

Estimating surface roughness of terrestrial laser scan data using orthogonal distance regression

Ryan M. Pollyea and Jerry P. Fairley

Department Geological Sciences, University of Idaho, Moscow, Idaho 83844-3022, USA

ABSTRACT

With the increasing accessibility of terrestrial light detection and ranging scanners (LiDAR), generating tools to elicit meaningful information from high-density point cloud data has become of paramount importance. Surface roughness is one metric that has gained popularity, largely due to the accuracy and density of LiDAR-derived point cloud data. Surface roughness is typically defined as a spread of point distances from a reference datum, the standard deviation of point distances from a model surface being a commonly employed model. Unfortunately, a recent literature review has found that existing surface roughness models are far from standardized and may be prone to error resulting from underlying surface topography. In the research presented here, we develop a surface roughness model that is robust to underlying topographic variability by segmenting the point cloud with a three-dimensional regular grid, establishing local (grid cell) reference planes by orthogonal distance regression, and estimating the surface roughness of each grid cell as the standard deviation of orthogonal point-to-plane distances. This surface roughness model is employed to identify fracture and rubble zone distributions within a terrestrial LiDAR scan from a basalt outcrop in southeast Idaho, and the results are compared to a more common model based on ordinary least-squares plane fitting. Results indicate that the orthogonal regression model is robust to outcrop orientation and that the ordinary least-squares model systematically overestimates surface roughness by contaminating estimates with spatially correlated errors that increase with decreasing grid size.

INTRODUCTION

Light detection and ranging (LiDAR) provides a powerful tool for high-resolution terrain surface mapping and is becoming increasingly accessible to the research community in the form of portable, ground-based terrestrial LiDAR scanners (Buckley et al., 2008). Surface roughness is one attribute that has gained popularity for analyzing high-resolution point cloud data, and has been used across a broad spectrum of geoscience disciplines to investigate such diverse topics as channel bed morphology (Cavalli et al., 2008), landslide morphology (Glenn et al., 2006; McKean and Roering, 2004), eolian mass transport following wildfire (Sankey et al., 2010), and alluvial fan mapping (Frankel and Dolan, 2007). Similarly, surface roughness has been measured explicitly to investigate displacement surface geometry along faults (Sagy et al., 2007), and to optimize model scaling of in situ rock joint roughness for hydromechanical analysis (Fardin et al., 2004). These studies represent only a fraction of those that use LiDAR data collection methods, and a review of the published literature indicates that this number is growing at an increasing rate.

Surface roughness is generally defined as the standard deviation of the distances of data points from a model surface datum (usually defined on the basis of local subsets of the full data set) or, in at least one case (Frankel and Dolan, 2007), of the changes in slope from a moving window across the data. Numerous methods have been

proposed for estimating surface roughness from a gridded point cloud data set using a variety of surface fitting procedures. These include digital elevation models (DEMs) (McKean and Roering, 2004; Cavalli et al., 2008), ordinary least-squares (OLS) regression planes (Fardin et al., 2004), neighborhood detrending (Davenport et al., 2004), a priori coordinate axes rotation to create an exposure-parallel principal plane (Sagy et al., 2007), and smoothing functions such as kriging (Frankel and Dolan, 2007), thin-plate splines (Glenn et al., 2006), and moving window means (Sankey et al., 2010). The common feature of these methods is that model surfaces are fit as a function of variability in one spatial coordinate, typically elevation. This implicitly assumes that the orientation of the principal plane is fixed, and that the application of a low-frequency filter over one coordinate will suffice to establish a local datum from which to measure point distances for estimating surface roughness. Although this assumption may be reasonable for the analysis of airborne LiDAR data sets or regions of negligible relief, it becomes tenuous for ground-based LiDAR applications. This is of particular concern with the increasing application of LiDAR to outcrop scanning, where one can reasonably expect outcrop orientations and local (grid scale) surfaces to deviate significantly from the principal planes. By restricting the model surface to be a function of one spatial coordinate, the chosen surface can inject underlying topographic vari-

ability into the surface roughness estimates, resulting in artificially high estimates and spatially correlated errors.

