# APPLICATION OF GROUND-BASED LIDAR FOR GULLY INVESTIGATION IN AGRICULTURAL LANDSCAPES

**Henrique G. Momm, Ronald L. Bingner, Robert R. Wells, and Seth M. Dabney**
United States Department of Agriculture
Agriculture Research Service
National Sedimentation Laboratory
598 McElroy Drive
Oxford, MS 38655
Henrique.Momm@ars.usda.gov
Ron.Bingner@ars.usda.gov
Robert.Wells@ars.usda.gov
Seth.Dabney@ars.usda.gov

## ABSTRACT

Detailed scientific investigation of gullies in agricultural fields requires accurate topographic information with adequate temporal and spatial resolution. New technologies, such as ground-based LIDAR systems, are capable of generating datasets with high temporal and spatial resolutions. The spatial resolution is dependent on the ground point sampling density, which is a result of operator controlled factors such as the area of data collection (scan angle), average point density of scans, and degree of overlap between scans. The selection of the appropriate point density sampling is especially important in research sites where the same location needs to be surveyed multiple times over lengthy periods of time as conditions change due to precipitation and runoff events, field management changes, and/or implementation of different conservation practices. This study investigated the relationship between point sampling density and topographic information through the use of variograms. A Monte Carlo-type experiment was performed by interactively and randomly reducing points to created reduced point sets and comparing theoretical variograms of reduced datasets to theoretical variogram of the original dataset. Results indicated point sampling density thresholds for different relief variation datasets (low, moderate, and high) that produced little or no additional topographic information when exceeded. Although variations in local relief can lead to different point sampling density requirements, the outcome of this study serves as guidance for future field surveys of gully evolution and erosion.

**KEYWORDS:** ephemeral gully, ground-based LIDAR, erosion, point sampling density, variograms

## INTRODUCTION

In agricultural fields, gully erosion is significant and often times is comparable to or exceeds sheet and rill erosion volumes (Bernard, et al. 2010). A large number of modeling tools have been developed over the years to estimate sediment transport from agricultural fields to streams and lakes (Wischmeier and Smith, 1978; Knisel, 1980; Renard et al., 1991; and Bingner et al., 2010). However, at the current stage of development, algorithms to estimate sediment from gullies in models are either limited or do not exist. Gullies in agriculture fields are often classified into either ephemeral or classical gullies.

Ephemeral gullies are defined as small channels located in agricultural fields eroded primarily from concentrated overland flow that can be easily filled by normal tillage, only to reform again in the same location by additional runoff events (Soil Science Society of America, 2001). Due to their small dimensions, producers often reshape the channel's topography, by re-filling the channel through tillage, to maintain regular farming operations (Quadros et al., 2004). Because the field topography is often unchanged between seasons, ephemeral gullies have a tendency to re-form at the same or nearby location (Casali et al., 1999). An ephemeral gully becomes permanent, referred to as classic gully, in situations where a headcut migrates upstream faster than the time interval between farmers tilling operations, followed by widening of the channel and which forces producers to operate around the gully channel.

Capturing this dynamic behavior poses a challenge to efforts of understanding and estimation of the erosional contribution of ephemeral and classical gullies in agricultural watersheds (Casali et al., 1999). In order to capture the microtopography impacting ephemeral gully formation, DEMs with spatial resolution ranging between 5mm to 5cm have been shown to be necessary (Schmid et al., 2004).

Recent developments in laser scanner technology, provide new opportunities for scientific investigation of ephemeral and classical gullies. These systems are capable of collecting information with a wide range of ground point densities as a result of operator controlled factors such as the selection of data collection (scan angles) area, average point density of scans, and degree of overlap between scans (Figure 1). Higher point density can be achieved by higher sensor resolution, smaller vertical field of view angles, and multiple scans of the same ground location. The instrument resolution is often controlled by an imaginary plane located at the middle range of the vertical field of view and is orthogonal to the sensor's normal sight. The vertical field of view angles also influence the point density as point sampling becomes sparser as vertical scan angle tends to the horizon (Figure 1). Multiple scans can be used to collect data over the same geographical location resulting in increased sampling density and avoiding problems such as shadowing and limited coverage due to vegetation.

The identification of the proper sampling density is especially important in research sites where the same location needs to be surveyed multiple times over lengthy periods of time as conditions change due to precipitation and runoff events, field management changes, and/or implementation of different conservation practices. However, as the number of data points increase (point density) with increased resolution-rich surveys, so does the time required for field data collection, post processing, and size of datasets. This scenario is aggravated when monitoring multiple sites with different topographic and farming practices. The objective of this study is the identification of the minimum sampling point density, which provides the necessary topographic information at individual agricultural gully scale and yields field surveys that are reproducible and efficient. The results of this study may be used as a guideline in using laser scanner technology to characterize topographic conditions associated with the evolution of ephemeral and classical gullies and, more importantly, accurate estimates of quantities of sediment eroded or conserved by erosion control practices.

