

available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/compag

Original paper

A post-processing step error correction algorithm for overlapping LiDAR strips from agricultural landscapes[☆]

Jeffrey Willers^{a,*}, Mingzhou Jin^b, Burak Eksioğlu^b, Andy Zusmanis^c,
Charles O'Hara^d, Johnie Jenkins^a

^a Genetics and Precision Agriculture Research Unit, USDA-ARS, Mississippi State, MS, United States

^b Department of Industrial and Systems Engineering, Mississippi State University, United States

^c Leica Geosystems, Integrated Solutions Group, Norcross, GA, United States

^d GeoResources Institute, Mississippi State, MS, United States

ARTICLE INFO

Article history:

Received 3 August 2007

Received in revised form

9 April 2008

Accepted 25 April 2008

Keywords:

LiDAR

Step error

Digital surface model

Agriculture

Optimization model

ABSTRACT

In the processing of light detection and ranging (LiDAR) data, a step error is an abrupt change in estimates of elevation between adjacent strips and must be reduced before building a digital surface model (DSM) of elevation. Existing methodologies in the literature for removing this artifact require an analyst to (1) utilize the sensor and aircraft information of the LiDAR mission, (2) isolate homologous flat surfaces within regions of overlap of adjoining LiDAR strips to estimate the mean offset, or (3) a combination of the two. In this application involving an agricultural landscape, a different methodology was required because the necessary information from the laser scanner or internal navigation system (INS) of the aircraft was unavailable and it was not possible to successfully identify homologous flat surfaces. Therefore, a post-processing, quadratic optimization model was formulated to reduce step artifacts. Using statistics obtained from the geographic overlap of the strips with a benchmark strip, it was possible to determine from the elevation values of the LiDAR point clouds two quantities: the strip variance and the total variance. Using these values and related statistics, the optimization model estimated correction constants, called decision variables, that minimized the among-group variance of the adjoining strips. When the values of these decision variables are added to the point cloud elevations of their respective LiDAR strips, the systematic step errors among adjoining strips are minimized with respect to the elevations provided by the point cloud of the benchmark strip. Decision variable values ranged between -0.087 and 0.078 m. The adjusted LiDAR strip point clouds were used to build a corrected DSM of a 638.2-ha agricultural landscape at a spatial resolution of 0.5 m. The elevation range of the DSM is approximately 44–81 m HAE (height above the ellipsoid), where the higher elevations are the tops of trees. Effectiveness of the optimization model approach to reduce the step errors was evaluated by comparing the DSM before and after adjustment. Several hillshade, gray scale image subsets, and profile plot comparisons between the before and after adjustment of the point clouds of the LiDAR strips illustrate the algorithm's performance in reducing step error effects.

Published by Elsevier B.V.

[☆] Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendations or endorsement by the U.S. Department of Agriculture.

* Corresponding author at: P.O. Box 5367, Mississippi State, MS 39762, United States. Tel.: +1 662 320 7383.

E-mail address: jeffrey.willers@ars.usda.gov (J. Willers).

0168-1699/\$ – see front matter. Published by Elsevier B.V.

doi:10.1016/j.compag.2008.04.013

1. Introduction

The capability to create digital surface models (DSMs) of elevation for rural or urban landscapes is facilitated by light detection and ranging (LiDAR) or laser scanning sensor systems (Axelsson, 1999; Baltsavias, 1999; Jensen, 2000; Wehr and Lohr, 1999). A DSM describing landscape elevations creates opportunities for solving many problems (Ackermann, 1999; Filin, 2004). DSMs of elevation have been widely applied in forestry (Kraus and Pfeifer, 1998; Means et al., 2000; Popescu et al., 2002), bare-earth extraction (Sithole and Vosselman, 2004), urban planning (Shan and Sampath, 2005) and many other applications (Barnes et al., 1990; Hollaus et al., 2005; Leyva et al., 2002). It is essential that LiDAR data be of high quality (Latypov, 2002) in all of these applications, particularly for agricultural fields that have low relief.

Sources of errors in LiDAR data can be apportioned into (1) random and/or (2) systematic causes and (3) blunders (Baltsavias, 1999; Huising and Pereira, 1998). Blunders and/or random errors are corrected by manual or automated filtering methods (Fritsch and Kilian, 1994). Systematic errors are due to characteristics of the laser scanner itself and/or the internal navigation system (INS) of the aircraft carrying the scanner. Systematic errors, like the other sources, must also be corrected in order to produce good quality surface models of elevational relief (Morin, 2002; Skaloud and Lichti, 2006; Vosselman, 2002).

An attempt to build a DSM of elevation for an agricultural landscape near Gunnison, MS, discovered that systematic step errors among the LiDAR strips compromised its quality (Fig. 1). A step artifact (Crombaghs et al., 2000) is one type of systematic error that represents an abrupt change in elevation that occurs among adjacent LiDAR strips. Luethy and Ingensand (2001) discussed issues of quality control for LiDAR data and stated that step errors were one of six characteristics limiting LiDAR quality, citing poor calibration or poor navigation data as the most probable causes. In an agricultural DSM,

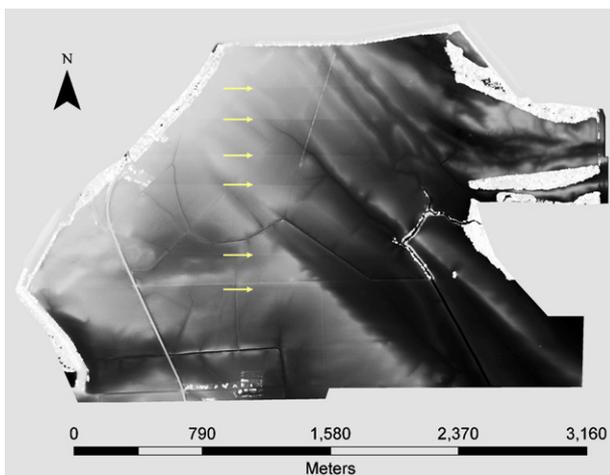


Fig. 1 – Grayscale image of the digital surface model (DSM) before step error corrections. Six of the largest step errors (identified by arrows) are apparent as horizontal stripes across the image.



Fig. 2 – View of the cultivated land showing the undulations of the seed bed and furrows with 1.02 m row spacing.

step errors create cliffs or terraces not actually found on the ground. When other attributes such as curvature, flow accumulation, and slope (Chang, 2006) are derived from the DSM, these artifacts are propagated, compromising the value of these hydrological derivatives (ESRI®, 2005). Our interest in LiDAR-based data products as topographical covariates for the analyses of site-specific, agricultural experiments (Willers et al., 2004, 2008) means the effects of step artifacts must be reduced.

According to Pfeifer (2005), algorithms (or models) used to control systematic errors are of two general types: (1) sensor system models and (2) data driven models. A sensor model approach (Burman, 2002; Filin, 2003; Kager, 2004; Morin, 2002) relies on information collected during the calibration of the scanner and/or at the time of data acquisition. For our LiDAR data, a sensor model was not applicable because the necessary information about the characteristics of the laser scanner and the INS data from the aircraft was unavailable. This meant that some type of data driven approach would have to be utilized as a post-processing method.