Here we present a surface roughness model based on plane fitting by orthogonal distance regression (ODR) that (1) builds on the existing definition of surface roughness as a function of point distances from a model surface; (2) effectively filters out topographic variability at the super-gridblock scale; and (3) is robust to principal axis orientation. These properties are demonstrated by applying the proposed ODR-based surface roughness model to a terrestrial LiDAR scan of a vertical basalt exposure and comparing the results to a more common model based on OLS plane fitting.

SURFACE ROUGHNESS ALGORITHM

The general workflow for the surface roughness algorithm proposed here consists of segmenting the point cloud data set using a regular three-dimensional grid, and within each grid cell: (1) establish a local (grid cell) reference plane by orthogonal distance regression, (2) compute orthogonal distances from the reference plane to each point in the grid cell, and (3) estimate surface roughness for each grid cell as the standard deviation of orthogonal point-to-plane distances. This process is performed over each grid cell in the point cloud; the mathematics of the procedure are presented in the remainder of this section.

The local (grid cell) reference datum is defined as a plane given by the standard plane equation:

$$z = \beta_0 + \beta_1 x + \beta_2 y, \quad (1)$$

where x , y , and z are spatial coordinates in \mathbb{R}^3 vector space and β_0 , β_1 , and β_2 are plane coefficients. Plane coefficients are found using orthogonal distance regression, which seeks a plane that minimizes orthogonal point-to-plane distances. Markovsky and VanHuffel (2007) provided a succinct and informative review of ODR methods (also called total least-squares regression) and provided an analytical expression for the orthogonal distance regression problem, restated here with respect to a subset (grid cell) of point cloud data in \mathbb{R}^3 vector space:

$$\boldsymbol{\beta} = (X^T X - \sigma^2 I)^{-1} X^T \mathbf{z}, \quad (2)$$

where $\boldsymbol{\beta}$ is the vector of plane coefficients, X is the design matrix, superscript T denotes the

matrix transpose, \mathbf{z} is the vector of z coordinates, σ is the smallest singular value of the augmented matrix $[X \mathbf{z}]$, and I is the \mathbb{R}^3 identity matrix.

Once the plane coefficients are determined, the orthogonal distance from each point in the grid cell to the regression plane is determined as follows: (1) translate each observation vector (data point) such that all observation vectors in the grid cell share a common origin in the regression plane, and (2) compute the scalar projection of the translated observation vectors onto the plane normal vector. Here we choose the z intercept of the regression plane (β_0) as the new vector origin as it is known explicitly. By scaling the plane normal vector, \mathbf{n} , to unit length, each orthogonal point-to-plane distance is the positive inner product of the observation vector and the unit normal vector:

$$D_i = |\mathbf{n}_u \cdot \mathbf{v}_i|, i = 1, \dots, N, \quad (3)$$

where D_i is an orthogonal point-to-plane distance within the grid cell, \mathbf{n}_u is the unit normal vector to the local regression plane, and \mathbf{v}_i is an observation vector (data point) with its origin at the z intercept. Once all orthogonal point-to-plane distances have been evaluated within a grid cell, the surface roughness (θ) is modeled as the standard deviation of these distances:

$$\theta = \sqrt{\frac{1}{N} \sum_{i=1}^N (D_i - \bar{D})^2}. \quad (4)$$

APPLICATION TO AN OUTCROP DATA SET

The proposed ODR-based surface roughness algorithm was applied to a terrestrial LiDAR scan of a vertical basalt exposure at the Box Canyon in southeast Idaho (Fig. 1). The scanned exposure consists of two adjacent basalt flows and portions of several additional flows. The outcrop measures ~90 m horizontal by 10 m vertical, exhibits significant fracturing, and has a rubbly flow margin near the center of the exposure (Fig. 1). LiDAR scanning was performed in June 2010, using a Leica ScanStation 2. Individual scans ($n = 14$) were acquired and merged with a mean registration error of 0.004 m. In total, ~65 million data points were obtained. After post-processing to remove vegetation, the data were exported such that exposure orientation was approximately parallel to the xz plane.