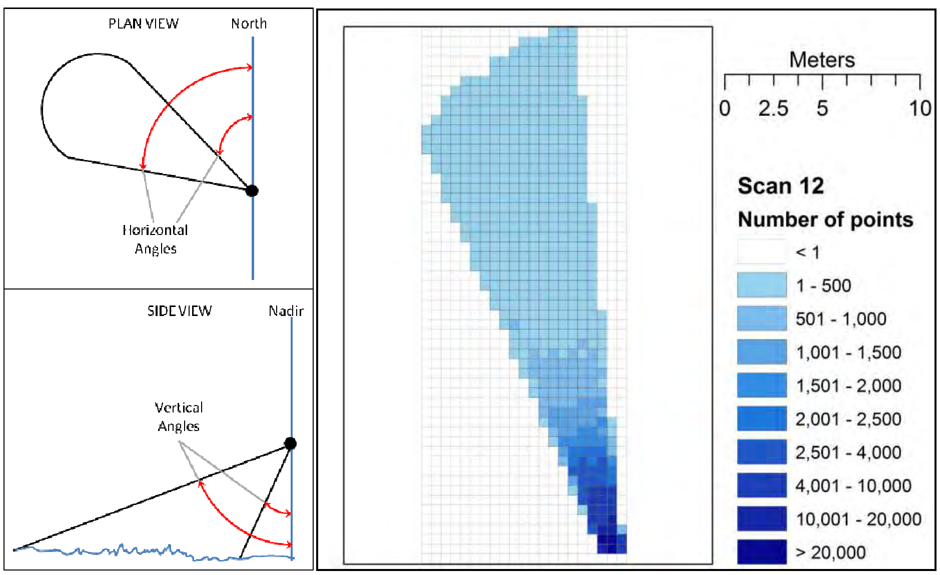


**Figure 1.** Map of spatial variation of ground point sampling density, represented as number of points per grid unit, with vertical scanning angle of ground-based laser scanners.

# METHODOLOGY

## Study Site

The study site selected for this investigation is located within the Cheney Lake Reservoir watershed near the town of Hutchinson in South Central Kansas. The predominant land use is agriculture (>73%) in the form of cropland and rangeland. The gully within the study site was 96 meters long oriented North-South, approximately 1.3 meters wide and from 10 to 50 cm deep. The channel is free of vegetation and crop residues, while the surrounding field is covered by crop residues resulting from no-till management used in winter wheat followed by sorghum (milo) in the 2010 crop rotation. Historical cultivation practice indicates that initially this ephemeral gully did not disrupt farming operations; however, as no-tillage practices were adopted in 2005, the channel grew wider and deeper to the point that the farming equipment could not be used to travel across the gully and the ensuing cropping activity was performed around the main channel (Figure 2) (Frees et al, 2009).

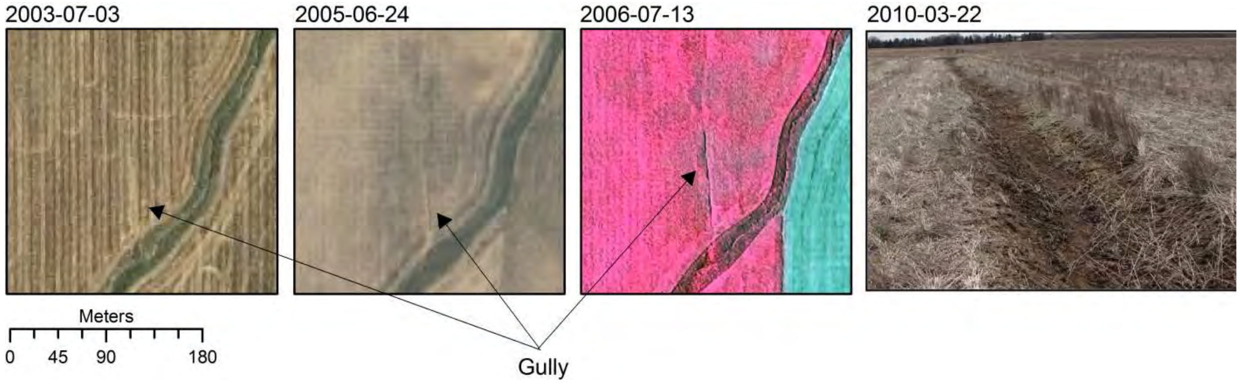


**Figure 2**. Ephemeral gully evolution into classical gully and its consequent disruption to producer's operations. Imagery data for years 2003, 2005, and 2006 obtained from the National Agriculture Imagery Program (NAIP) and 2010 from field visit.

Two locations with known geographic coordinates within the study site were used to provide reference geographical coordinates. This is an important step to translate the equipment local coordinates into geographic coordinates, thus providing a means to compare surveys performed at different times. Initially, the operator scans the pre-defined targets installed at the reference points. Based on the known geometry of the targets, the instrument is then capable of calculating its location in relation to the reference points (geographical coordinates). Four standard targets were installed in the far outmost corners of the gully being investigated. These four static targets are surveyed and their coordinates computed and recorded. Each subsequent scan starts with surveying the four targets to locate the laser scanner in the local coordinate system. A total of eleven scans in eleven set ups (one scan for each equipment set up) were used to describe the gully. In the post processing steps, each scan with local coordinates are translated into geographical coordinates using the relation between the four targets and the reference points. During point collection verifying the point sampling coverage is very important to avoid gaps in the survey.

## Sampling Density Investigation

Studies have been performed to identify the optimum balance between point density at small gully scales and volume of data with the goal of optimizing data collection and cost. Guo (2010) provides a detailed investigation of the relationship associated with airborne LIDAR point density reduction, interpolation methods, and resolution of digital elevation models. There have not been studies that investigate the relationship of laser point sampling density collected using ground-based laser scanners on topographic information tailored to gully investigations in agricultural fields. Ground-based systems differ from airborne systems as a result of encountering a wider range of incidence angles (Soudarissanane, et al., 2009). Also, the finer resolution in investigations of gullies formed in agricultural fields requires the determination of micro-topography, and the presence of crop residues produces various levels of terrain roughness, posing a challenge to interpolation techniques.