Data driven (e.g., Crombaghs et al., 2000; Pfeifer, 2005) and some sensor system (e.g., Filin, 2003; Skaloud and Lichti, 2006) approaches use numerous homologous areas of flat terrain such as parking lots, flat rooftops, or sport fields to estimate step error correction values. The issue of flatness should not be lightly regarded, because as Skaloud and Lichti (2006) remark, it can be difficult to find sufficiently flat natural terrain surfaces, even on features such as athletic fields. In agricultural landscapes, identifying homologous patches of flat terrain among LiDAR strips is more difficult. Tillage patterns (Fig. 2), interference from center pivots (Fig. 3), vegetation on field roads (Fig. 4) or other irregularities on unpaved roads (Fig. 5) cause changes in relief that are often greater than the sizes of the step errors.

Since a sensor model approach was impossible to employ and homologous areas of flat terrain could not be isolated for use with published data driven models, we needed to investigate an alternative approach. Non-linear optimization models are widely used in many areas of problem solving



Fig. 3 – View of a center irrigation pivot.

(Himmelblau, 1972) as an alternative to regression analyses when its basic assumptions cannot be satisfied (Ignizio, 1982). Our objective is to describe the development, application, and results of a quadratic optimization model as an alternative data driven methodology to reduce LiDAR step errors. To accomplish this objective, Section 2 first describes the LiDAR data set and its general processing into a DSM. We next present the formulation of the quadratic optimization model. Section 3 discusses several case studies that helped develop and test the capabilities of the optimization model. The effect of the model's corrections upon the elevation value of the LiDAR point clouds used to build a corrected DSM is described. These impacts are examined by comparing two smaller subsets geographically extracted from the pre- and post-correction DSMs. A representative set of profile plots from another geographic location compared the before and after adjustment corrections upon LiDAR point clouds with respect to point elevations from an orthogonal LiDAR strip, called the Tie Line. In Section 4, additional comments describe the optimization model



Fig. 4 – View of a non-paved field road showing that, at times, vegetation residues prevent them from being used as a source of flat terrain for data driven step error correction methods.



Fig. 5 – View of a barren field road showing wetter (dark) and drier (light) regions due to effects of small depressions and ridges, several days after a rain. Recent traffic patterns are also shown. These are other effects that exclude agricultural roads as a source of flat terrain for data driven step error correction methods.

and its application. One benefit was the discovery of an additional source of random error. Histograms of point elevations from the geographic overlap of the fifth strip and the Tie Line are compared. This exercise illustrates that estimates of mean deviations cannot reduce step errors whenever homologous areas between strips are not sufficiently flat. A correlation analysis further demonstrated that an optimization model approach is superior for reducing step artifacts when homologous flat areas cannot be utilized.

2. Materials and methods

2.1. Study site and LiDAR data set

LiDAR acquisition (3 June 2003) over a part of Perthshire Farms, Gunnison, MS and preliminary processing of the strips were completed by Earthdata Aviation[®] (Hagerstown, MD) and delivered as binary (Schuckman, 2003) files. The LiDAR point clouds of the strips described the elevational relief in meters above the ellipsoidal height (HAE) using the NAVD88 datum (using GEOID99) in the Universal Transverse Mercator (UTM) co-ordinate system.

To reduce costs, the vendor was requested to not process the LiDAR data into a DSM. The binary files were converted into point theme shapefiles to allow for wider usage in research. As shapefile point themes, the original eight digit Earthdata[®] labels for the east–west LiDAR strips were re-named (from north to south) as Strips 1 through 13 (Fig. 6). Also, the north–south LiDAR strip near the center of the east–west strips was re-named as the Tie Line.

Using Geographic Information System (GIS) processing, the next step was to reduce the size of the original strips, nearly 6 km in length, into smaller lengths that spanned a specific collection of cotton fields (Fig. 6). These fields were

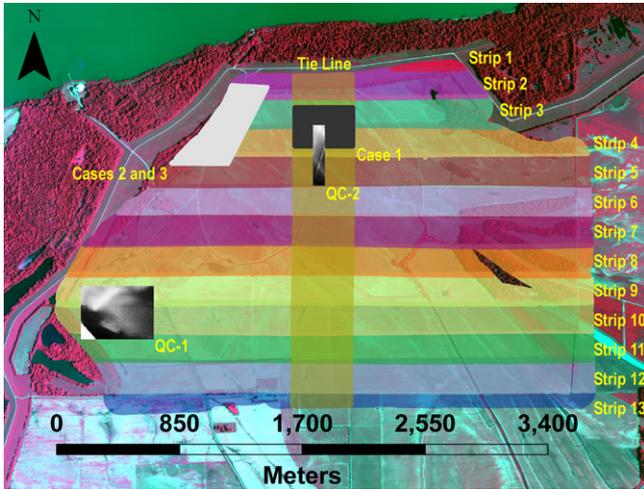


Fig. 6 – Illustration showing the relationships among the LiDAR strips (east–west) and the Tie Line (north–south) strip for an agricultural area of interest located on Perthshire Farms, Bolivar County, MS. Locations of the Cases 1–3 and two quality control (QC-1 and 2) areas of interest are also indicated.

the requested acquisition target of the LiDAR mission on the cooperating farm. After this, using commercially available software, the elevation attributes from the dBASE (DBF) table of these shapefiles (ESRI, 1998) were filtered for blunders, merged, and processed into a DSM. Results were visually assessed using a hillshade image that simulates shadows as if the sun was at a particular orientation. While the vendor provided a LiDAR data set of excellent quality, numerous step artifacts of small magnitudes were readily apparent throughout the hillshade image of the uncorrected DSM. Fig. 1 represents a gray scale, raster, image of the same area as the hillshade image—it also shows several of the larger-sized step errors.

2.2. Development of the optimization model

Preliminary inspections of the LiDAR point clouds estimated that these step errors were up to 15 cm in size and were of varying sizes among the strips. Removing them by manual methods or by focal processing (or other smoothing) techniques was considered, but not attempted. Manual methods would be too laborious and/or unreliable. While focal or smoothing techniques may improve the visual appearance of the DSM, they would not necessarily provide a better product for analysis. To overcome these limitations, a data driven methodology using an optimization model was developed.

This effort began with GIS processing to (1) isolate the points (or returns) of each strip that geographically overlapped with the points of the Tie Line (Fig. 6), (2) label them by strip, (3) merge them, (4) calculate their co-ordinate attributes, and (5) then create a new shapefile of point elevations. By definition, the overlap area of all strips with the Tie Line contains point observations from a total of L strips. In the optimization model, the index i is used to indicate a given strip. It is assumed that strip $i=L$ is the Tie Line which defines the benchmark

elevations at its geographic intersection with each strip. The x -coordinate, y -coordinate, raw elevations of point j in strip i (x_{ij} , y_{ij} , and e_{ij}), and the decision variables, a_i , of the optimization model are all measured in meters, where a_i is the elevation adjustment of strip i . There is no decision variable associated with the Tie Line because it is not adjusted; the other strips ($i=1, 2, \dots, L-1$) are adjusted to reduce step error effects.