A FORTRAN 2003 program was written to execute the ODR-based surface roughness algorithm, and a similar program was used for modeling surface roughness based on fitting grid cell planes by OLS regression. In a statistical sense, OLS regression assumes that all deviation from the regression model is contained in one variable, the response or dependent variable, and that the remaining (independent) variables are



Figure 1. Box Canyon, Idaho, basalt exposure used for this investigation. Exposure is ~90 m horizontal by 10 m vertical. Note that this image is not orthorectified.

error free. The OLS-based roughness algorithm therefore fits a grid cell regression plane by minimizing point-to-plane distances in one coordinate direction, which, for this analysis, is the direction of the exposure-normal principal axis. The OLS algorithm then evaluates surface roughness as the standard deviation of these point-to-plane distances. The OLS surface roughness algorithm therefore performs similar to surface roughness models that fit a DEM to the point cloud, then estimate surface roughness from the statistics of point elevations above the model surface.

The ODR and OLS algorithms were executed for the entire Box Canyon data set using four successively smaller grid intervals (1 m, 0.5 m, 0.25 m, and 0.125 m), and the summary statistics are presented in Table 1. In evaluating the summary statistics, it becomes evident that the OLS-based algorithm systematically returns greater surface roughness and the discrepancy between the ODR and OLS methods increases with each grid refinement. Fracture distribution maps of the Box Canyon exposure were generated for each method on the 0.125 m grid using binary transform plots (0.05 cutoff) for normalized surface roughness (Fig. 2). The lack of detail present in the OLS map results from

systematic overestimation of surface roughness in regions of high topographic variability (e.g., surfaces with significant departure from the underlying principal plane). This effect is analyzed in the following.

DISCUSSION

The larger roughness values returned by the OLS algorithm in areas of greater topographic variability arise because OLS assumes a baseline surface common to all subsets of the data. This datum is assumed to exist without error, and the error (i.e., the distances from the best-fit plane to the data point) is assigned wholly to the dependent axis. In the Box Canyon outcrop, for example, the baseline surface was chosen a priori to be the principal coordinate plane approximately parallel with the trend of the basalt exposure, which, in this instance, is the xz plane. Subsequent OLS plane fitting was then performed for each grid cell by minimizing the point-to-plane distances with respect to the exposure-normal principal axis (y). In contrast, the ODR-based roughness algorithm calculates each best-fit plane independent of the principal axes orientations by fitting a regression plane that minimizes orthogonal point-to-plane distances in each grid cell. Using minimum orthogonal point-to-plane

TABLE 1. SURFACE ROUGHNESS SUMMARY STATISTICS

	Total gridblocks (N)	Average points per grid cell	Minimum*	Maximum*	Mean*	Variance†	Standard Deviation*
0.125 m grid							
ODR surface roughness	53508	1222	0.0000	0.1202	0.0069	4.32×10^{-5}	0.0066
OLS surface roughness			0.0000	0.6486	0.0184	0.0018	0.0427
0.25 m grid							
ODR surface roughness	13755	4752	0.0000	0.1324	0.0146	1.57×10^{-4}	0.0125
OLS surface roughness			0.0000	0.5920	0.0285	0.0023	0.0481
0.5 m grid							
ODR surface roughness	3565	18334	1.00×10^{-8}	0.1963	0.0332	5.50×10^{-4}	0.0235
OLS surface roughness			7.80×10^{-7}	0.4651	0.0478	0.0030	0.0545
1.0 m grid							
ODR surface roughness	948	68947	6.69×10^{-4}	0.3245	0.0696	0.0016	0.0600
OLS surface roughness			0.0017	0.4482	0.0793	0.0036	0.0406

Note: ODR—orthogonal distance regression; OLS—ordinary least squares.