Point sampling density was investigated by tiling the entire LIDAR point cloud into one meter square grids. Sampling density was computed by counting the number of points in each tile. This information can be utilized when verifying spatial coverage of sampled points to identify gaps or under-sampled regions. Areas with specific features, such as gully headcuts, should be scanned with higher point density, whereas featureless areas can be scanned at lower point density. An area with high point density designed to detail the gully active headcut as accurately as possible was obtained (Figure 3). In contrast, there is an under-sampled region in the mid-section of the gully.

A total of 5,032 tiles were generated (many of them containing no points) (Figure 3). Tiles were ranked by standard deviation of elevation values and divided in three groups based on the data quartile values. In each group, the tiles with the highest number of points were selected, 2175 (155,373 points with $\sigma_{elev}$=0.13), 2177 (40,923 points with $\sigma_{elev}$=0.06), and 2304 (17,144 points with $\sigma_{elev}$=0.01). The same variation in elevation represented by standard deviation values can be observed on histograms (Figure 4). Tile 2175 has the largest elevation range (~ 50 cm) as the gully active headcut is located in this tile. Histogram plots of the distance values to the nearest neighbor depict point density of each plot. The vast majority of the points are within 5 mm of other points.

Investigation of the relationship between sampling density and topographic information was performed through the use of variograms of the laser-produced point clouds to minimize the uncertainty introduced when interpolating irregularly spaced points into regularly spaced grid. The variogram explores spatial independence (Kitanidis, 1997) and quantitatively relates variance to space separation (Curran and Atkison, 1998). Experimental semivariograms for each of the three tiles selected were computed using the algorithm *gamv* available in the Geostatistical Software LIBrary (GSLIB) due to the irregularly spaced nature of the laser points (Deutsch and Journel, 1998). The lag separation distance (distance between two points used to create the point pair database) and lag tolerance were selected to be 2 cm and 1 cm respectively (2cm±1cm). The omnidirectional variogram was considered throughout our investigation.

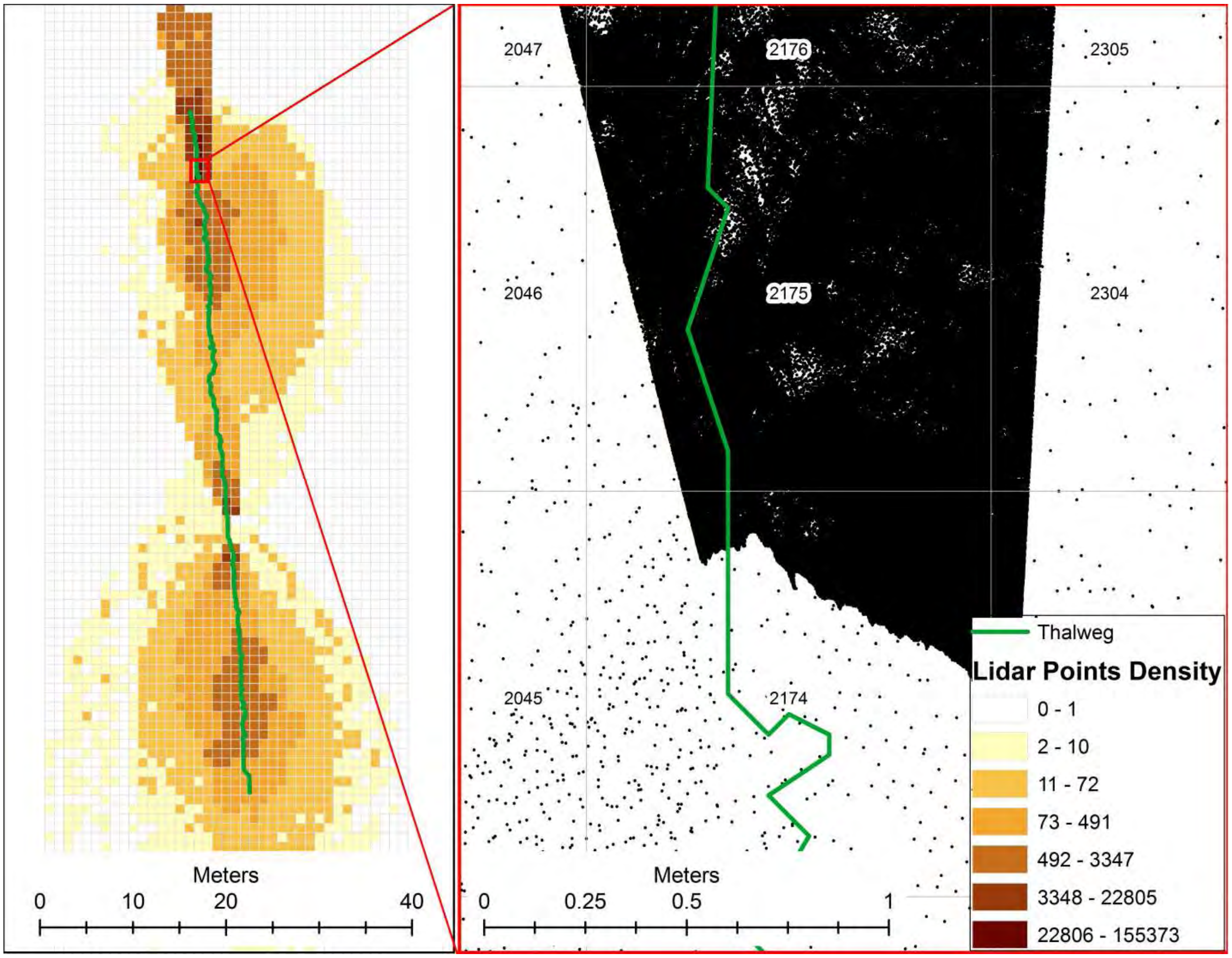


**Figure 3.** Ground-based lidar point density investigation. Left hand-side map shows the spatial variation of the sampling density in square meter cell size grid. The right hand-side map illustrates the differences in point density between tiles.