Using the new shapefile, points from each strip, along with the points from the Tie Line, are divided into groups. These groups ensure that the method is robust and is not affected by the lack of identical co-ordinate locations among the points from different LiDAR strips. To accomplish this, the LiDAR points from a geographically defined area of overlap are categorized into K groups indexed by k ; where the set of points in group k is such that $S_k = \{(i,j) | \text{Sub } X_k \leq x_{ij} < \text{Sup } X_k \text{ and } \text{Sub } Y_k \leq y_{ij} < \text{Sup } Y_k\}$. The number of points in strip i in group k is denoted by n_{ik} , and the total number of points in group k is denoted by n_k ; thus, $n_k = \sum_{i=1}^L n_{ik}$. A convenient way to conceptualize this step is to consider all the points found in the region of overlap as a list sorted by their co-ordinate positions. This list is binned (or divided) into K binning groups each containing n_k points (approximately 3000–5000 points). The strip label attribute of the point is used to determine the number of points (n_{ik}) in the k th group ($k=1, 2, \dots, K$) of the i th strip ($i=1, 2, \dots, L-1$).

The next step is to exclude from further analysis any binning groups whose raw elevations have a standard deviation greater than a threshold value (t , in meters) specified by the analyst according to the characteristics of the landscape. In other words, the collection of K groups is reduced to a smaller set of groups, K' , according to the following criteria:

$$K' = \left\{ k \mid \left| \frac{\sqrt{\sum_{i=1}^L \sigma_{ik}^2}}{L} \right| \leq t \right\} \quad (1a)$$

where

$$\sigma_{ik}^2 = \frac{\sum_{i,j \in S_k} (e_{ij} - E_{ik})^2}{n_{ik} - 1} \quad (1b)$$

and

$$E_{ik} = \frac{\sum_{i,j \in S_k} e_{ij}}{n_{ik}} \quad (1c)$$

The criteria of steps (1a)–(1c) exclude binning groups having variances larger than the threshold value. Large variance values in a binning group may occur due to the blending of the point elevations from different topographical features. For example, some of the points in a binning group could be returns from a power pole, a center irrigation pivot, or other tall items that were captured by a single strip but not detected by the others within the geographic location of a particular binning group. In this instance, that binning group would have a standard deviation greater than the threshold and would be excluded from the analysis. The employment of binning groups resolved the difficulty caused by the lack of numerous, small patches of homogenous flat terrain in areas of overlap among the LiDAR strips. The above steps (1a)–(1c) do not filter

blunders or random errors—other standard processing steps accomplished this task during the building of a DSM.

Once the final set of binning groups is determined, let B_k be the mean elevation of group k before adjustment and A_k be the mean elevation for the same group after adjustment ($k \in K'$). These mean values are found using:

$$B_k = \frac{\sum_{(i,j) \in S_k} e_{ij}}{n_k} \quad \text{and} \quad A_k = \frac{\sum_{(i,j) \in S_k} (e_{ij} + a_i)}{n_k} = B_k + \sum_{i=1}^{L-1} \frac{n_{ik}}{n_k} a_i \tag{2}$$

Let V_k be the sum of squared residuals in group $k \in K'$ after adjustment:

$$V_k = \sum_{(i,j) \in S_k} (e_{ij} + a_i - A_k)^2 \tag{3}$$

Note that with the inclusion of the decision variable, a_i , expression (3) is not equivalent to the squared residuals term of ordinary least squares regression.

In order to minimize the step errors, the sum of squared residuals of the groups is minimized by determining the values of decision variables a_i , $i = 1, 2, \dots, L - 1$, which leads to the following unconstrained optimization problem:

$$\min_{a_i} \sum_k V_k = \sum_k \sum_{(i,j) \in S_k} (e_{ij} + a_i - A_k)^2 \tag{4}$$

Since (4) is a convex objective function (Himmelblau, 1972), existing optimization solvers work well to obtain a unique optimal solution. Notice that the estimated strip adjustment, i.e., the value of the decision variable a_i , does not affect the within-strip variances but will minimize the variances among the strips. In effect, the values of the decision variables a_i , when added to the elevations of the respective point clouds, shifts the i th strip closer to the elevations of the Tie Line's point cloud. By iteration, the optimal solution for each strip is determined. Once a strip is adjusted, the value of its decision variable is treated as a constant in subsequent iterations that estimate the decision variable for the next strip.

The minimization problem (4) is equivalent to

$$\min_{a_i} \sum_k \sum_{(i,j) \in S_k} \left[2(e_{ij} - B_k) \left(a_i - \sum_{l=1}^{L-1} \frac{n_{lk}}{n_k} a_l \right) + \left(a_i - \sum_{l=1}^{L-1} \frac{n_{lk}}{n_k} a_l \right)^2 \right] \tag{5}$$

which is equal to

$$= \sum_{l=1}^{L-1} 2 \left\{ \sum_k \sum_{(i,j) \in S_k} (e_{ij} - B_k) - \sum_k \sum_{(i,j) \in S_k} (e_{ij} - B_k) \frac{n_{lk}}{n_k} \right\} a_l + \sum_{l=1}^{L-1} \left(n_l - \sum_k \frac{n_{lk}^2}{n_k} \right) a_l^2 - \sum_{l=1}^{L-1} \sum_{s=1, s \neq l}^{L-1} 2a_l a_s \left(\sum_k \frac{n_{lk} n_{sk}}{n_k} \right)$$

The first order derivatives after dividing by 2 are

$$\frac{\partial \sum_k V_k}{2 \partial a_l} = \sum_k \left\{ \sum_{(i,j) \in S_k} (e_{ij} - B_k) - \sum_{(i,j) \in S_k} (e_{ij} - B_k) \frac{n_{lk}}{n_k} \right\} + \left(n_l - \sum_k \frac{n_{lk}^2}{n_k} \right) a_l - \sum_{s=1, s \neq l}^{L-1} a_s \left(\sum_k \frac{n_{lk} n_{sk}}{n_k} \right) \tag{6}$$

$l = 1, \dots, L - 1$

Expression (5) is a convex function, so one can either solve it or the linear system (6) set equal to zero.

Decision variables for each LiDAR strip with respect to their geographic overlap with the Tie Line were estimated by this optimization model methodology. Strip 1 was excluded because it did not contain any points from the cultivated land.

2.3. Evaluation of the optimization model

We established several case studies to evaluate the quadratic optimization model (Eq. (4)). The first set of test data was subset from Strips 3–5, and the Tie Line by GIS procedures. The specific task of this test, named Case 1 (Fig. 6), was to determine if the optimization model could iteratively estimate up to three decision variables with respect to the Tie Line, which was the benchmark strip. The second test, Case 2 (Fig. 6), tested the generality of the Case 1 solutions to another set of points from the same three strips located to the west of Case 1. At the Case 2 location, there were no Tie Line points. The third test, named Case 3 (Fig. 6), while spatially coincident with Case 2, evaluated the filtering steps (Eq. (1a)–(1c)) when sharp contrasts in elevation occurred between woods and cultivated land. However, in Case 3, Strip 4 now functioned as the benchmark strip. New decision variables were estimated for Strips 3 and 5 using the overlaps along the edges of these point clouds with the edges of Strip 4's point cloud. These lateral regions of overlap defined the geographic location of the binning groups used in Case 3.