*Surface roughness units are given in meters.

†Units in square meters.

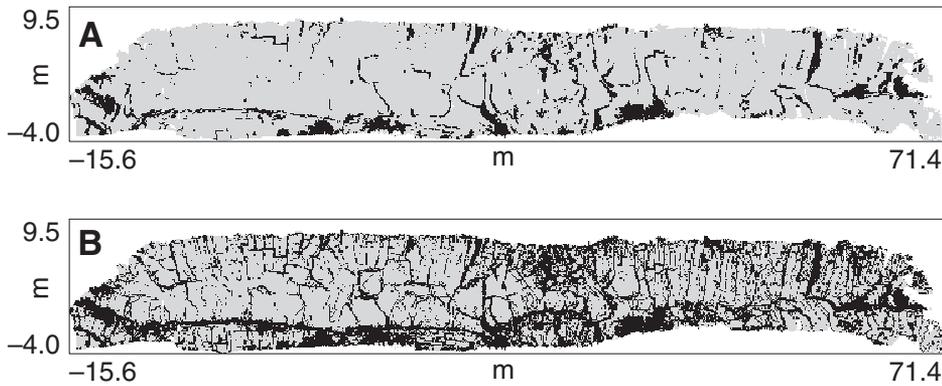


Figure 2. Binary transform maps of normalized surface roughness estimates from Box Canyon LiDAR (light detection and ranging) data evaluated on 0.125 m grid. Values ≥ 0.05 are denoted in black and values < 0.05 are denoted in gray. Distances are in meters. A: Surface roughness model generated using the ordinary least-squares plane-fitting algorithm. B: Surface roughness model generated using orthogonal distance regression plane-fitting algorithm.

distances forces the ODR algorithm to return the smallest possible roughness values for any procedure that models surface roughness as a function of point-to-plane distances. It is therefore intuitive that, for grid cells in regions where surface expressions deviate significantly from the exposure-parallel principal plane, the OLS method will result in overall greater point-to-plane distances (i.e., greater roughness values) than the ODR method (Table 1).

Systematic overestimation of OLS-based surface roughness in regions of high topographic variability is of concern because this effectively suppresses fine-scale detail by decreasing the influence of smaller roughness intensities relative to the entire distribution. This effect is illustrated by plotting histograms of normalized (over their respective ranges) OLS and ODR surface roughness distributions from the 0.125 m grid (Fig. 3), where $\sim 86\%$ of the OLS roughness estimates are contained in smallest roughness bin (0–0.05) (Fig. 3A). This is

in stark contrast to the ODR histogram, which holds $\sim 63\%$ of the normalized roughness values in the smallest roughness bin (Fig. 3B). In other words, by normalizing each roughness data set over its range and comparing the OLS and ODR roughness distributions, it becomes evident that the artificially high OLS-based estimates decrease the signal-to-noise ratio, severely affecting the overall roughness distribution by masking the influence of low to moderate roughness intensities. This effect can be seen in binary transform maps of normalized surface roughness at the 0.05 threshold (Fig. 2). As shown in Figure 2A, the OLS method accurately reproduces coarse features (wide aperture fractures and rubby basalt flow margins), but cannot render fine-scale intraflow fractures. In contrast, for the same normalized roughness threshold the ODR method, being more sensitive to local topographic deviation, is capable of reproducing coarse features as well as the finer details which, in this case, are the column-