The experimental semivariograms were computed using all the available laser points in each tile with the *gamv* algorithm. The theoretical semivariograms were generated using the Levenberg-Marquardt optimization algorithm (Levenberg, 1944) for determining the set of parameters that provides the best fit to the experimental variogram through minimization of the sum of squares of the residuals. Different mathematical models were selected to represent the theoretical semivariogram curves. For tile 2175 a composite Gaussian model was used and variations of the standard Gaussian model were used to generate theoretical semivariogram curves for the remaining two tiles, 2304 and 2177. These theoretical semivariograms were considered as reference in the subsequent spatial continuity experiment.

A Monte Carlo type investigation was performed by creating a series of independent simulations of reduced datasets containing a smaller number of laser points than the original number in each tile. The reduced dataset was generated by randomly selecting laser points based on a pre-defined percentage. A percentage of 100% represents all the laser points available in the tile while a reduced set using a percentage of 50% would yield half of the available points in the tile. For each pre-defined percentage, a total of 100 independent realizations were performed (100 independent randomly selected reduced sets). Each reduced set was used in the computation of experimental and theoretical semivariogram curves. Multiple percentage threshold values were used (10 for tiles 2175 and 2304 and 18 for tile 2177). Smaller percentage threshold values introduce higher levels of uncertainty represented by the increased variability of the curves (Figure 5).

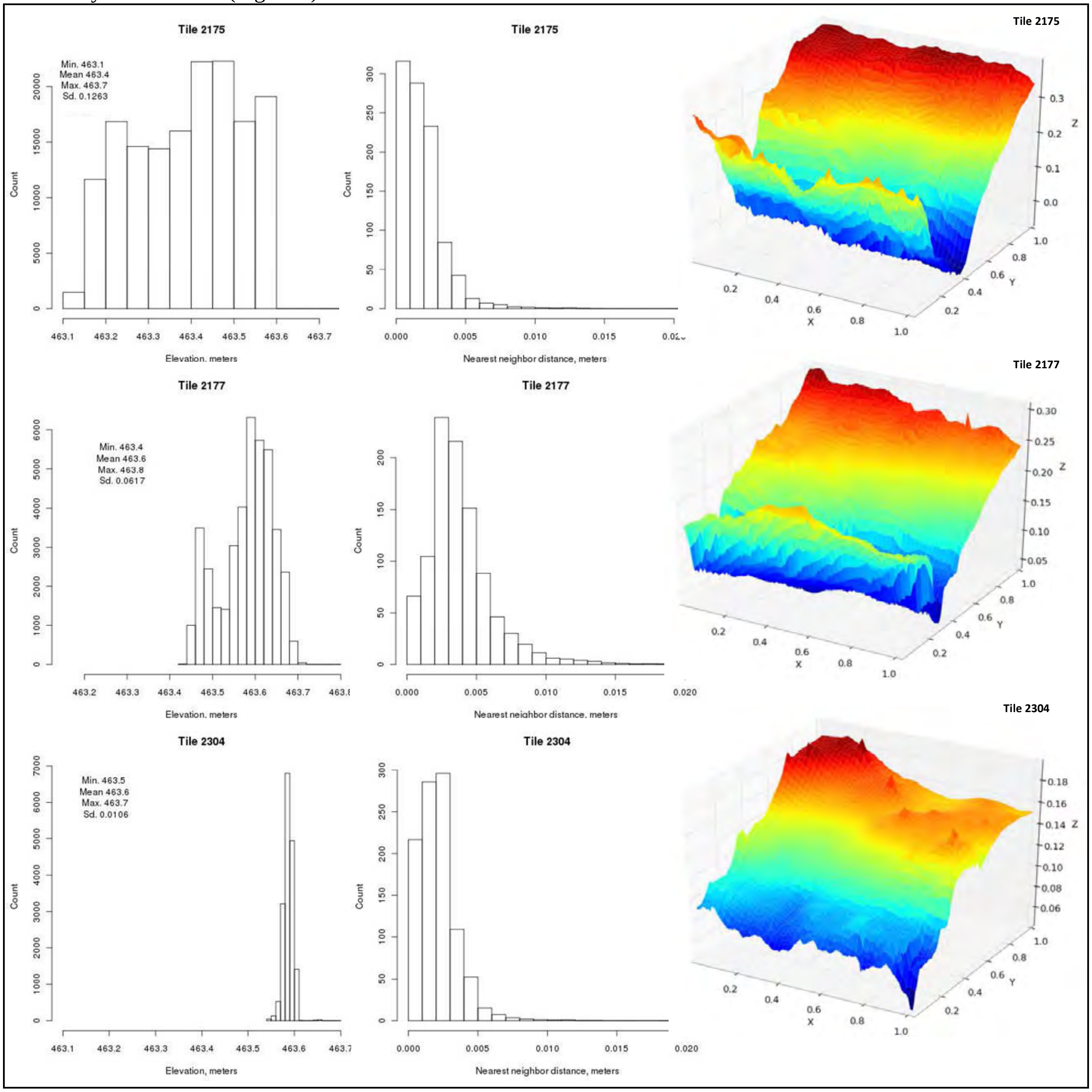


**Figure 4**. Three tiles selected for the point sampling density investigation. Left hand-side column presents the histograms of elevation values, center column the distance to the nearest neighbor in each of the square meter tile considered, and the right hand-side column three-dimensional grids of each tile.