The General Case (Fig. 6) estimated decision variables for Strips 2–13 with respect to their geographic overlaps with the Tie Line. The estimated values of the decision variables were added to all of the point elevations of their corresponding strip. This process defined a new elevation attribute used to build a corrected DSM. Comparison of results was accomplished by visual inspection of the pre- and post-correction DSMs (Figs. 1 and 9, respectively).

2.4. Assessments using subsets

Additional visual inspections evaluated the optimization model's effectiveness by comparing smaller images that were subset from the pre- and post-correction DSMs. Profile plots in two co-ordinate directions from a small set of points extracted from the point clouds of Strips 4–6 and the Tie Line provided additional assessments. Comparative histograms of the unadjusted point elevations from the geographic overlap between Strip 5 and the Tie Line show the effects of scanning angle, scanning rate, and flight direction upon the shape of the data

distributions. Finally, a correlation analysis was performed between (1) the simple mean deviation estimates of the strip and Tie Line elevations and (2) the respective decision variable values from all regions of geographic overlap.

3. Results

Using a color infrared image from June 2004 as a background, Fig. 6 illustrates the arrangement of the LiDAR strips with respect to the Tie Line. Also shown are the geographic extents of three case studies (Cases 1–3) and two quality control (QC-1 and 2) locations.

3.1. Case 1 (estimation of decision variables for three LiDAR strips that overlap the Tie Line)

The Case 1 data subset (Fig. 6, center) from Strips 3, 4, 5, and the Tie Line involved a ‘four-fold’ overlap of points. The sample size of extracted points was 86,305 for Strip 3, 111,035 for Strip 4, 94,210 for Strip 5, and 149,171 for the Tie Line. Using 3000–5000 LiDAR points in any one binning group (Section 2.2), a large number of them were established to estimate the decision variables. For conciseness, the actual sample sizes are only reported for Case 1, since similar, if not larger, sample sizes arose for the three additional cases.

Three decision variables were iteratively estimated for Strip 3 ($a_1 = -0.049$ m), Strip 4 ($a_2 = -0.033$ m), and Strip 5 ($a_3 = 0.059$ m). A visual inspection of pre- and post-correction hillshade surfaces (not presented) for a DSM for the Case 1 area of interest showed the step artifacts were successfully reduced.

3.2. Case 2 (applicability of solutions beyond the geographic extent of the Tie Line overlap)

In Case 2 (Fig. 6, upper left), the generality of Case 1 solutions was tested by applying them to a new set of points selected from Strips 3, 4, and 5 at a location approximately 716 m to the west. Hillshade surfaces of these point elevations before (Fig. 7A) and after (Fig. 7B) correction again showed the successful removal of the step errors. The Case 2 LiDAR point clouds did not geographically overlap with the Tie Line point cloud.

For extremely long LiDAR strips, another investigator (Maas, 2002) found curvi-linear trends in the offset bias as their length increased. However, for our LiDAR data, where the strip lengths were less than 3 km, there is no evidence of a curvi-linear trend. If there had been a curvi-linear trend among adjacent strips of this LiDAR acquisition, the estimated decision variables provided by the quadratic optimization model would fail to reduce them.

3.3. Case 3 (effects of elevational contrasts among feature types upon obtainment of solutions)

The two previous cases (1 and 2) applied correction values using LiDAR point clouds acquired over cultivated land. A few returns due to a center irrigation pivot feature and the grassy vegetation below it (Fig. 3) were present within the Case 1

area of interest. However, the returns from the pivot or the grasses on the pivot service road were a very small proportion of the total number of points found in that area of interest. Their effect upon the estimated values of the decision variables would be small, even if returns from these features were not excluded. As a consequence, Case 1 was not the best scenario to determine if the binning groups and/or the application of Eq. (1a)–(1c) actually were robust concepts. Therefore, we thought it was necessary to develop a more rigorous test where features of mixed elevations comprised a large proportion of the total number of points.

Utilizing Strip 4 as the benchmark elevation at the geographic location of Case 2 (Fig. 6), it is seen (Fig. 7A) that a large proportion of points are from both high (treetops) and low elevation (seed beds or furrows) features. In this type of overlap with Strips 3 and 5, some of the binning groups contained points from both types of features, particularly at the interface of different terrain features. As described in Section 2.2, all binning groups having variances smaller than the threshold value (Eq. (1a)) would be used to estimate the decision variable regardless of whether the feature type was the tops of tree in the forest or from the cultivated land. Other binning groups containing mixed features having large elevational differences would exceed the threshold value and be excluded from the optimization model.

New decision variables were estimated for Strip 3 ($a_1 = -0.022$ m) and Strip 5 ($a_2 = 0.077$ m) with respect to Strip 4. The values of these Case 3 decision variables are different than the values used for Strips 3 and 5 in Cases 1 and 2 because here there was no region of geographic overlap with the Tie Line. After correction, the hillshade surface (equivalent to Fig. 7B) of the DSM showed the algorithm estimated appropriate values for the two decision variables to reduce the step errors. The average difference in elevation between the Case 1 and Case 3 DSMs would be small and approximate the value of the Case 1 decision variable (-0.033 m) for Strip 4. Case 3 also demonstrated that it is not necessary for a benchmark strip to be orthogonal.

3.4. General case

3.4.1. Uncorrected digital surface model

For this agricultural landscape, where the majority of the slopes with the cultivated land are less than 1%, step artifacts were apparent in the uncorrected DSM (Fig. 1). These artifacts essentially disappeared wherever changes in relief were larger than the size of the step error biases. This is evident by visually tracing the step features into the wooded areas along the edges of the cultivated land.

3.4.2. Corrected digital surface model

Estimates for correction values of the strips ranged between -0.087 and 0.078 m (Fig. 8) and were smaller than the square of the root mean squared error (RMSE) of 3.0 cm of the parent data (a vendor provided specification). These estimates, when added to the original elevations of the corresponding strips by GIS processing, successfully reduced the systematic step errors. The grayscale image of the corrected DSM (Fig. 9) clearly shows improvements in quality.