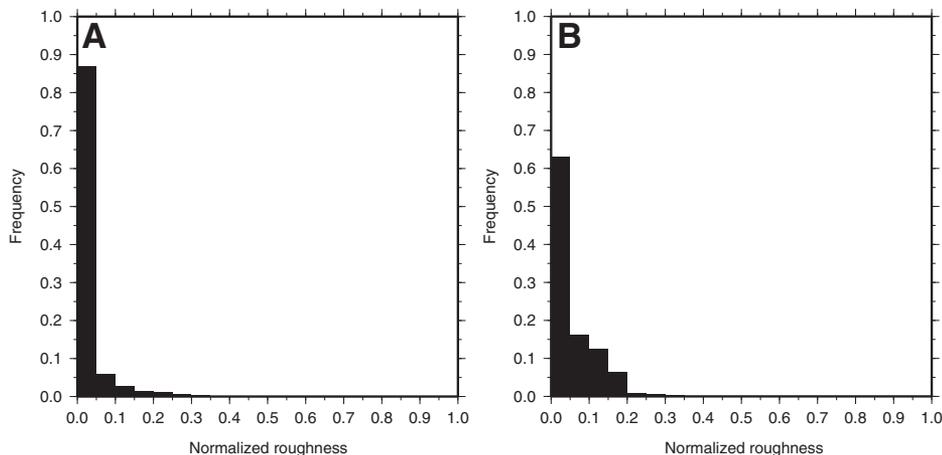


Figure 3. Histograms of normalized surface roughness for the Box Canyon LiDAR (light detection and ranging) data segmented on 0.125 m grid. A: Histogram of normalized surface roughness for ordinary least-squares-based method. B: Histogram of normalized surface roughness for orthogonal distance regression-based method.

normal and column-bounding fractures within the flow interior (Fig. 2B).

Because the ODR algorithm minimizes point-to-plane distances with respect to all three coordinate axes, the ODR surface roughness measurements represent minimum error estimates for comparison with the OLS-based surface roughness measurements. The OLS model error, E_{OLS} , can therefore be defined as the absolute value of the difference between the roughness estimates:

$$E_{OLS} = |\theta_{OLS} - \theta_{ODR}|, \quad (5)$$

where θ is the roughness estimate for a given grid cell (developed by OLS or ODR methods as indicated by the subscript).

Semivariogram analyses were performed for the Box Canyon ODR and OLS surface roughness estimates, as well as the E_{OLS} , for grid discretizations of 1 m, 0.5 m, 0.25 m, and 0.125 m (Fig. 4). Semivariograms for the 1 m grid (Fig. 4A) show that the OLS (triangles), ODR (circles), and E_{OLS} (plus symbols) are in general agreement, indicating that errors due to the OLS method are negligible for large grid cells. In fact, the grid cell dimension at which this general agreement of the OLS and ODR semivariograms occurs may be considered a characteristic length scale of the underlying topography that the OLS method is unable to filter. As the grid size is successively reduced for subsequent grid refinements, the OLS and E_{OLS} semivariograms begin deviating from the ODR semivariogram, and this deviation becomes increasingly pronounced with additional refinement (Figs. 4B–4D). Similarly, as the OLS and E_{OLS} semivariograms deviate further from the ODR semivariogram, their spatial dependence increases (lower semivariogram values indicate stronger spatial dependence), resulting in artificially suppressed uncertainty, i.e., the OLS semivariogram suggests a lower nugget effect (implying greater certainty) than the ODR semivariogram. In addition, with each grid refinement the OLS and E_{OLS} semivariograms become closer to unity, indicating that for highly refined grids, error is a dominant contribution to the OLS surface roughness estimates and that these errors become increasingly correlated with each grid refinement. This spatial correlation of the OLS errors arises as a result of the residual topographic variability that the OLS plane-fitting algorithm is unable to filter. The presence of spatially correlated errors implies that the OLS method of estimating roughness will propagate underlying topographic variability of the surface in the guise of roughness into any subsequent analysis that makes use of the surface roughness data.