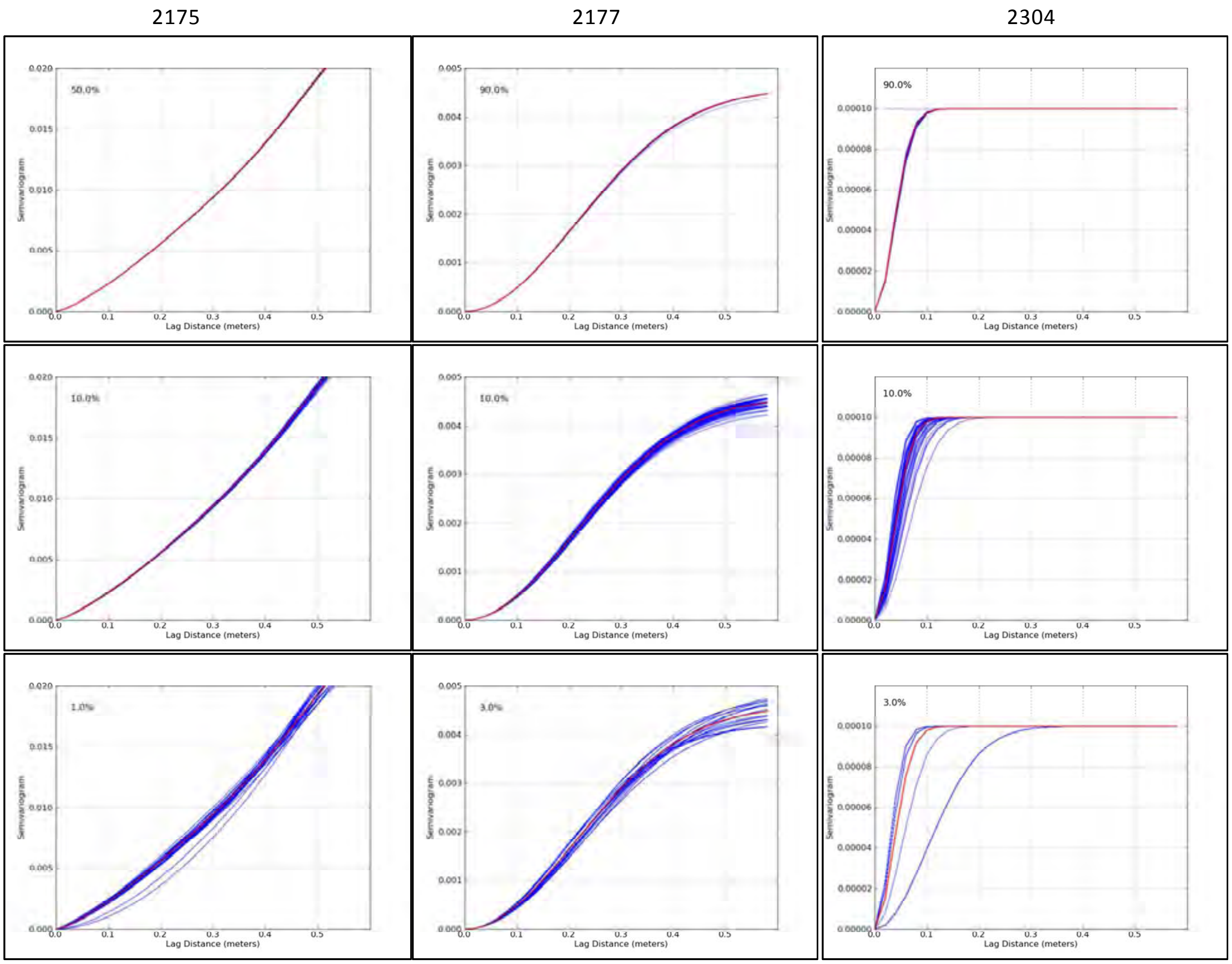


**Figure 5.** Comparison of theoretical variograms generated using all available data points (red) and reduced dataset (blue). Percentage values indicate the amount of points randomly selected for variance analysis. A total of 100 realizations were performed for each percentage of points considered.

## DISCUSSION OF RESULTS

The theoretical semivariogram curves generated with reduced data points were quantitatively evaluated by individual comparison to the theoretical semivariogram curve, obtained using all collected laser points, through the calculation of root mean squared deviation (RMSD) as shown in Equation 1.

$$RMSD_{(V_{100},V_P)} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{100,i} - X_{P,i}\right)^2}{n}} \quad \textbf{[1]}$$

In this equation, $V_{100}$ represents the theoretical semivariogram curve developed using all available laser points, *n* is the total number of points in the curve (total number of lag intervals considered), $V_P$ represents theoretical semivariogram curves generated using a reduced dataset with percentage *P*. A total of 100 RMSD values for each percentage threshold were calculated and averaged. The resulting set of averaged RMSD values are graphically displayed in Figure 6 for each tile.