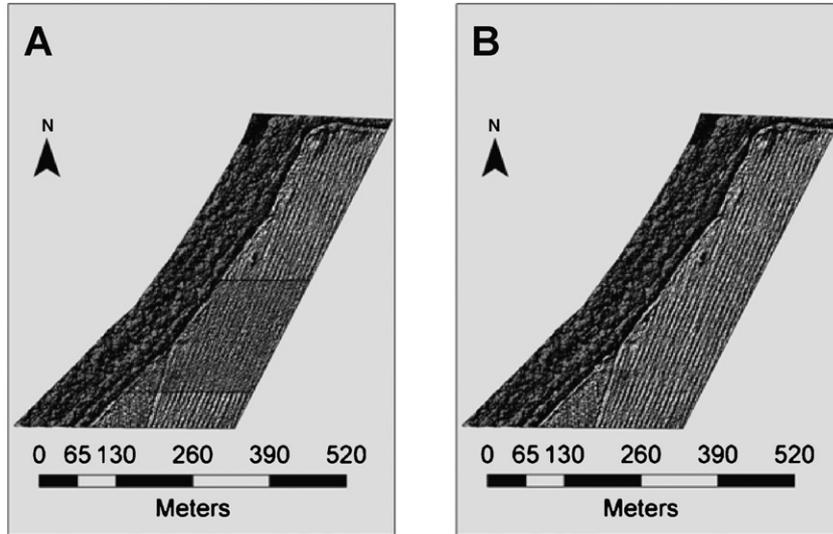


Fig. 7 – Hillshade of the digital surface model for the Case 2 and 3 areas of interest. Panel A shows the step artifacts before correction. Panel B indicates successful removal of the step artifacts after adding the decision variable corrections to the LiDAR points.

The elevation range of the DSM is approximately 44–81 m HAE (height above the ellipsoid), where the higher elevations are the tops of trees, and spanned a 638.2-ha agricultural landscape. During acquisition, the LiDAR strips were planned to have a 3-fold overlap for each unit of surface area across this

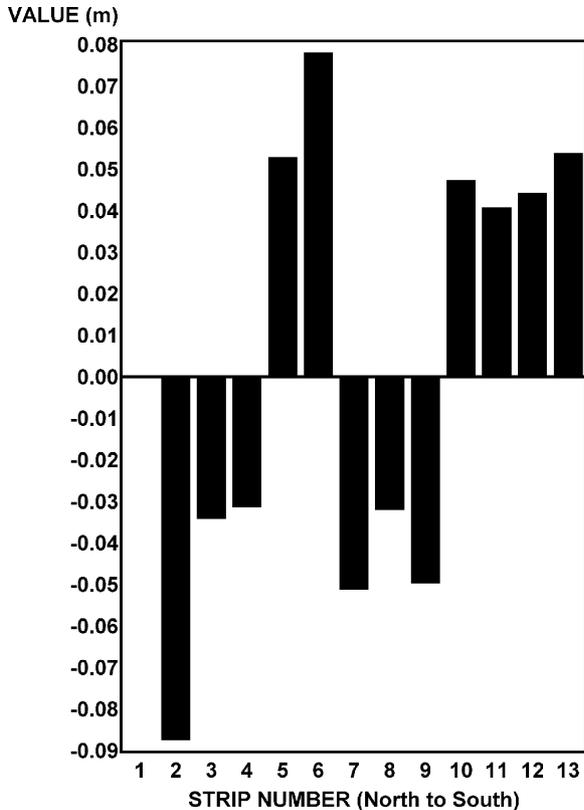


Fig. 8 – Bar graph depicting the magnitudes of the General Case decision variables for each LiDAR strip estimated by the optimization model. No estimate was obtained for Strip 1.

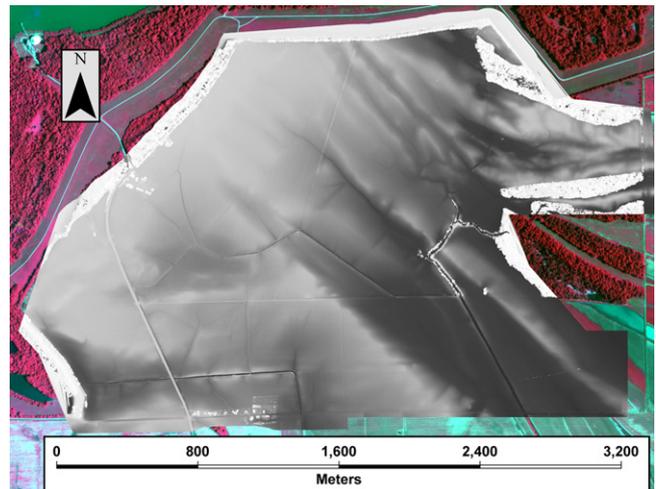


Fig. 9 – Grayscale image of the digital surface model after application of the step error corrections. Compare to Fig. 1.

area of interest, which provided a data density great enough to build a DSM at a spatial resolution of 0.5 m.

3.5. Quality control (QC) assessments

A subset of the unadjusted DSM (Fig. 10) was extracted from the location labeled QC-1 (Fig. 6). A visual examination of this raster surface subset indicates lower elevations (dark gray or black) generally occur to the southwest; a higher elevation (white) ridge rises from the center and runs to the northeast, while an indistinct drainage feature runs north to south along the western edge. Along the bottom eastern half, alternating dark gray (lower elevations) and lighter gray (higher elevations) bands run parallel to one another in an east–west orientation. The feature in the extreme southwest corner are the tops of trees.

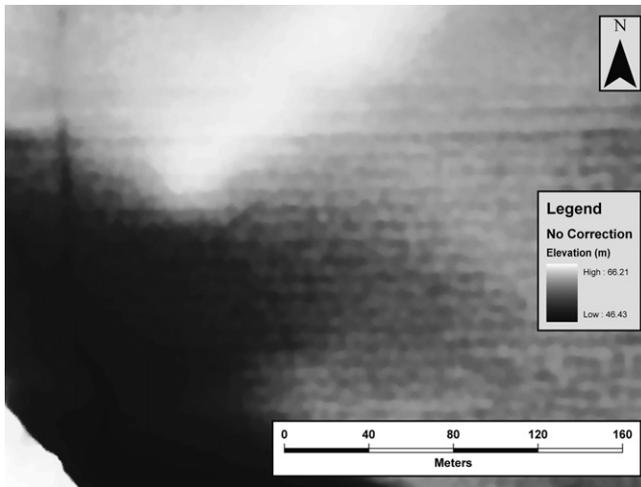


Fig. 10 – Grayscale image of a subset of the uncorrected DSM for the quality control site QC-1 (see Fig. 6). Note the blurred appearance of drainage features, roadways and tillage patterns.

For the same geographic location, Fig. 11 shows how the post-processing of the original LiDAR point clouds reduced the step artifacts and improved the quality of the DSM. The appearance of this surface is less blurred and various features are more distinctive. While the major features of the south-westerly basin and northeasterly ridge are still seen, it is now obvious that the north-south drainage feature along the western edge consisted of two parallel lines of different widths. Inspection on site showed this to be a drainage feature built by a tractor drawn implement that created two parallel ditches to improve surface drainage. The western cut received more water erosion than the eastern cut, causing it to widen with time.

