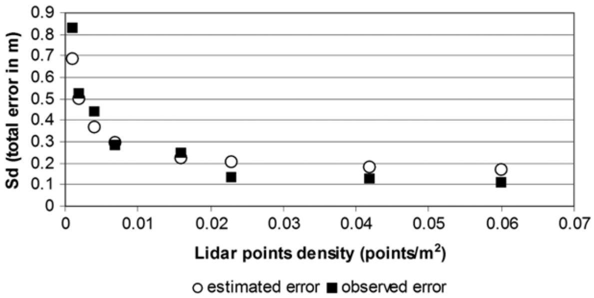

Fig. 7. Validation results corresponding to the Bristol dataset. Sd is the standard deviation of residuals at check points for every LiDAR point density tested.

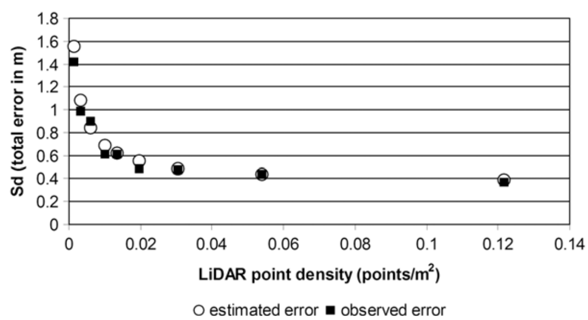

Fig. 8. Validation results corresponding to the Gador mountain range dataset. Sd is the standard deviation of residuals at check points for every LiDAR point density tested.

dataset. Therefore, the experimental data derived from Gador can be considered as more reliable than that from Bristol. Given the exhaustive number of subsets tested, and the careful methods used to filter the LiDAR raw data, these results may be deemed as highly significant to validate the proposed model.

Considering the simplification of the formulae to provide a more practitioner-suited approach, it should be highlighted that Eq. (13) has been developed to deal with the IDW method, and that the way the sample point error is propagated to the grid points is by means of this interpolation method. Hence, similar formulas could be deduced to deal with other interpolation methods. It is, therefore, non-trivial to simplify the general equation because the contribution of every term depends on the values for slope, SDE_{LC}, and point density. The normal values for M under operational conditions use between 4 and 10 closest neighbours, with the final total error being largely insensitive to variations in M when it is within this range. It is likely that the most troublesome variable to select, in the absence of an adequate set of check points, would be SDE_{LC}. In this case, and merely as a rule of thumb, using values that satisfy the US National Digital Elevation Program (Stoker et al., 2007) could be recommended (i.e. around 0.15 m vertical error (RMSE) in the case of open terrain and 0.37 m for non-open terrain). Regarding the point density estimation over non-open terrain, it is suggested, from an operative point of view, to start from the nominal point density based on the flying height and applying a reduction rate for that density based on the nature of the land cover (e.g. 25% penetration rate for densely vegetated land cover). Obviously, such a penetration rate should be based on user experience.

It is worthy to note that the proposed model could be considered as a suitable method to carry out what is called terrain generalization (Weibel, 1992), or more widely expressed as surface generalization. That is, through model application, the amount of data points required to correctly model the terrain can be reduced without practically losing accuracy (or, at least, knowing the error which is being introduced through the generalization process). Such an approach presents important advantages regarding the computer efficiency for the management of initially very dense point clouds like those obtained by laser scanning, where a previous mesh optimization process is almost always needed. Very CPU-time consuming DEM applications such as terrain visualization (Kraak, 2003), surface matching and registration for change detection (Miller et al., 2008), earthwork computation (Agüera et al., 2007), hydrological analysis (Gong and Xie, 2009), etc., can be optimized by controlled data reduction based on the generalization methods. After the corresponding generalization, and taking into account that this study has dealt with grid DEMs, data compression techniques can be utilized to address the problem of using less disk space to store the final raster format file.

5. Conclusions

This article has described the development of a hybrid theoretical–empirical model for modelling the error of LiDAR-derived DEMs under non-open terrain. Both terrain slope and sampling density were incorporated in the empirical component through the calibration phase, resulting in a very good agreement between predicted and observed data ($R^2 = 0.9856$; $p < 0.001$). Regarding the validation results, Bristol observed vertical errors, corresponding to different LiDAR point densities, offered a reasonably good fit to predicted errors. Even better results were achieved in the more rugged morphology of the Gador mountain range dataset.