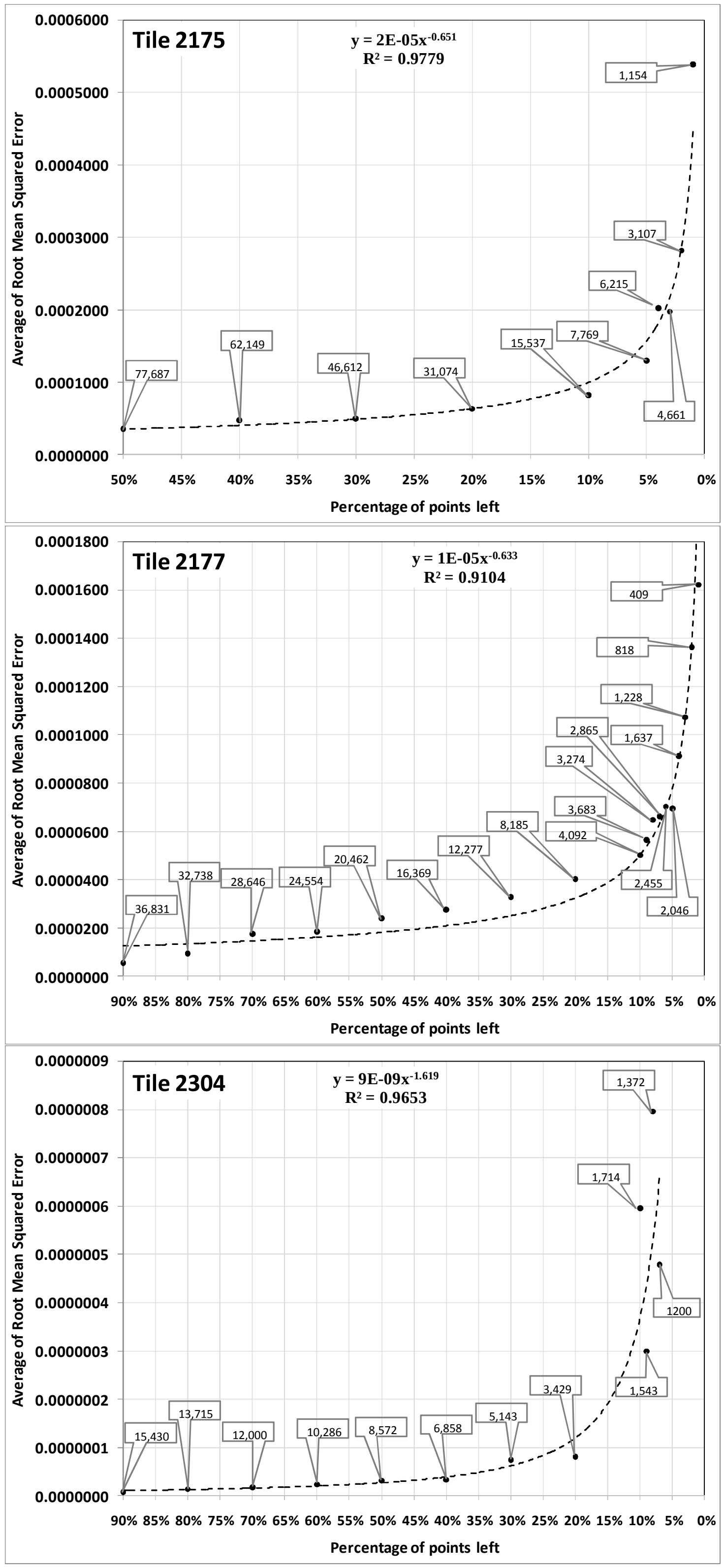


**Figure 6.** Comparison of the goodness of fit between theoretical variogram using 100% of the available laser points in each tile and 100 realizations of reduced sets of points. Reduced number of points increases uncertainty of theoretical variograms.

The three curves display similar shape with the largest discontinuities found in the plot for tile 2304. Points representing the percentage of 10% and 8% yielded higher averaged RMSD values than the point with the lowest number of points (7%). This can be partially explained by the procedure from which a reduced set was created. A standard random sampling technique was used, therefore, it is possible that selected points were not spatially distributed throughout the tile (forming clusters) and as result, the theoretical variogram curve differs from the reference, which yielded large RMSD values. Just a few realizations of clustered points could significantly increase the average value. Nonetheless, despite these two discontinuities it is possible to identify a general trend. The curves start with a gentle slope and as the number of points becomes smaller, curves tend towards to increase vertically. In other words, results indicate that, in the scale considered, there is an upper threshold of point density where topographic information provided by the LIDAR point cloud does not increase (or increases very little) despite the increased point sampling density. Additionally, a positive relationship between this minimum number of points and the tile standard deviation of elevation can be observed, as higher sampling densities are needed to topographically describe locations with higher relief, as expected.

To further evaluate the effect of point sampling on topographic information, these curves were used to select three sampling density threshold values to reduce the remaining tiles in the survey, 7,500, 4,000, and 3,500 from tile 2175, 2177, and 2304 respectively. Using these values, the number of points in a tile was reduced to the threshold of 7,500/$m^2$ if the standard deviation of elevation was $\geq$ 0.03617, to 4,000/$m^2$ if the standard deviation of elevation was < 0.03617 and $\geq$ 0.0106, and to 3,500/$m^2$ if the standard deviation of elevation was $\leq$ 0.0106. A total of 25 tiles were reduced.