The east-west bands lying parallel to one another along the eastern quadrant of the bottom half of both the uncorrected

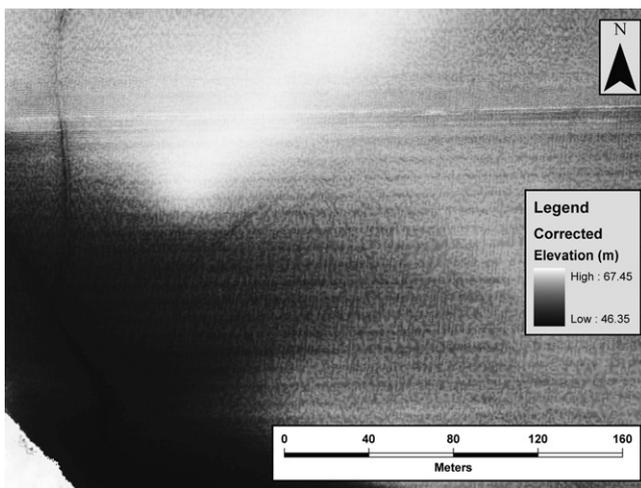


Fig. 11 – Grayscale image of a subset of the corrected DSM for the quality control site QC-1 (see Fig. 6). Note the increased clarity of drainage features, roadways and tillage patterns. Compare to Fig. 10.

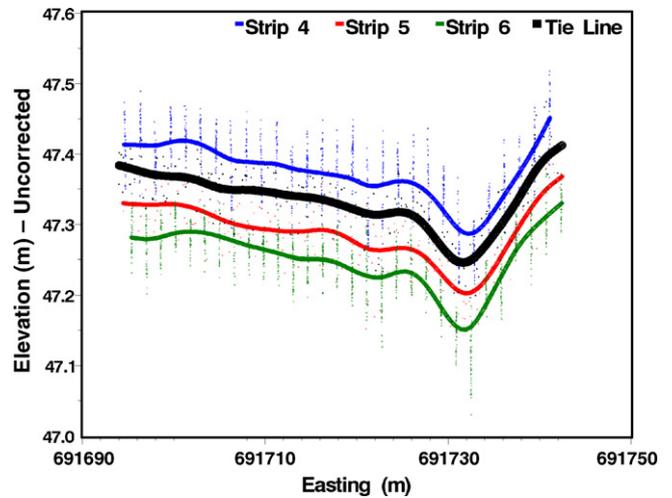


Fig. 12 – Profile plot of the uncorrected strips with respect to the Tie Line viewed by the Easting co-ordinate for the quality control site QC-2. The sharp dip on the right is the drainage feature visible just beneath the middle pivot section in Fig. 3. Linear splines were fit to the point clouds of each strip for additional clarity.

and corrected DSMs (Figs. 10 and 11) are the results of long term tillage operations. These parallel bands are not as distinct in the northern part of these DSM subsets, because the slope of the land there is gentler, reducing the water flow rate. In contrast, the slope increases east to west moving toward the south to produce the darker bands at approximately 10m spacings, where tillage passes by eight-row equipment abut one another. A probable explanation for these parallel bands is that the rate of erosion is greater along the margins than the middle of the tillage passes. Parallel patterns corresponding to tillage operations are found throughout the corrected DSM (see also Fig. 7B).

Profile plots from a second location labeled QC-2 (Fig. 6) provided another examination of the optimization model's ability to estimate decision variables to alleviate step errors. The before (Figs. 12 and 13) and after (Figs. 14 and 15) profile plot comparisons for Strips 4-6 clearly indicate reductions in step error magnitudes. These effects are highlighted by spline lines fit to the different point clouds. The closely spaced points occurring as linear patterns in the profile plots viewed along the Easting co-ordinates (Figs. 12 and 14) belong to similar scan lines within the strips. The remaining width of the point cloud profiles for these lateral views of the point clouds in the Easting and Northing directions, even after correction, is due primarily to the alternating undulation of the crop seed bed and the furrows (Fig. 2), and/or returns from the grassy vegetation of the center pivot service road (Fig. 3). Returns from the center pivot itself (Fig. 3) have been excluded from all profile plots to enhance the scaling of the y-axis.

4. Discussion

Case 1 estimates for Strips 3, 4, and 5 (Section 3.1) were slightly different than their corresponding estimates in the General

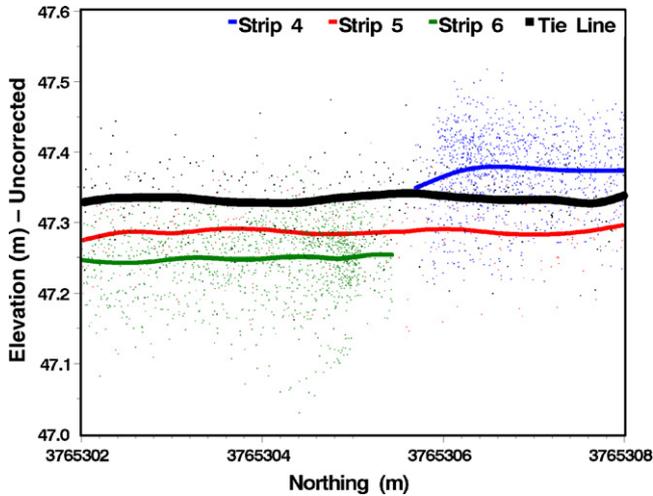


Fig. 13 – Profile plot of the uncorrected strips with respect to the Tie Line viewed by the Northing co-ordinate. Note the obvious step error discrepancy between the point clouds of Strips 4 and 6.

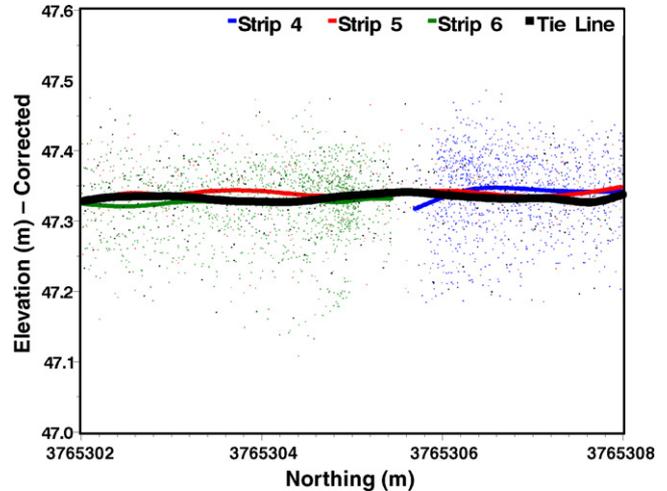


Fig. 15 – Profile plot of the uncorrected strips with respect to the Tie Line viewed by the Northing co-ordinate. Note the removal of the obvious step error discrepancies among the point clouds of the strips in comparison to Fig. 13.

Case (Fig. 8). An explanation is the fact that these strips had to be reconciled to other adjoining strips that were to the north and south of them. For example, Strip 3 geographically overlapped Strip 2 at the same time it overlapped Strip 4. Strip 5 geographically overlapped Strip 6 at the same time it overlapped with Strip 4, and so on. We demonstrated in Case 1 that the optimization model concurrently works with sets of adjoining strips. Strips 2 and 6 were excluded during Case 1, but were included in the General Case application.