Very little work has been done to determine the minimum data requirements for specific applications of DEMs, although there is an ever increasing tendency to collect larger volumes of elevation data. In the majority of cases, it is preferable to have an optimized DEM adapted to user needs rather than to have a vast amount of data, which will be more difficult to handle and process. In the particular case of LiDAR data, it must be emphasized that higher LiDAR resolutions generally require an increment in the overall survey costs (e.g. sensor with a higher pulse rate, lower altitude over-flight, narrower scan angle, and so more flight-lines to cover the same area). Despite the fact that the results reported should be regarded as preliminary, the model requiring further testing against more datasets, the findings presented in this article could be used as a guide for the selection of appropriate operational parameters, essentially point density, in order to optimize survey costs as a function of terrain complexity, in projects related to LiDAR survey in non-open terrain, for example those projects dealing with forestry applications.

Acknowledgments

The authors are grateful to the Spanish Government for financing the first author's fellowship at Newcastle University through the Spanish Research Mobility Programme "Estancias de profesores e investigadores españoles en centros de enseñanza superior e investigacion extranjeros". Thanks are due to the Ordnance Survey for facilitating access to the Bristol LiDAR dataset. Elements of this research have been made possible through the project entitled "Integration and exploratory analysis of multi-source geospatial data for monitoring and modelling the evolution and vulnerability of coastal areas" (Proyecto de Excelencia P08-RNM-03575), financed by Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía.

References

Adams, J.C., Chandler, J.H., 2002. Evaluation of lidar and medium scale photogrammetry for detecting soft-cliff coastal change. Photogrammetric Record 17 (99), 405–418.
Agüera, F., Aguilar, F.J., Aguilar, M.A., Carvajal, F., 2007. Minimising the earthwork cost in the construction of irrigation offstream reservoirs. Water Resources Management 21 (2), 375–397.
Aguilar, F.J., Agüera, F., Aguilar, M.A., Carvajal, F., 2005. Effects of terrain morphology, sampling density and interpolation methods on grid DEM accuracy. Photogrammetric Engineering & Remote Sensing 71 (7), 805–816.