The two point clouds, original and reduced, were converted to Triangular Irregular Network (TIN) format to facilitate volume computations. A third TIN, with artificially filled channel, was created by manually digitizing the edges of the gully channel to form a polygon and then subsequent removal of all the laser points within the channel polygon. Through the use of differencing technique, the original and reduced TINs were subtracted from the TIN with artificially filled channel yielding volumes estimate of 18.154 $m^3$ and 18.146 $m^3$ respectively. There is a difference of less than 0.04% between the two estimates. Additionally, visual comparison of the thalweg profiles for both datasets confirms the agreement between the original and reduced dataset (Figure 5). In multi-temporal research efforts, it is important to obtain accurate horizontal and vertical characterization of the gully's thalweg in order to precisely characterize gully changes over time leading to improved understanding of gully evolution.

## CONCLUSION

This study used the semi-variogram concept to quantitatively investigate the relationship between LIDAR point sampling density and topographic modeling needed to evaluate ephemeral and classic gullies in agricultural fields. The impact of gullies in agricultural fields can be studied at different scales, such as watershed scale, field scale, and individual gully scales. In this study, we addressed effects of point sampling density on the topographic information at the individual gully scale. The gully investigated was partitioned into square meter tiles and the sampling density of each tile was computed by counting the number of laser points in each tile. This experiment revealed a large variation in the LIDAR point sampling density throughout the gully. Tiles were ranked by standard deviation of elevation values and partitioned into three groups based on quartile values. The tile with the highest number of points in each group was selected for the sensitivity analysis. Multiple realizations of subsets of randomly selected points at pre-defined percentages were used to identify the minimum point sampling density in which the data set retains the original spatial characteristic. Using the minimum number of points per square meter thresholds, a reduced point cloud dataset was developed and compared to the original dataset yielding not significant discrepancy. This indicates that data could be collected with smaller sampling density while retaining the original spatial characteristics.

This is important because surveys that collect less data result in less time needed in the field to characterize ephemeral and classical gully evolution. Although results indicated that the reduced dataset did not significantly differ from the original dataset in terms of topographic information, and thus these tiles could be considered over-sampled, the reduced tiles represent only a small percentage of the entire dataset. Out of 2,085 tiles containing laser points, only 25 were reduced because they originally had more laser points than the defined thresholds. And, out of the 25 reduced tiles only 14 were located in and around the gully channel. Despite the oversampling of 14 tiles in and around the gully channel, still there are 175 tiles (out of 189) located in and around the gully channel that presented fewer points per square meter than the defined thresholds, as consequence of the large variation in sampling density.

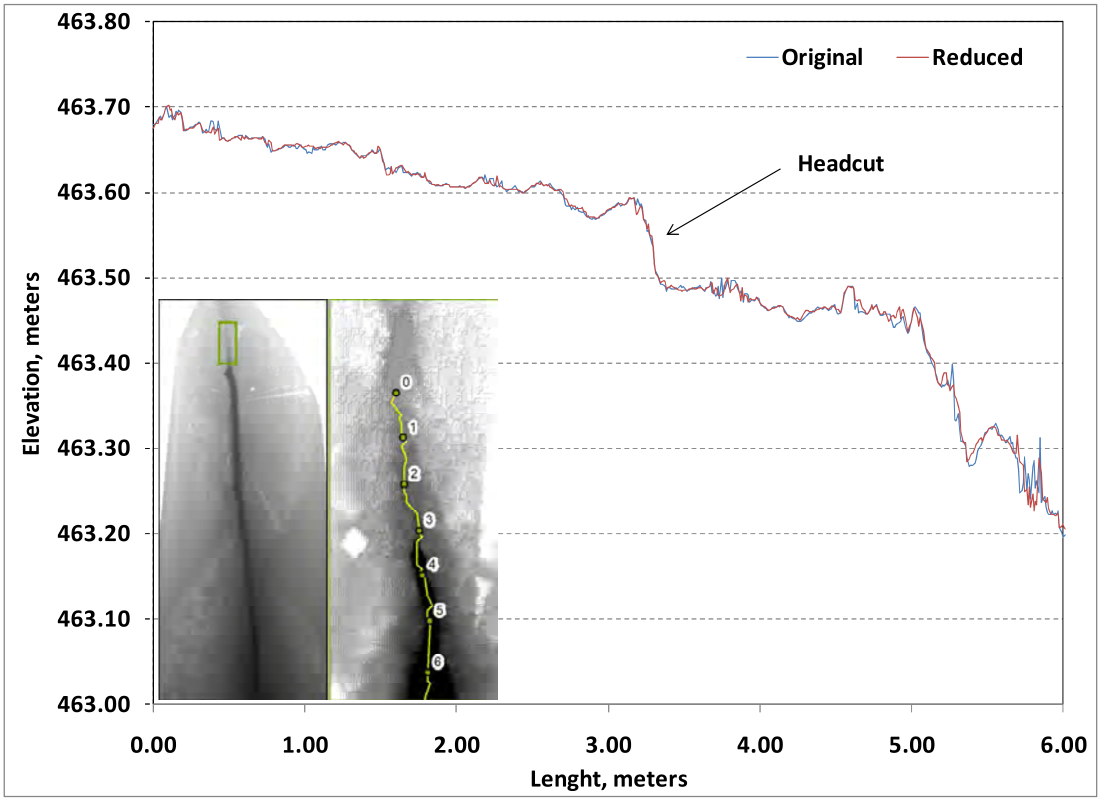


**Figure 5.** Thalweg topographic profile of no till gully collected using ground-based LIDAR. Two profiles are compared: blue from a TIN considering all the points collected in the survey and red from a TIN created using a reduced dataset.