The removal of the step errors as shown in the corrected DSM (Fig. 9) and the quality control illustrations (Figs. 11, 14 and 15) were due to the addition of the estimated decision variable values to the elevation attribute of each point

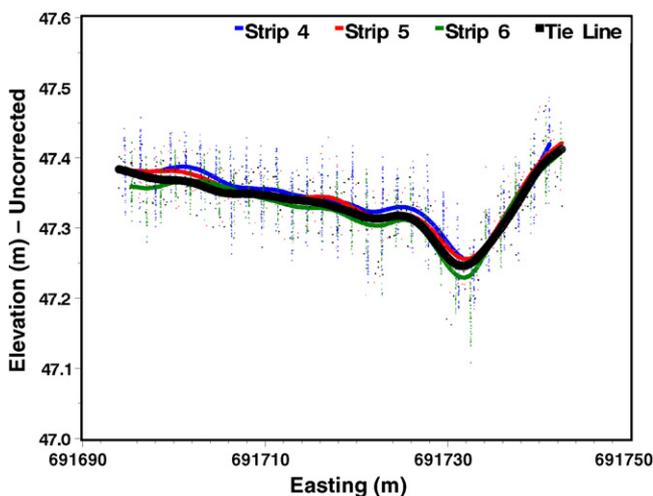


Fig. 14 – Profile plot of the corrected strips with respect to the Tie Line viewed by the Easting co-ordinate for the quality control site QC-2. The point clouds of the strips better conform to the elevations of the Tie Line in contrast to Fig. 12. Linear splines were fit to the point clouds of each strip for additional clarity.

of the respective LiDAR strip. No additional smoothing or focal processing (Chang, 2006; Theobald, 2003) algorithms were applied while building the corrected DSM with the adjusted elevation values of the LiDAR point clouds.

One supplemental benefit of our processing efforts was the discovery of a source of random error in these data. Varying point density (Fig. 16) due to geographic gaps between adjoining strips also created linear features at some locations which primarily show up as textural differences in the corrected DSM. The locations of a few gaps due to this cause were easily found by visual inspection of the hillshade image of corrected elevations (not shown). The effect of gap errors was most severe wherever the distance between adjacent LiDAR strips exceeded 4 m. Thus, this corrected DSM will require additional GIS processing in those areas wherever the point sample intensity varied due to gaps of random widths. Hu and Tao (2005) and Filin and Pfeifer (2005) discussed processing LiDAR data that vary in point density.

Whenever homogenous areas of flat terrain are unavailable it is not possible to optimally estimate simple mean deviations between strips. When a common geographical area is not flat, effects of scan angle and rate, reflectance signatures, and direction of strip acquisition (Burman, 2002; Morin, 2002) affect the frequency distribution of point elevations from different strips (Fig. 17). On non-flat terrain, these cumulative effects determine the shape of the frequency distribution of elevation attributes of the different strips obtained from a common area of geographic overlap and influence the estimates of their respective mean elevations. As a consequence, the difference between the strip means for areas of geographic overlap will not optimally estimate the size of the step error bias. These comparative histograms also demonstrate that LiDAR elevation data from homologous, non-flat, areas are not normally distributed. An optimization model methodology that seeks to minimize the among-strip variance will perform better (Ignizio, 1982) when all these types of influences arise.

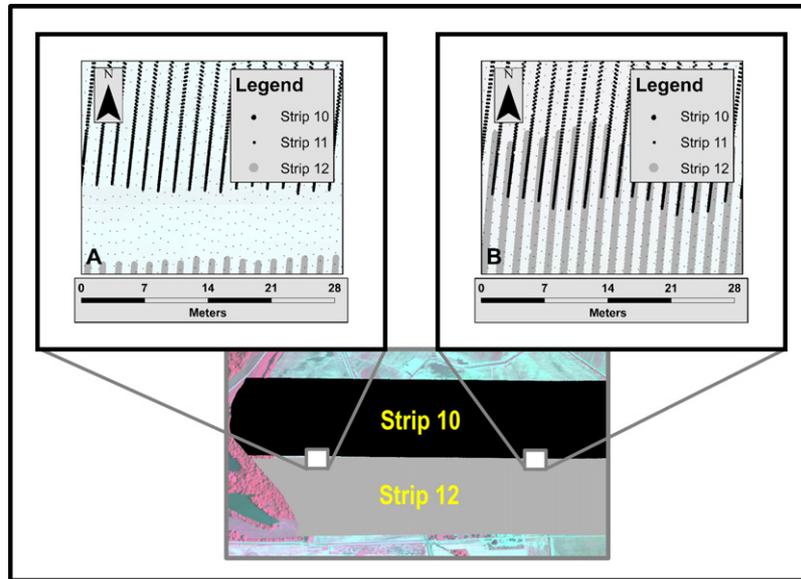


Fig. 16 – Illustration of gap (or sampling intensity, Panel A) errors between Strips 10 and 12 at a selected location along their lengths. This type of random error may cause textural differences in the digital surface model after step error biases are corrected. Panel B shows the absence of the gap error between Strips 10 and 12 at another location. Strip 11 is centered between them as shown in the detailed views.

To further examine this issue, estimates of simple mean deviations of the thirteen strips and the Tie Line elevations across the fullest possible extent of the various geographic overlaps were obtained. A correlation analysis between the respective decision variables and simple differences between the mean estimates of the strips and Tie Line found a non-significant correlation of -0.26 ($P=0.4214$). This result illustrates the difficulty of estimating mean strip deviations by algebraic methods for agricultural landscapes. Therefore, the major benefit of this optimization model approach was the estimation of correction constants without the requirement for highly planar and homologous planar features.

In conclusion, the corrected DSM of the agricultural landscape was accepted as an accurate representation of relative

elevation since drainage patterns, tree heights, building relationships, etc. corresponded to features observed on site. Any remaining error due to the vertical bias of the Tie Line itself to the actual elevations of this landscape is likely to be small since the RMSE (3.0cm) of the unadjusted data was also small. The quadratic optimization model approach successfully reduced systematic step errors of small magnitudes.

Acknowledgements

Financial support was provided by Advanced Spatial Technologies for Agriculture (ASTA-322-298) and the USDA Area-Wide Tarnished Plant Bug Management Project (thru ARS CRIS Project 6406-21610-006-00D). Additional funding was provided by the USDA-ARS Genetics and Precision Agriculture Research Unit, Mississippi State, MS (ARS CRIS Project 6406-21610-007-00D). Appreciation is expressed to Dr. R.O. Bowden, Department of Industrial and Systems Engineering, Mr. Ronald E. Britton, USDA-ARS, Genetics and Precision Agriculture Research Unit, Mississippi State, Dr. D.B. Reynolds, Department of Plant and Soil Sciences, Mississippi State, MS, and Mr. Kenneth Hood, Perthshire Farms, Gunnison, MS, for their support. Approved for publication as Journal Article No. J-10876 of the Mississippi Agricultural and Forestry Experiment Station, Mississippi State University.