Aguilar, F.J., Carvajal, F., Aguilar, M.A., Agüera, F., 2007a. Developing digital cartography in rural planning applications. Computers and Electronics in Agriculture 55 (2), 89–106.
Aguilar, F.J., Agüera, F., Aguilar, M.A., 2007b. A theoretical approach to modeling the accuracy assessment of digital elevation models. Photogrammetric Engineering & Remote Sensing 73 (12), 1367–1379.
Aguilar, F.J., Mills, J.P., 2008. Accuracy assessment of LiDAR-derived digital elevation models. Photogrammetric Record 23 (122), 148–169.
Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing 33 (Part B4), 110–117.
Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, New York.
Clark, M.L., Clark, D.B., Roberts, D.A., 2004. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. Remote Sensing of Environment 91 (1), 68–89.
Daniel, C., Tennant, K., 2001. DEM quality assessment. In: Maune, D.F. (Ed.), Digital Elevation Models Technologies and Applications: The DEM Users Manual. American Society of Photogrammetry and Remote Sensing, Bethesda, MA, pp. 395–440.
FEMA,, 2007. LiDAR Specifications for Flood Hazard Mapping, Appendix 4B: Airborne Light Detection and Ranging Systems, http://www.fema.gov/plan/prevent/fhm/lidar_4b.shtm (accessed 27.05.2009).
Fisher, P.F., Tate, N.J., 2006. Causes and consequences of error in digital elevation models. Progress in Physical Geography 30 (4), 467–489.
Gong, J., Xie, J., 2009. Extraction of drainage networks from large terrain datasets using high throughput computing. Computers & Geosciences 35 (2), 337–346.
Goodwin, N.R., Coops, N.C., Culvenor, D.S., 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. Remote Sensing of Environment 103 (2), 140–152.
Hodgson, M.E., Jensen, J.R., Schmidt, L., Schill, S., Davis, B., 2003. An evaluation of lidar and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. Remote Sensing of Environment 84 (2), 295–308.
Hodgson, M.E., Bresnahan, P., 2004. Accuracy of airborne lidar-derived elevation: Empirical assessment and error budget. Photogrammetric Engineering & Remote Sensing 70 (3), 331–339.
Hodgson, M.E., Li, X., Cheng, Y., 2004. A parameterization model for transportation feature extraction. Photogrammetric Engineering & Remote Sensing 70 (12), 1399–1404.
Höhle, J., 2007. The EUROSDR Project "automated checking and improving of digital terrain models". In: Proc. ASPRS Annual Conference, Tampa, 7-11 May. 12 p..
Höhle, J., Höhle, M., 2009. Accuracy assessment of Digital Elevation Models by means of robust statistical methods. ISPRS Journal of Photogrammetry and Remote Sensing 64 (4), 398–406.
Hopkinson, C., Chasmer, L.E., Zsigovics, G., Creed, I.F., Sitar, M., Treitz, P., Maher, R.V., 2004. Errors in LiDAR ground elevation and wetland vegetation height estimates. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W2), 108–113.
Hu, P., Liu, X., Hu, H., 2009. Accuracy assessment of digital elevation models based on approximation theory. Photogrammetric Engineering & Remote Sensing 75 (1), 49–56.
Huang, Y.D., 2000. Evaluation of information loss in digital elevation models with digital photogrammetric systems. Photogrammetric Record 16 (95), 781–791.
James, T.D., Murray, T., Barrand, N.E., 2006. Extracting photogrammetric ground control from lidar DEMs for change detection. Photogrammetric Record 21 (116), 312–328.
Karel, W., Kraus, K., 2006. Quality parameters of digital terrain models. In: Checking and improving of digital terrain models, No. 51. European Spatial Data Research EuroSDR, pp. 125–139.
Kraak, M.J., 2003. Geovisualization illustrated. ISPRS Journal of Photogrammetry and Remote Sensing 57 (5–6), 390–399.
Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53 (4), 193–203.
Kraus, K., Karel, W., Briese, C., Mandlburger, G., 2006. Local accuracy measures for digital terrain models. Photogrammetric Record 21 (116), 342–354.
Li, Z., 1993. Mathematical models of the accuracy of digital terrain model surfaces linearly constructed from square gridded data. Photogrammetric Record 14 (82), 661–674.
Lim, K., Treitz, P., Wulder, M., St-Onge, B., Flood, M., 2003. LiDAR remote sensing of forest structure. Progress in Physical Geography 27 (1), 88–106.
Miller, P., Mills, J.P., Edwards, S., Bryan, P., Marsh, S., Mitchell, H., Hobbs, P., 2008. A robust surface matching technique for coastal geohazard assessment and management. ISPRS Journal of Photogrammetry and Remote Sensing 63 (5), 529–542.
Raber, G.T., Jensen, J.R., Hodgson, M.E., Tullis, J.A., Davis, B.A., Berglund, J., 2007. Impact of LiDAR nominal post-spacing on DEM accuracy and flood zone delineation. Photogrammetric Engineering & Remote Sensing 73 (7), 793–804.
Smith, S.L., Holland, D.A., Longley, P.A., 2005. Quantifying interpolation errors in urban airborne laser scanning models. Geographical Analysis 37 (2), 200–224.
Shi, W.Z., Li, Q.Q., Zhus, C.Q., 2005. Estimating the propagation error of DEM from higher-order interpolation algorithms. International Journal of Remote Sensing 26 (14), 3069–3084.
Stoker, J., Parrish, J., Gisclair, D., Harding, D., Haugerud, R., Flood, M., Andersen, H.E., Schuckman, K., Maune, D., Rooney, P., Waters, K., Habib, A., Wiggins, E., Ellingson, B., Jones, B., Nechero, S., Navegandhi, A., Saultz, T., Lee, G., 2007. Report of the first national LiDAR initiative meeting, February 14-16, Reston, Va. USGS Land Remote Sensing Program, Geography Discipline. Report Series OF 2007-1189. http://lidar.cr.usgs.gov/presentations/NLImeetingReport.pdf (accessed 15.06.2009).
Stoker, J., Harding, D., Parrish, J., 2008. The need for a national Lidar dataset. Photogrammetric Engineering & Remote Sensing 74 (9), 1066–1068.
Su, J., Bork, E., 2006. Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. Photogrammetric Engineering & Remote Sensing 72 (11), 1265–1274.
Véga, C., St-Onge, B., 2008. Height growth reconstruction of a boreal forest canopy over a period of 58 years using a combination of photogrammetric and lidar models. Remote Sensing of Environment 112 (4), 1784–1794.
Voltz, M., Webster, R.A., 1990. A comparison of kriging cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science 41 (3), 473–490.
Weibel, R., 1992. Models and experiments for adaptive computer-assisted terrain generalization. Cartography and Geographic Information Science 19 (3), 133–153.