Although the ideal situation would be to survey gullies with the highest possible sampling density, this is often not practical because sampling density varies with factors such as resolution of the instrument, vertical scan angle, number of overlapping scans, and land coverage. Furthermore, scientific investigation to quantify and to understand the development of ephemeral and classic gullies in agricultural fields over time often requires multi-temporal surveys of multiple locations throughout the watershed.

Based on the findings of this study, a more consistent distribution of sampling density should be the ultimate goal. During the field collection the laser scanner is mounted on a tripod that can be elevated allowing the possibility of collecting data far away from the nadir situation (large vertical angles). Collection of data with such large vertical angles leads to lower sampling densities and shadowing when investigating gullies with deep channels. One possible alternative would be to survey the same location using multiple overlapping scans with lower point density. Although the instrument would be set to collect data at a lower point density, the combined set of scans would yield higher point density. Additionally, the overlapping dataset could be used to evaluate the point cloud by identifying pairs of points with high elevation difference what could be a potential cue to remove anomalies from the data cloud. The objective should be to design a targeted surveying plan where different minimum sampling density threshold values should used based on local conditions such as surface roughness and vegetation cover, such as the findings found herein relating point sampling density and surface roughness.

The use of ground-based LIDAR for ephemeral and classical gully investigations in agricultural fields is relatively new and research in this field is expected to continue to grow as technology becomes less expensive and new applications are developed. The use of such technology can help in collecting detailed micro-topography information that can be used in many different research areas such as ephemeral and classical gully modeling, soil water depression storage capacity, terrain roughness measurements, and many others. Additionally, the use of ground-based laser scanners is a competitive alternative to traditional methods such as total station and photogrammetry due to the increased accuracy and reduced field time required.

## REFERENCES


Bernard, J., Bingner, R. L., Dabney, S. M., Langendoen, E. J., Lemunyon, J., Merkel, W., Theurer, F., Wells, R. R., Widman, N. and Wilson, G. V., 2010. Ephemeral Gully Erosion – A Natural Resource Concern. USDA-ARS National Sedimentation Laboratory Research Report. No. 69. US Department of Agriculture, Agricultural Research Service, National Sedimentation Laboratory, Oxford, MS.

Bingner, R. L., R. Wells, H. G. Momm, F. D. Theurer, and L. D. Frees, 2010. Development and Application of Gully Erosion Components within the USDA AnnAGNPS Watershed Model for Precision Conservation, *10th International Conference on Precision Agriculture*, Denver, Colorado, USA.

Casali, J., J. Lopez, and J. Giráldez, 1999. Ephemeral gully erosion in southern Navarra (Spain), *Catena*, *36:* 65-84.

Curran, P. J., and P. M. Atkinson, 1998. Geostatistics and Remote Sensing. *Progress in Physical Geography* 22:61–78.

Deutsch, C. V. and A. G. Journel, 1998. *GSLIB: Geostatistical software library and user's guide*, New York, 369.

Frees, Lyle, Lisa French, and Fred Theurer. (2009). Conservation Effects Assessment Program: Special Emphasis Watersheds, Cheney Lake Watershed, Kansas. 17 March 2009. USDA, NRCS, Kansas. 50 pp.

Guo Q., W. Li, H. Yu, and O. Alvarez, 2010. Effects of Topographic Variability and LIDAR Sampling Density on Several DEM Interpolation Methods. *Photogrammetric Engineering and Remote Sensing*, 76(6) 701-712.

Knisel, W.G., ed. 1980. CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems. U.S. Dept of Agric., Conservation Research Report No. 26, 640 pp.

Kitanidis, P. K., 1997. *Introduction to Geostatistics: Application in Hydrogeology*, Cambridge University Press, Cambridge, England.

Levenberg, K., 1944. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *The Quarterly of Applied Mathematics* 2: 164–168.

Quadros, R. M. B., C. Wisneiwski, and E. Passos, E., 2004. Modelagem de canais incisivos – revisao, RA'E GA ISSN:1516-4136, 8, 69-81. (Portuguese)

Renard, K. G., G. R. Foster, G. A. Weesies, and J. P. Porter. 1991. RUSLE: revised universal soil loss equation. *Journal of Soil and Water Conservation*, 46(1): 30-33.

Soil Science Society of America, 2001. Glossary of Soil Terms. Soil Science Society of America, Madison, USA, pp: 134.

Schmid, T., H. Schack-Kirchner, and E. Hildebrand, 2004. A case study of terrestrial laser-scanning in erosion research: calculation of roughness indices and volume balance at a logged forest site. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(8):114-118.

Soudarissanane, S., R. Lindenbergh, M. Menenti, P. Teunissen, and T. Delft, 2009, Incidence angle influence on the quality of terrestrial laser scanning points , ISPRS, Proceedings ISPRS Workshop Laser scanning, 1-2 Sept, Paris, France.

Wischmeier, W.H. and D.D. Smith. 1978. Predicting rainfall erosion losses. Agriculture Handbook 537. USDA. Agricultural Research Service, Washington, DC. 58 pp.