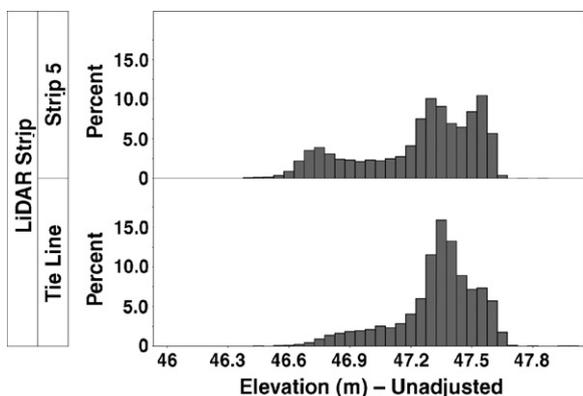


Fig. 17 – Comparative histograms of point elevations in the region of overlap between Strip 5 and the Tie Line before removing the step error. A small number of returns from the center pivot (see Fig. 3) between elevations of 48–52 m were excluded for clarity.

REFERENCES

Ackermann, F., 1999. Airborne laser scanning—present status and future expectations ISPRS. *Journal of Photogrammetry Remote Sensing* 54, 64–67.

- Axelsson, P., 1999. Processing of laser scanning data—algorithms and applications. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 138–147.
- Baltsavias, E.P., 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 199–214.
- Barnes, F.J., Karl, R.J., Kunkel, K.E., Stone, G.L., 1990. Lidar determination of horizontal and vertical variability in water vapor over cotton. *Remote Sensing of Environment* 32, 81–90.
- Burman, H., 2002. Laser strip adjustment for data calibration and verification. *International Archives of Photogrammetry and Remote Sensing* 34 (3), 67–72.
- Chang, K., 2006. *Introduction to Geographic Information Systems*, 3rd ed. McGraw Hill, Boston, MA.
- Crombaghs, M.J.E., Brügelmann, R., de Min, E.J., 2000. On the adjustment of overlapping strips of laser altimeter height data. *International Archives of Photogrammetry and Remote Sensing* 33 (B3/1), 230–237.
- ESRI, 1998. *ESRI Shapefile Technical Description*. ESRI, Redlands, CA.
- ESRI, 2005. *Arc Hydro Tools Overview, Version 1.1*. ESRI, Redlands, CA.
- Filin, S., 2003. Recovery of systematic biases in laser altimetry data using natural surfaces. *Photogrammetric Engineering & Remote Sensing* 69 (11), 1235–1242.
- Filin, S., 2004. Surface classification from airborne laser scanning data. *Computers & Geosciences* 30, 1033–1041.
- Filin, S., Pfeifer, N., 2005. Neighborhood systems for airborne laser data. *Photogrammetric Engineering & Remote Sensing* 71 (6), 743–755.
- Fritsch, D., Kilian, J., 1994. Filtering and calibration of laser scanner measurements. *International Archives of Photogrammetry and Remote Sensing* 30 (3), 227–234.
- Himmelblau, D.M., 1972. *Applied Nonlinear Programming*. McGraw-Hill, New York.
- Hollaus, M., Wagner, W., Kraus, K., 2005. Airborne laser scanning and usefulness for hydrological models. *Advances in Geosciences* 5, 57–63.
- Hu, Y., Tao, C.V., 2005. Hierarchical recovery of digital terrain models from single and multiple return lidar data. *Photogrammetric Engineering & Remote Sensing* 71 (4), 425–433.
- Huising, E.J., Pereira, L.M., 1998. Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications. *ISPRS Journal of Photogrammetry Remote Sensing* 53 (5), 245–261.
- Ignizio, J.P., 1982. *Linear Programming in Single- & Multiple Objective Systems*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Jensen, J.R., 2000. *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice-Hall, New Jersey, pp. 285–332.
- Kager, H., 2004. Discrepancies between overlapping laser scanner strips—simultaneous fitting of aerial laser scanner strips. In: *International Archives of Photogrammetry and Remote Sensing, Proceedings of the XX ISPRS Congress, Istanbul, Turkey, July 12–23 (CD-ROM)*.
- Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry Remote Sensing* 53 (4), 193–203.
- Latypov, D., 2002. Estimating relative lidar accuracy information from overlapping flight lines. *ISPRS Journal of Photogrammetry Remote Sensing* 56, 236–245.
- Leyva, R.I., Henry, R.J., Graham, L.A., Hill, J.M., 2002. Use of LiDAR to determine vegetation vertical distribution in areas of potential black-capped vireo habitat at Fort Hood, Texas. *Endangered Species Monitoring and Management at Fort Hood, Texas, 2002 Annual Report*. The Nature Conservancy, Fort Hood, TX.
- Luethy, L., Ingensand, H., 2001. How to evaluate the quality of airborne laser scanning data. *International Archives of Photogrammetry, Remote Sensing Spatial Information Sciences XXXVI (8/W2)*, 313–317.
- Maas, H., 2002. Methods for measuring height and planimetry discrepancies in airborne laserscanner data. *Photogrammetric Engineering & Remote Sensing* 68 (9), 933–940.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering & Remote Sensing* 54, 95–104.
- Morin, K.W., 2002. Calibration of airborne laser scanners. UCGE Reports No. 20179, Department of Geomatics Engineering, University of Calgary, Calgary, Alberta, CA [URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>] (Verified 20 March 2007).
- Pfeifer, N., 2005. Airborne laser scanning strip adjustment and automation of tie surface measurement. *Boletim de Ciências Geodésicas* 11 (1), 3–22.
- Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37, 71–95.
- Schuckman, K., 2003. Announcement of the proposed ASPRS binary lidar data file format standard. *Photogrammetric Engineering & Remote Sensing* 69 (1), 13–19.
- Shan, J., Sampath, A., 2005. Urban DEM generation from raw lidar data: a labeling algorithm and its performance. *Photogrammetric Engineering & Remote Sensing* 71 (2), 217–226.
- Sithole, G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 59, 85–101.
- Skaloud, J., Lichti, D., 2006. Rigorous approach to bore-sight self-calibration in airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 61, 47–59.
- Theobald, D.M., 2003. *GIS Concepts and ArcGIS® Methods*. Conservation Planning Technologies, Fort Collins, CO.
- Vosselman, G., 2002. On the estimation of planimetric offsets in laser altimetry data. *International Archives of Photogrammetry Remote Sensing* 34 (3A), 375–380.
- Wehr, A., Lohr, U., 1999. Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry Remote Sensing* 54, 68–82.
- Willers, J.L., Milliken, G.A., Jenkins, J.N., O'Hara, C.G., Gerard, P.D., Reynolds, D.B., Boykin, D.L., Good, P.V., Hood, K.B., 2008. Defining the experimental unit for the design and analysis of site-specific experiments in commercial cotton fields. *Agricultural Systems* 96, 237–249.
- Willers, J.L., Milliken, G.A., O'Hara, C.G., Jenkins, J.N., 2004. Information technologies and the design and analysis of site-specific experiments within commercial fields. In: *Proceedings of the 16th Applied Statistics in Agriculture Conference, Manhattan, Kansas, pp. 41–73*.