The propagation of spatially correlated errors in OLS-based surface roughness estimates may be of negligible import for situations where there is little underlying topographic variability,

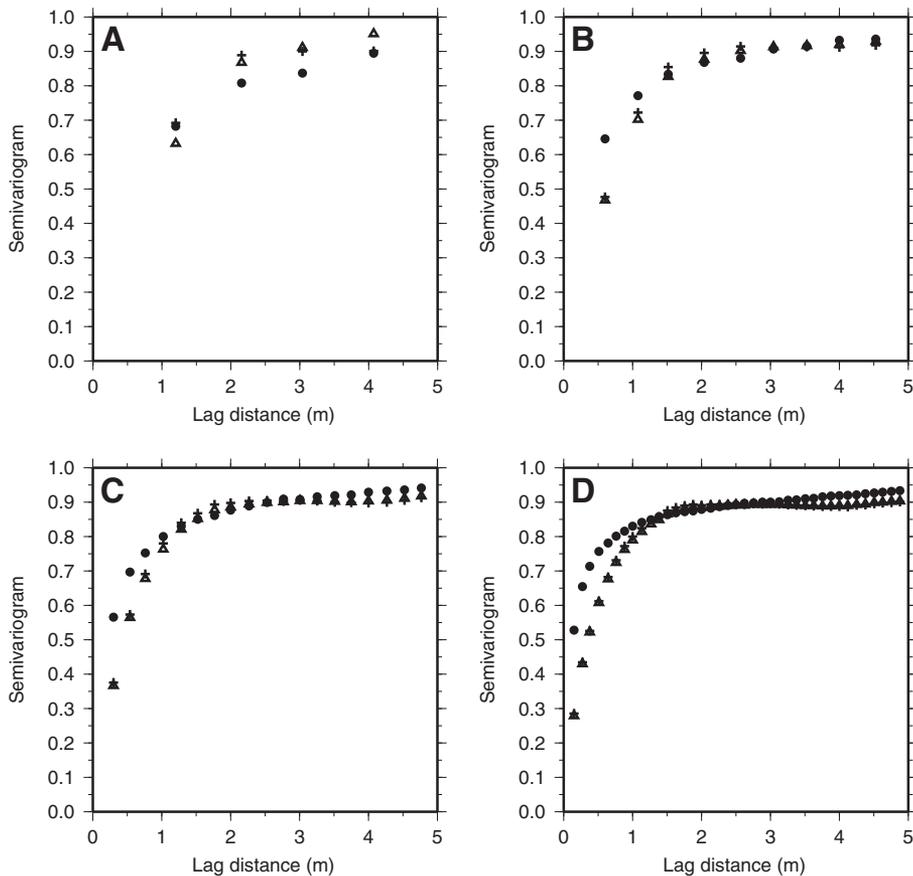


Figure 4. Normalized experimental semivariograms of orthogonal distance regression-based surface roughness (circles), ordinary least-squares (OLS)-based surface roughness (triangles), and OLS-based surface roughness errors (plus symbols). A: For 1 m grid. B: For 0.5 m grid. C: For 0.25 m grid. D: For 0.125 m grid.

or the super-gridblock variability can be filtered out before making roughness estimates. However, most real-world applications of LiDAR technology involve the scanning of surfaces that are of variable orientation (e.g., outcrops that include changes in slope, contain prominences or depressions). To date, these situations have been approached in an ad hoc manner using various OLS-type roughness algorithms that fit reference datum as a function of one spatial coordinate and/or constrain reference datum such that adjacent planes share a common edge; this latter constraint is a notable DEM feature representing the primary difference between DEM reference datum and the OLS regression plane used in this analysis. By imposing these constraints on the grid cell reference datum (best-fit planes) from which surface roughness is modeled, the subsequent roughness estimates are contaminated with super-gridblock-scale topographic variability containing spatially correlated errors. The case of an amphitheater-shaped outcrop, where the exposure wraps around an observation point, is of special interest because the ODR-based estimates of roughness are invariant to the orientation of the scanned face. As a result, using the ODR algorithm allows all data

to be analyzed simultaneously, without detrending or piecemeal rotations of the outcrop principle axes, while avoiding the outcrop orientation errors to which OLS-type methods are prone.

CONCLUSION

In this paper we have sought to present a general tool for estimating surface roughness that builds on previous efforts and is robust to exposure orientation, while minimizing the effects of underlying surface topography. We have shown that orthogonal distance regression and orthogonal point-to-plane distance measurements satisfy these requirements and avoid the propagation of spatially correlated errors that are associated with OLS-type surface roughness models commonly in use. The method of ODR presented here is somewhat more computationally expensive than OLS methods; however, the greater versatility and defensibility of the ODR method provide a standard against which future LiDAR-based investigations involving surface roughness estimates may be measured. In addition, applying the ODR-based surface roughness models to point cloud data from a scanned outcrop is demonstrated to be an effective tool for mapping discontinuities on a scanned out-

crop surface. This technique may have broad appeal for practitioners seeking data to support discrete fracture network models, training images for multiple-point geostatistics in fractured systems, and fracture density maps for geomechanical investigations.

ACKNOWLEDGMENTS

We thank J. Hinds for her invaluable assistance in the field. We also thank T. Candela, M. Wilkinson, and one anonymous reviewer for their insightful comments and suggestions. This work was supported by the Center for Advanced Energy Studies under contract DE-AC07-05ID14517.

REFERENCES CITED

- Buckley, S.J., Howell, J., Enge, H., and Kurz, T., 2008, Terrestrial laser scanning in geology: Data acquisition, processing and accuracy consideration: *Geological Society of London Journal*, v. 165, p. 625–638, doi:10.1144/0016-76492007-100.
- Cavalli, M., Tarolli, P., Marchi, L., and Dalla Fontana, G., 2008, The effectiveness of airborne LiDAR data in the recognition of channel-bed morphology: *Catena*, v. 73, p. 249–260, doi:10.1016/j.catena.2007.11.001.
- Davenport, I.J., Holden, N., and Gurney, R.J., 2004, Characterizing errors in airborne laser altimetry data to extract soil roughness: *IEEE Transactions on Geoscience and Remote Sensing*, v. 42, no. 10, p. 2130–2141, doi:10.1109/TGRS.2004.834648.
- Fardin, N., Feng, Q., and Stephansson, O., 2004, Application of a new in situ 3D laser scanner to study the scale effect on the rock joint surface roughness: *International Journal of Rock Mechanics and Mining Sciences*, v. 41, p. 329–335, doi:10.1016/S1365-1609(03)00111-4.
- Frankel, K.L., and Dolan, J.F., 2007, Characterizing arid region alluvial fan surface roughness with air-borne laser swath mapping digital topographic data: *Journal of Geophysical Research*, v. 112, F02025, doi:10.1029/2006JF000644.
- Glenn, N.F., Strecker, D.R., Chadwick, J.D., Thackray, G.D., and Dorsch, S.J., 2006, Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity: *Geomorphology*, v. 73, p. 131–148, doi:10.1016/j.geomorph.2005.07.006.
- Markovsky, L., and VanHuffel, S., 2007, Overview of total least-squares methods: *Signal Processing*, v. 87, p. 2283–2302, doi:10.1016/j.sigpro.2007.04.004.
- McKean, J., and Roering, J., 2004, Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry: *Geomorphology*, v. 57, p. 331–351, doi:10.1016/S0169-555X(03)00164-8.
- Sagy, A., Brodsky, E.E., and Axen, G.J., 2007, Evolution of fault-surface roughness with slip: *Geology*, v. 35, p. 286–286, doi:10.1130/G23235A.1.
- Sankey, J.B., Glenn, N.F., Germino, M.J., Gironella, A.I., and Thackray, G.D., 2010, Relationships of aeolian erosion and deposition with LiDAR-derived landscape surface roughness following wildfire: *Geomorphology*, v. 119, p. 135–145, doi:10.1016/j.geomorph.2010.03.013.

Manuscript received 14 January 2011
 Revised manuscript received 15 February 2011
 Manuscript accepted 18 February 2011

Printed in USA